Workgroup: ippm
Internet-Draft:
draft-ietf-ippm-ioam-ipv6-options-10
Published: 28 February 2023
Intended Status: Standards Track
Expires: 1 September 2023
Authors: S. Bhandari, Ed. F. Brockners, Ed.
Thoughtspot Cisco

In-situ OAM IPv6 Options

Abstract

In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network. This document outlines how IOAM data fields are encapsulated in IPv6.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network. IOAM concepts and associated nomenclature, as well as IOAM data fields are defined in [RFC9197]. IOAM is a functionality that applies to limited domains [RFC8799]. This document outlines how IOAM data fields are encapsulated in IPv6 [RFC8200] and discusses deployment requirements for networks that use IPv6-encapsulated IOAM data fields.

2. Conventions

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Abbreviations

Abbreviations used in this document:

E2E: Edge-to-Edge

IOAM: In-situ Operations, Administration, and Maintenance as defined in [RFC9197]

OAM: Operations, Administration, and Maintenance

POT: Proof of Transit

3. In-situ OAM Metadata Transport in IPv6

IOAM in IPv6 is used to enhance diagnostics of IPv6 networks. It complements other mechanisms designed to enhance diagnostics of IPv6 networks, such as the IPv6 Performance and Diagnostic Metrics Destination Option described in [RFC8250].

At the time this document was written, several implementations of IOAM for IPv6 exist, e.g., IOAM for IPv6 in the Linux Kernel (supported from Kernel version 5.15 onwards IPv6 IOAM in Linux Kernel), IOAM for IPv6 in VPP .

IOAM data fields can be encapsulated with two types of extension headers in IPv6 packets - either the hop-by-hop options header or the destination options header. Multiple options with the same option type MAY appear in the same hop-by-hop options or destination options header, with distinct content.

An IPv6 packet carrying IOAM data in an extension header can have other extension headers, compliant with [RFC8200].

IPv6 hop-by-hop and destination option format for carrying IOAM data fields:

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Option Type  |  Opt Data Len |   Reserved    |   IOAM Type   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<-+
|                                                               |  |
.                                                               .  I
.                                                               .  O
.                                                               .  A
.                                                               .  M
.                                                               .  .
.                          Option Data                          .  O
.                                                               .  P
.                                                               .  T
.                                                               .  I
.                                                               .  O
.                                                               .  N
|                                                               |  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<-+
```

Option Type: 8-bit option type identifier as defined in Section 6.

Opt Data Len: 8-bit unsigned integer. Length of this option, in octets, not including the first 2 octets.

Reserved: 8-bit field MUST be set to zero by the source.

IOAM Type: 8-bit field as defined in section 7.1 in [RFC9197].

Option Data: Variable-length field. Option-Type-specific data.

IOAM Option data is inserted as follows:

1. Pre-allocated Trace Option: The IOAM Preallocated Trace Option-Type defined in Section 4.4 of [RFC9197] is represented as an IPv6 option in the hop-by-hop extension header:

   Option Type: TBD_1_1 8-bit identifier of the IPv6 Option Type for IOAM.

   IOAM Type: IOAM Pre-allocated Trace Option-Type.

2. Proof of Transit Option: The IOAM POT Option-Type defined in Section 4.5 of [RFC9197] is represented as an IPv6 option in the hop-by-hop extension header:

   Option Type: TBD_1_1 8-bit identifier of the IPv6 Option Type for IOAM.

   IOAM Type: IOAM POT Option-Type.

3. Edge to Edge Option: The IOAM E2E option defined in Section 4.6 [RFC9197] is represented as an IPv6 option in destination extension header:

   Option Type: TBD_1_0 8-bit identifier of the IPv6 Option Type for IOAM.

   IOAM Type: IOAM E2E Option-Type.

4. Direct Export (DEX) Option: The IOAM Direct Export Option-Type defined in Section 3.2 of [I-D.ietf-ippm-ioam-direct-export] is represented as an IPv6 option in the hop-by-hop extension header:

   Option Type: TBD_1_0 8-bit identifier of the IPv6 Option Type for IOAM.

   IOAM Type: IOAM Direct Export (DEX) Option-Type.

All the IOAM IPv6 options defined here have alignment requirements. Specifically, they all require 4n alignment. This ensures that fields specified in [RFC9197] are aligned at a multiple-of-4 offset from the start of the hop-by-hop and destination options header.

IPv6 options can have a maximum length of 255 octets. Consequently, the total length of IOAM Option-Types including all data fields is also limited to 255 octets when encapsulated into IPv6.

4. IOAM Deployment In IPv6 Networks

4.1. Considerations for IOAM deployment and implementation in IPv6 networks

IOAM deployments in IPv6 networks MUST take the following considerations and requirements into account:

C1 IOAM MUST be deployed as a limited domain feature as defined in [RFC8799].

C2 Implementations of IOAM MUST ensure that the addition of IOAM data fields does not alter the way routers forward packets or the forwarding decisions they make. Packets with added IOAM information must follow the same path within the domain as an identical packet without IOAM information would, even in the presence of Equal-Cost Multi-Path (ECMP). This behavior is important for deployments where IOAM data fields are only added "on-demand". Implementations of IOAM MUST ensure that ECMP behavior for packets with and without IOAM data fields is the same. In order for IOAM to work in IPv6 networks, IOAM MUST be explicitly enabled per interface on every node within the IOAM domain. Unless a particular interface is explicitly enabled (i.e., explicitly configured) for IOAM, a router MUST ignore IOAM Options.

C3 In order to maintain the integrity of packets in an IOAM domain, the Maximum Transmission Unit (MTU) of transit routers and switches must be configured to a value that does not lead to an ICMP Packet Too Big error message being sent to the originator and the packet being dropped. The PMTU tolerance range must be identified and IOAM encapsulation operations or data field insertion must not exceed this range. Control of the MTU is critical to the proper operation of IOAM and is one of the reasons it is considered a limited domain feature (see also [RFC8799]). The PMTU tolerance must be identified through configuration and IOAM operations must not exceed the packet size beyond PMTU.

C4 [RFC8200] precludes insertion of IOAM data directly into the original IPv6 header of in-flight packets. IOAM deployments which do not encapsulate/decapsulate IOAM on the host but desire to encapsulate/decapsulate IOAM on transit nodes MUST add an additional IPv6 header to the original packet. IOAM data is added to this additional IPv6 header.

4.2. IOAM domains bounded by hosts

For deployments where the IOAM domain is bounded by hosts, hosts will perform the operation of IOAM data field encapsulation and decapsulation. IOAM data is carried in IPv6 packets as hop-by-hop or destination options as specified in this document.

4.3. IOAM domains bounded by network devices

For deployments where the IOAM domain is bounded by network devices, network devices such as routers form the edge of an IOAM domain. Network devices will perform the operation of IOAM data field encapsulation and decapsulation.

5. Security Considerations

This document describes the encapsulation of IOAM data fields in IPv6. Security considerations of the specific IOAM data fields for each case (i.e., Trace, Proof of Transit, and E2E) are described and defined in [RFC9197].

As this document describes new options for IPv6, these are similar to the security considerations of [RFC8200] and the weakness documented in [RFC8250].

5.1. Applicability of AH

IOAM is a limited domain feature (refer to [RFC8799]). The network devices in an IOAM domain are trusted to add, update and remove IOAM options according to the constraints specified in [RFC8200]. IOAM domain does not rely on the Authentication Header (AH) as defined in [RFC4302] to secure IOAM options. The use of IOAM options with AH and its processing is not defined in this document. Future documents may define use of IOAM with AH and its processing.

6. IANA Considerations

This draft requests the following IPv6 Option Type assignments from the destination options and hop-by-hop options sub-registry of Internet Protocol Version 6 (IPv6) Parameters.

http://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml#ipv6-parameters-2

```
Hex Value     Binary Value      Description             Reference
              act chg rest
-------------------------------------------------------------------
TBD_1_0       00   0  TBD_1     IOAM                    [This draft]
                                destination option
                                and
                                IOAM hop-by-hop option
TBD_1_1       00   1  TBD_1     IOAM                    [This draft]
                                destination option
                                and
                                IOAM hop-by-hop option
```

7. Acknowledgements

The authors would like to thank Tom Herbert, Eric Vyncke, Nalini Elkins, Srihari Raghavan, Ranganathan T S, Karthik Babu Harichandra Babu, Akshaya Nadahalli, Stefano Previdi, Hemant Singh, Erik Nordmark, LJ Wobker, Mark Smith, Andrew Yourtchenko and Justin Iurman for the comments and advice. For the IPv6 encapsulation, this document leverages concepts described in [I-D.kitamura-ipv6-record-route]. The authors would like to acknowledge the work done by the author Hiroshi Kitamura and people involved in writing it.

8. References

8.1. Normative References

[I-D.ietf-ippm-ioam-direct-export]
Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In Situ Operations, Administration, and

Maintenance (IOAM) Direct Exporting", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-direct-export-11, 23 September 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-ippm-ioam-direct-export-11>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC9197] Brockners, F., Ed., Bhandari, S., Ed., and T. Mizrahi, Ed., "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", RFC 9197, DOI 10.17487/RFC9197, May 2022, <https://www.rfc-editor.org/info/rfc9197>.

8.2. Informative References

[I-D.kitamura-ipv6-record-route]
Kitamura, H., "Record Route for IPv6 (PR6) Hop-by-Hop Option Extension", Work in Progress, Internet-Draft, draft-kitamura-ipv6-record-route-00, November 2000, <https://tools.ietf.org/id/draft-kitamura-ipv6-record-route-00.txt>.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302, DOI 10.17487/RFC4302, December 2005, <https://www.rfc-editor.org/info/rfc4302>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8250] Elkins, N., Hamilton, R., and M. Ackermann, "IPv6 Performance and Diagnostic Metrics (PDM) Destination Option", RFC 8250, DOI 10.17487/RFC8250, September 2017, <https://www.rfc-editor.org/info/rfc8250>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

Contributors

This document was the collective effort of several authors. The text and content were contributed by the editors and the co-authors listed below. The contact information of the co-authors appears at the end of this document.

*Carlos Pignataro

*Hannes Gredler

*John Leddy

*Stephen Youell

*Tal Mizrahi

*Aviv Kfir

*Barak Gafni

*Petr Lapukhov

*Mickey Spiegel

*Suresh Krishnan

*Rajiv Asati

*Mark Smith

Contributors' Addresses

Carlos Pignataro
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States
Email: cpignata@cisco.com

Hannes Gredler
RtBrick Inc.
Email: hannes@rtbrick.com

John Leddy
Email: john@leddy.net

Stephen Youell
JP Morgan Chase
25 Bank Street
London E14 5JP
United Kingdom
Email: stephen.youell@jpmorgan.com

Tal Mizrahi
Huawei Network.IO Innovation Lab
Israel
Email: tal.mizrahi.phd@gmail.com

Aviv Kfir
Mellanox Technologies, Inc.
350 Oakmead Parkway, Suite 100
Sunnyvale, CA 94085
U.S.A.
Email: avivk@mellanox.com

Barak Gafni
Mellanox Technologies, Inc.
350 Oakmead Parkway, Suite 100
Sunnyvale, CA 94085
U.S.A.
Email: gbarak@mellanox.com

Petr Lapukhov
Facebook

1 Hacker Way
Menlo Park, CA 94025
US
Email: petr@fb.com

Mickey Spiegel
Barefoot Networks, an Intel company
4750 Patrick Henry Drive
Santa Clara, CA 95054
US
Email: mickey.spiegel@intel.com

Suresh Krishnan
Kaloom
Email: suresh@kaloom.com

Rajiv Asati
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709
US
Email: rajiva@cisco.com

Mark Smith
PO BOX 521
HEIDELBERG, VIC 3084
AU
Email: markzzzsmith+id@gmail.com

Authors' Addresses

Shwetha Bhandari (editor)
Thoughtspot
3rd Floor, Indiqube Orion, 24th Main Rd, Garden Layout, HSR Layout
Bangalore, KARNATAKA 560 102
India

Email: shwetha.bhandari@thoughtspot.com

Frank Brockners (editor)
Cisco Systems, Inc.
Hansaallee 249, 3rd Floor
40549 DUESSELDORF
Germany

Email: fbrockne@cisco.com