ippm,6man
Internet-Draft
Intended status: Standards Track
Expires: September 30, 2020

S. Bhandari
F. Brockners
Cisco
T. Mizrahi
Huawei Network.IO Innovation Lab
A. Kfir
B. Gafni
Mellanox Technologies, Inc.
M. Spiegel
Barefoot Networks, an Intel company
S. Krishnan
Kaloom
M. Smith
March 29, 2020

Deployment Considerations for In-situ OAM with IPv6 Options
draft-ioametal-ippm-6man-ioam-ipv6-deployment-03

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 can be enabled in an IPv6 network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 30, 2020.

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. [I-D.ioametal-ippm-6man-ioam-ipv6-options] defines how IOAM data fields are encapsulated in the IPv6 [RFC8200]. This document discusses deployment options for networks which leverage IOAM data fields encapsulated in the IPv6 protocol.

Deployment considerations differ, whether the IOAM domain starts and ends on hosts or whether the IOAM encapsulating and decapsulating nodes are network devices that forward traffic, such as routers.

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

ION: IOAM Overlay Network

OAM: Operations, Administration, and Maintenance

POT: Proof of Transit

3. Considerations for IOAM deployment in IPv6 networks

IOAM deployment in an IPv6 network should take the following considerations and requirements into account:

C1 It is desirable that the addition of IOAM data fields neither changes the way routers forward the packets, nor the forwarding decision the routers takes. The packet with the added OAM information should follow the same path within the domain that the same packet without the OAM information would follow within the domain even in the presence of ECMP. Such a behavior is particularly interesting for deployments where IOAM data fields are only added "on-demand", e.g. to provide further insights in case of undesired network behavior for certain flows. Implementations of IOAM should ensure that ECMP behavior for packets with and without IOAM data fields is the same.

C2 Given that IOAM data fields increase the total size of the packet, the size of the packet including the IOAM data could exceed the PMTU. In particular, the incremental trace IOAM HbH Option, which is proposed to support hardware implementations of IOAM, changes Option Data Length en-route. Operators of an IOAM domain are to ensure that the addition of OAM information does not lead to fragmentation of the packet, e.g. by configuring the MTU of

transit routers and switches to a sufficiently high value. Careful control of the MTU in a network is one of the reasons why IOAM is considered a domain specific feature, see also [I-D.ietf-ippm-ioam-data]. In addition, the PMTU tolerance range in the IOAM domain should be identified (e.g. through configuration) and IOAM encapsulation operations and/or IOAM data field insertion (in case of incremental tracing) should not be performed if it exceeds the packet size beyond PMTU.

C3 Packets with IOAM data or associated ICMP errors, should not arrive at destinations which have no knowledge of IOAM. Consider using IOAM in transit devices; misleading ICMP errors due to addition and/or presence of OAM data in the packet can confuse a source of the packet that did not insert the OAM information.

C4 OAM data leaks may affect the forwarding behavior and state of network elements outside an IOAM domain. IOAM domains SHOULD provide a mechanism to prevent data leaks or be able to assure that upon leak network elements outside the domain are not affected i.e they continue to process other valid packets.

C5 The source of that inserted and leaked the IOAM data must be easy to identify for the purpose of troubleshooting, due to the high complexity of troubleshooting a source that inserted the IOAM data and did not remove it when the packet traversed across an AS. Such a troubleshooting process may require coordination between multiple operators, complicated configuration verification, packet capture analysis, etc.

C6 Compliance with [RFC8200] would require OAM data to be encapsulated instead of header/option insertion directly into in-flight packets using the original IPv6 header.

4. 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, see [I-D.ioametal-ippm-6man-ioam-ipv6-options].

5. 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.

5.1. Deployment options

This section lists out possible deployment options that can be employed to meet the requirements listed in Section 3.

5.1.1. IPv6-in-IPv6 encapsulation

Leverage an IPv6-in-IPv6 approach: Preserve the original IP packet and add an IPv6 header including IOAM data fields in an extension header in front of it, to forward traffic within and across the IOAM domain. The overlay network formed by the additional IPv6 header with the IOAM data fields included in an extension header is referred to as IOAM Overlay Network (ION) in this document.

1. Perform an IPv6-in-IPv6 approach. The source address of the outer IPv6 header is that of the IOAM encapsulating node. The destination address of the outer IPv6 header is the same as the inner IPv6 destination address, i.e. the destination address of the packet does not change.

2. To simplify debugging in case of leaked IOAM data fields in packets, consider a new IOAM E2E destination option to identify the Source IOAM domain (AS, v6 prefix). Insert this option into the IOAM destination options EH attached to the outer IPv6 header. This additional information would allow for easy identification of an AS operator that is the source of packets with leaked IOAM information. Note that leaked packets with IOAM data fields would only occur in case a router would be misconfigured. [I-D.ioametal-ippm-6man-ioam-ipv6-options] requires that by default, packets with extension headers which carry IOAM data fields are dropped unless the router's interfaces are explicitly configured for IOAM.

3. All the IOAM options are defined with type "00 - skip over this option and continue processing the header. So presence of the options must not cause packet drop in the network elements that do not understand the option. In addition [I-D.ietf-6man-hbh-header-handling] should be considered.

5.1.2. IP-in-IPv6 encapsulation with ULA

The "IP-in-IPv6 encapsulation with ULA" [RFC4193] approach can be used to apply IOAM to an IPv6 as well as an IPv4 network. In addition, it fulfills requirement C4 (avoid leaks) by using ULA for the ION. Similar to the IPv6-in-IPv6 encapsulation approach above, the original IP packet is preserved. An IPv6 header including IOAM data fields in an extension header is added in front of it, to forward traffic within and across the IOAM domain. IPv6 addresses

for the ION, i.e. the outer IPv6 addresses are assigned from the ULA space. Addressing and routing in the ION are to be configured so that the IP-in-IPv6 encapsulated packets follow the same path as the original, non-encapsulated packet would have taken. This would create an internal IPv6 forwarding topology using the IOAM domain's interior ULA address space which is parallel with the forwarding topology that exists with the non-IOAM address space (the topology and address space that would be followed by packets that do not have supplemental IOAM information). Establishment and maintenance of the parallel IOAM ULA forwarding topology could be automated, e.g. similar to how LDP [RFC5036] is used in MPLS to establish and maintain an LSP forwarding topology that is parallel to the network's IGP forwarding topology.

Transit across the ION could leverage the transit approach for traffic between BGP border routers, as described in [RFC1772], "A.2.3 Encapsulation". Assuming that the operational guidelines specified in Section 4 of [RFC4193] are properly followed, the probability of leaks in this approach will be almost close to zero. If the packets do leak through IOAM egress device misconfiguration or partial IOAM egress device failure, the packets' ULA destination address is invalid outside of the IOAM domain. There is no exterior destination to be reached, and the packets will be dropped when they encounter either a router external to the IOAM domain that has a packet filter that drops packets with ULA destinations, or a router that does not have a default route.

5.1.3. x-in-IPv6 Encapsulation that is used Independently

In some cases it is desirable to monitor a domain that uses an overlay network that is deployed independently of the need for IOAM, e.g., an overlay network that runs Geneve-in-IPv6, or VXLAN-in-IPv6. In this case IOAM can be encapsulated in as an extension header in the tunnel (outer) IPv6 header. Thus, the tunnel encapsulating node is also the IOAM encapsulating node, and the tunnel end point is also the IOAM decapsulating node.

6. Security Considerations

This document discusses the deployment of IOAM with IPv6 options. Security considerations of the specific IOAM data fields are described in [I-D.ietf-ippm-ioam-data].

7. IANA Considerations

There are no IANA considerations that apply to this document.

8. Acknowledgements

The authors would like to thank Mark Smith, 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, and Andrew Yourtchenko 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.

9. References

9.1. Normative References

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-04 (work in progress), October 2018.

[I-D.ioametal-ippm-6man-ioam-ipv6-options]
Bhandari, S., Brockners, F., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Kfir, A., Gafni, B., Lapukhov, P., Spiegel, M., and S. Krishnan, "In-situ OAM IPv6 Options", draft-ioametal-ippm-6man-ioam-ipv6-options-01 (work in progress), October 2018.

[RFC1772] Rekhter, Y. and P. Gross, "Application of the Border Gateway Protocol in the Internet", RFC 1772, DOI 10.17487/RFC1772, March 1995, <https://www.rfc-editor.org/info/rfc1772>.

[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>.

9.2. Informative References

[I-D.ietf-6man-hbh-header-handling]
Baker, F. and R. Bonica, "IPv6 Hop-by-Hop Options Extension Header", March 2016.

[I-D.kitamura-ipv6-record-route]
Kitamura, H., "Record Route for IPv6 (PR6) Hop-by-Hop Option Extension", draft-kitamura-ipv6-record-route-00 (work in progress), November 2000.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, DOI 10.17487/RFC4193, October 2005, <https://www.rfc-editor.org/info/rfc4193>.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <https://www.rfc-editor.org/info/rfc5036>.

[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>.

Authors' Addresses

Shwetha Bhandari
Cisco Systems, Inc.
Cessna Business Park, Sarjapura Marathalli Outer Ring Road
Bangalore, KARNATAKA 560 087
India

Email: shwethab@cisco.com

Frank Brockners
Cisco Systems, Inc.
Kaiserswerther Str. 115,
RATINGEN, NORDRHEIN-WESTFALEN 40880
Germany

Email: fbrockne@cisco.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

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

Mark Smith
PO BOX 521
HEIDELBERG, VIC 3084
AU

Email: markzzzsmith+id@gmail.com