RFC 9486
IPv6 Options for In Situ Operations, Administration, and Maintenance (IOAM)

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
This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9486.

Copyright Notice
Copyright (c) 2023 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.
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]. 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.
The terms "encapsulation" and "decapsulation" are used in this document in the same way as in [RFC9197]: An IOAM encapsulating node incorporates one or more IOAM Option-Types into packets that IOAM is enabled for.

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] 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 (PDM) 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 onward, IPv6 IOAM in Linux Kernel) and IOAM for IPv6 in Vector Packet Processing (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].
Figure 1: IPv6 Hop-by-Hop and Destination Option Format for Carrying IOAM Data-Fields

<table>
<thead>
<tr>
<th>Option-Type</th>
<th>Opt Data Len</th>
<th>Reserved</th>
<th>IOAM Opt-Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 Option-Type: Abbreviated to "IOAM Opt-Type" in the diagram above: 8-bit field as defined in Section 4.1 of [RFC9197].

Option Data: Variable-length field. The data is specific to the Option-Type, as detailed below.

Pre-allocated Trace Option: The IOAM Pre-allocated 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: 0x31 (8-bit identifier of the IPv6 Option-Type for IOAM).

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

Proof of Transit Option-Type: 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: 0x31 (8-bit identifier of the IPv6 Option-Type for IOAM).

IOAM Type: IOAM POT Option-Type.

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

Option-Type: 0x11 (8-bit identifier of the IPv6 Option-Type for IOAM).

IOAM Type: IOAM E2E Option-Type.
Direct Export (DEX) Option: The IOAM Direct Export Option-Type, defined in Section 3.2 of [RFC9326], is represented as an IPv6 option in the hop-by-hop extension header:

Option-Type: 0x11 (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 alignment on multiples of 4 bytes. 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 in an IOAM-Domain. An IOAM-Domain is a set of nodes that use IOAM. An IOAM-Domain is bounded by its perimeter or edge. The set of nodes forming an IOAM-Domain may be connected to the same physical infrastructure (e.g., a service provider's network). They may also be remotely connected to each other (e.g., an enterprise VPN or an overlay). It is expected that all nodes in an IOAM-Domain are managed by the same administrative entity. Please refer to [RFC9197] for more details on IOAM-Domains.

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 Multipath (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
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, i.e., hosts will place the IOAM Data-Fields directly in the IPv6 header or remove the IOAM Data-Fields directly from the IPv6 header. 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. Network devices will encapsulate IOAM Data-Fields in an additional, outer, IPv6 header that carries the IOAM Data-Fields.

5. Security Considerations

This document describes the encapsulation of IOAM Data-Fields in IPv6. For general IOAM security considerations, see [RFC9197]. Security considerations of the specific IOAM Data-Fields for each case (i.e., Trace, POT, and E2E) are also described and defined in [RFC9197].

As this document describes new options for IPv6, the security considerations of [RFC8200] and [RFC8250] apply.

From a network-protection perspective, there is an assumed trust model such that any node that adds IOAM to a packet, removes IOAM from a packet, or modifies IOAM Data-Fields of a packet is assumed to be allowed to do so. By default, packets that include IPv6 extension headers with IOAM information MUST NOT be leaked through the boundaries of the IOAM-Domain.

IOAM-Domain boundary routers MUST filter any incoming traffic from outside the IOAM-Domain that contains IPv6 extension headers with IOAM information. IOAM-Domain boundary routers MUST also filter any outgoing traffic leaving the IOAM-Domain that contains IPv6 extension headers with IOAM information.

In the general case, an IOAM node only adds, removes, or modifies an IPv6 extension header with IOAM information, if the directive to do so comes from a trusted source and the directive is validated.
Problems may occur if the above behaviors are not implemented or if the assumed trust model is violated (e.g., through a security breach). In addition to the security considerations discussed in [RFC9197], the security considerations associated with IPv6 extension headers listed in [RFC9098] apply.

5.1. Applicability of Authentication Header (AH)

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 AH as defined in [RFC4302] to secure IOAM options. The use of IOAM options with AH and its processing are not defined in this document. Future documents may define the use of IOAM with AH and its processing.

6. IANA Considerations

IANA has assigned the IPv6 Option-Types from the “Destination Options and Hop-by-Hop Options” subregistry of “Internet Protocol Version 6 (IPv6) Parameters” <https://www.iana.org/assignments/ipv6-parameters/>.

<table>
<thead>
<tr>
<th>Hex Value</th>
<th>Binary Value</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x11</td>
<td>00 0 10001</td>
<td>IOAM Destination Option and IOAM Hop-by-Hop Option</td>
<td>RFC 9486</td>
</tr>
<tr>
<td>0x31</td>
<td>00 1 10001</td>
<td>IOAM Destination Option and IOAM Hop-by-Hop Option</td>
<td>RFC 9486</td>
</tr>
</tbody>
</table>

Table 1

7. References

7.1. Normative References


7.2. Informative References


Acknowledgements

The authors would like to thank Tom Herbert, Éric 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 [IPV6-RECORD-ROUTE]. The authors would like to acknowledge the work done by the author Hiroshi Kitamura and people involved in writing it.

Contributors

This document was the collective effort of several authors. The text and content were contributed by the editors and the coauthors listed below.
Carlos Pignataro
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States of America
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
United States of America
Email: avivk@mellanox.com

Barak Gafni
Mellanox Technologies, Inc.
350 Oakmead Parkway, Suite 100
Sunnyvale, CA 94085
United States of America
Email: gbarak@mellanox.com

Petr Lapukhov
Facebook
1 Hacker Way
Menlo Park, CA 94025
United States of America
Email: petr@fb.com
Mickey Spiegel  
Barefoot Networks, an Intel company  
4750 Patrick Henry Drive  
Santa Clara, CA 95054  
United States of America  
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  
United States of America  
Email: rajiva@cisco.com

Mark Smith  
PO BOX 521  
Heidelberg VIC 3084  
Australia  
Email: markzzzsmith+id@gmail.com

Authors' Addresses

Shwetha Bhandari (EDITOR)  
Thoughtspot  
3rd Floor, Indiqube Orion  
24th Main Rd, Garden Layout, HSR Layout  
Bangalore 560 102  
Karnataka  
India  
Email: shwetha.bhandari@thoughtspot.com

Frank Brockners (EDITOR)  
Cisco Systems, Inc.  
Hanssaallee 249, 3rd Floor  
40549 Duesseldorf  
Germany  
Email: fbrockne@cisco.com