N. Elkins
B. Jouris
Inside Products
K. Haining
U. S. Bank
M. Ackermann
BCBS Michigan

Intended Status: Proposed Standard
Expires: July 2014
January 27, 2014

IPv6 Performance and Diagnostic Metrics Destination Option
draft-elkins-6man-ipv6-pdm-dest-option-05

Abstract

To diagnose performance and connectivity problems, metrics on real (non-synthetic) transmission are critical for timely end-to-end problem resolution. Such diagnostics may be real-time or after the fact, but must not impact an operational production network. The base metrics are: packet sequence number and packet timestamp. Metrics derived from these will be described separately. This document solves these problems with a new destination option, the Performance and Diagnostic Metrics destination option (PDM).

Status of this Memo

Copyright and License Notice

Table of Contents

1 Introduction

To diagnose performance and connectivity problems, metrics on real (non-synthetic) transmissions are critical for timely end-to-end problem resolution. Such diagnostics may be real-time or after the fact, but must not impact an operational production network. The base metrics are: packet sequence number and packet timestamp.

For background, please see draft-ackermann-ntp-pdm-ntp-usage-00 [NTPPDM], draft-elkins-v6ops-ipv6-packet-sequence-needed-01 [ELKPSN], draft-elkins-v6ops-ipv6-pdm-recommended-usage-01 [ELKUSE], draft-elkins-v6ops-ipv6-end-to-end-rt-needed-01 [ELKRSP] and draft-elkins-ippm-pdm-metrics-03 [ELKIPPM]. These drafts are companions to this document.

As discussed in the above Internet Drafts, current methods are inadequate for these purposes because they assume unreasonable access to intermediate devices, are cost prohibitive, require infeasible changes to a running production network, and/or do not provide timely data.   This document provides a solution for these problems.

As defined in RFC2460 [RFC2460], destination options are carried by the IPv6 Destination Options extension header. Destination options include optional information that need be examined only by the IPv6 node given as the destination address in the IPv6 header, not by routers or other "middle boxes". This document specifies a new destination option, the Performance and Diagnostic Metrics destination option (PDM).

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2 Performance and Diagnostic Metrics Destination Options

2.1 Destination Options Header

The IPv6 Destination Options Header is used to carry optional information that need be examined only by a packet's destination node(s). The Destination Options Header is identified by a Next Header value of 60 in the immediately preceding header and is defined in RFC2460 [RFC2460].

2.2 PDM Types

The IPv6 Performance and Diagnostic Metrics Destination Option (PDM) is an implementation of the Destination Options Header (Next Header value = 60). Two types of PDM are defined. PDM type 1 requires time synchronization. PDM type 2 does not require time synchronization.

PDM type 1 and PDM type 2 are mutually exclusive. That is, a 5-tuple can either both send PDM type 1 or both send PDM type 2.

2.3 Performance and Diagnostic Metrics Destination Option (Type 1)

PDM type 1 is used to facilitate diagnostics by including a packet sequence number and timestamp.

The PDM type 1 is encoded in type-length-value (TLV) format as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Option Type  | Option Length | PSN This Packet               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                         TimeStamp This Packet (64-bit)        +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   PSN Last Packet             |   Reserved                    |
|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                         TimeStamp Last Packet (64-bit)        +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Option Type

TBD = 0xXX (TBD) [To be assigned by IANA] [RFC2780]

Option Length

8-bit unsigned integer. Length of the option, in octets, excluding the Option Type and Option Length fields. This field MUST be set to 22.

Packet Sequence Number This Packet (PSNTP)

16-bit unsigned integer. This field will wrap. It is intended for human use.

Initialized at a random number and monotonically incremented for packet on the 5-tuple. The 5-tuple consists of the source and destination IP addresses, the source and destination ports, and the upper layer protocol (ex. TCP, ICMP, etc).

Operating systems MUST implement a separate packet sequence number counter per 5-tuple. Operating systems MUST NOT implement a single counter for all connections.

Note: This is consistent with the current implementation of the IPID field in IPv4 for many, but not all, stacks.

TimeStamp This Packet (TSTP)

A 64-bit unsigned integer field containing a timestamp that this packet was sent by the source node. The value indicates the number of seconds since January 1, 1970, 00:00 UTC, by using a fixed point format. In this format, the integer number of seconds is contained in the first 32 bits of the field, and the remaining 32 bits resolve to picoseconds.

This follows timestamp formats used in Network Time Protocol (NTP) [RFC5905] and SEND [RFC3971]. A discussion of why NTP is used in preference to Precision Time Protocol (PTP) is in draft-elkins-v6ops-ipv6-end-to-end-rt-needed-01 [ELKRSP]. A discussion of how to implement NTP for use with PDM header type 1 is in draft-ackermann-ntp-pdm-ntp-usage-00 [NTPPDM].

Implementation note: This format is compatible with the usual representation of time under UNIX, although the number of bits available for the integer and fraction parts in different Unix implementations vary.

Packet Sequence Number Last Received (PSNLR)

16-bit unsigned integer. This is the PSN of the packet last received

on the 5-tuple.

TimeStamp Last Received (TSLR)

A 64-bit unsigned integer field containing a timestamp. This is the timestamp of the packet last received on the 5-tuple. Format is the same as TSTP.

2.4 Performance and Diagnostic Metrics Destination Option (Type 2)

The second type of IPv6 Performance and Diagnostic Metrics Destination Option (PDM) is as follows. PDM type 1 and PDM type 2 are mutually exclusive. That is, a 5-tuple can either both send PDM type 1 or both send PDM type 2.

PDM type 2 contains the following fields:

```
PSNTP    : Packet Sequence Number This Packet
PSNLR    : Packet Sequence Number Last Received
DELTALR  : Delta Last Received
PSNLS    : Packet Sequence Number Last Sent
DELTALS  : Delta Last Sent
```

PDM destination option type 2 is encoded in type-length-value (TLV) format as follows:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Option Type  | Option Length | PSN This Packet               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   PSN Last Received           |  PSN Last Sent                |
|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Delta Last Received         |  Delta Last Sent              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| TType |
+-+-+-+-+
```

Option Type

TBD = 0xXX (TBD) [To be assigned by IANA] [RFC2780]

Option Length

8-bit unsigned integer. Length of the option, in octets, excluding the Option Type and Option Length fields. This field MUST be set to 22.

Packet Sequence Number This Packet (PSNTP)

16-bit unsigned integer. This field will wrap. It is intended for human use.

Initialized at a random number and monotonically incremented for packet on the 5-tuple. The 5-tuple consists of the source and destination IP addresses, the source and destination ports, and the upper layer protocol (ex. TCP, ICMP, etc).

Operating systems MUST implement a separate packet sequence number counter per 5-tuple. Operating systems MUST NOT implement a single counter for all connections.

Note: This is consistent with the current implementation of the IPID field in IPv4 for many, but not all, stacks.

Packet Sequence Number Last Received (PSNLR)

16-bit unsigned integer. This is the PSN of the packet last received on the 5-tuple.

Packet Sequence Number Last Sent (PSNLS)

16-bit unsigned integer. This is the PSN of the packet last sent on the 5-tuple.

Delta TimeStamp Type (TIMETYPE)

4-bit unsigned integer. This is the type of time contained in the delta fields below.

0 - unknown 1 - time is in units of nanoseconds 2 - time is in units of microseconds 3 - time is in units of milliseconds 4 - time is in units of seconds 5 - time is in units of minutes 6 - time is in units of hours 7 - time is in units of days

The values 5 - 7 are relevant for Delay Tolerant Networks (DTN) which may operate with long delays between packets.

Delta Last Received (DELTALR)

A 16-bit unsigned integer field. This is server delay.

DELTALR = Send time packet 2 - Receive time packet 1

The value is according to the scale in TIMETYPE.

Delta Last Sent (DELTALS)

A 16-bit unsigned integer field. This is round trip or end-to-end time.

Delta Last Sent = Receive time packet 2 - Send time packet 1

The value is in according to the scale in TIMETYPE.

Option Type

The two highest-order bits of the Option Type field are encoded to indicate specific processing of the option; for the PDM destination option, these two bits MUST be set to 00. This indicates the following processing requirements:

00 - skip over this option and continue processing the header.

RFC2460 [RFC2460] defines other values for the Option Type field. These MUST NOT be used in the PDM. The other values are as follows:

01 - discard the packet.

10 - discard the packet and, regardless of whether or not the packet's Destination Address was a multicast address, send an ICMP Parameter Problem, Code 2, message to the packet's Source Address, pointing to the unrecognized Option Type.

11 - discard the packet and, only if the packet's Destination Address was not a multicast address, send an ICMP Parameter Problem, Code 2, message to the packet's Source Address, pointing to the unrecognized Option Type.

In keeping with RFC2460 [RFC2460], the third-highest-order bit of the Option Type specifies whether or not the Option Data of that option can change en-route to the packet's final destination.

In the PDM, the value of the third-highest-order bit MUST be 0. The possible values are as follows:

0 - Option Data does not change en-route

1 - Option Data may change en-route

The three high-order bits described above are to be treated as part of the Option Type, not independent of the Option Type. That is, a particular option is identified by a full 8-bit Option Type, not just the low-order 5 bits of an Option Type.

2.5 Header Placement

The PDM destination option MUST be placed as follows:

- Before the upper-layer header. That is, this is the last extension header.

This follows the order defined in RFC2460 [RFC2460]

IPv6 header

Hop-by-Hop Options header

Destination Options header

Routing header

Fragment header

Authentication header

Encapsulating Security Payload header

Destination Options header

upper-layer header

For each IPv6 packet header, the PDM MUST NOT appear more than once. However, an encapsulated packet MAY contain a separate PDM associated with each encapsulated IPv6 header.

The inclusion of a PDM in a packet affects the receiving node's processing of only this single packet. No state is created or modified in the receiving node as a result of receiving a PDM in a packet.

2.6 Implementation Considerations

The PDM destination options extension header SHOULD be turned on by each stack on a host node.

2.6.1 Dynamic Configuration Options

If implemented, each operating system MUST have a default configuration parameter, e.g. diag_header_sys_default_value=yes/no. The operating system MAY also have a dynamic configuration option to change the configuration setting as needed.

If the PDM destination options extension header is used, then it MAY be turned on for all packets flowing through the host, applied to an upper-layer protocol (TCP, UDP, SCTP, etc), a local port, or IP address only. These are at the discretion of the implementation.

The PDM MUST NOT be changed dynamically via packet flow as this may create potential security violation or DoS attack by numerous packets turning the header on and off.

As with all other destination options extension headers, the PDM is for destination nodes only. As specified above, intermediate devices MUST neither set nor modify this field.

2.6.2 Data Length Filtering

Different results for derived metrics, such as, server delay, will be obtained if calculations are done including or excluding packets which have a data length of 0 or 1. Some protocols, for example, TCP, provide acknowledgements which have a length of 0. Keep-alive packets have a data length of 0 or 1.

Operating systems may provide the user a choice of whether to include or exclude packets with a 0 or 1 byte data length.

2.6.3 5-tuple Aging

Within the operating system, metrics must be kept on a 5-tuple basis. As discussed before, these are:

```
PSNTP : Packet Sequence Number This Packet
TSTP  : Timestamp This Packet
PSNLR : Packet Sequence Number Last Received
TSLR  : Timestamp Last Received
PROTC : Protocol for Upper Layer (ex. TCP, UDP, ICMP, etc)
```

The question comes of when to stop keeping data or restarting the

numbering for a 5-tuple. For example, in the case of TCP, at some point, the connection will terminate. Keeping data in control blocks forever, will have unfortunate consequences for the operating system.

So, the recommendation is to use a known aging parameter such as Max Segment Lifetime (MSL) as defined in Transmission Control Protocol [RFC0793]. The choice of aging parameter is left up to the implementation.

3 Backward Compatibility

The scheme proposed in this document is backward compatible with all the currently defined IPv6 extension headers. According to RFC2460 [RFC2460], if the destination node does not recognize this option, it should skip over this option and continue processing the header.

4 Security Considerations

The PDM MUST NOT be changed dynamically via packet flow as this creates a possibility for potential security violations or DoS attacks by numerous packets turning the header on and off.

5 IANA Considerations

An option type must be assigned by IANA for the Performance and Diagnostic Metrics destination option.

6 References

6.1 Normative References

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2780] Bradner, S. and V. Paxson, "IANA Allocation Guidelines For Values In the Internet Protocol and Related Headers", BCP 37, RFC 2780, March 2000.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.

6.2 Informative References

[NTPPDM] Ackermann, M., "draft-ackermann-ntp-pdm-ntp-usage-00", Internet Draft, January 2014.

[ELKPSN] Elkins, N., "draft-elkins-v6ops-ipv6-packet-sequence-needed-01", Internet Draft, September 2013.

[ELKRSP] Elkins, N., "draft-elkins-v6ops-ipv6-end-to-end-rt-needed-01", Internet Draft, September 2013.

[ELKUSE] Elkins, N., "draft-elkins-v6ops-ipv6-pdm-recommended-usage-01", Internet Draft, September 2013

[ELKIPPM] Elkins, N., "draft-elkins-ippm-pdm-metrics-03", Internet Draft, January 2014.

7 Acknowledgments

The authors would like to thank Sigfrido Perdomo and Fred Baker for their comments.

Authors' Addresses

Nalini Elkins
Inside Products, Inc.
36A Upper Circle
Carmel Valley, CA 93924
United States
Phone: +1 831 659 8360
Email: nalini.elkins@insidethestack.com
http://www.insidethestack.com

William Jouris
Inside Products, Inc.
36A Upper Circle
Carmel Valley, CA 93924
United States
Phone: +1 925 855 9512
Email: bill.jouris@insidethestack.com
http://www.insidethestack.com

Michael S. Ackermann
Blue Cross Blue Shield of Michigan
P.O. Box 2888
Detroit, Michigan 48231
United States
Phone: +1 310 460 4080
Email: mackermann@bcbsmi.com
http://www.bcbsmi.com

Keven Haining
US Bank
16900 W Capitol Drive
Brookfield, WI 53005
United States
Phone: +1 262 790 3551
Email: keven.haining@usbank.com
http://www.usbank.com