IPPM
Internet-Draft
Intended status: Standards Track
Expires: January 5, 2020

T. Mizrahi
Huawei Network.IO Innovation Lab
F. Brockners
S. Bhandari
R. Sivakolundu
C. Pignataro
Cisco
A. Kfir
B. Gafni
Mellanox Technologies, Inc.
M. Spiegel
Barefoot Networks
J. Lemon
Broadcom

July 04, 2019

In-situ OAM Flags
draft-mizrahi-ippm-ioam-flags-00

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 presents new flags in the IOAM Trace Option headers. Specifically, the document defines the Loopback, Active, and Immediate Export flags.

Table of Contents

1. Introduction

IOAM [I-D.ietf-ippm-ioam-data] is used for monitoring traffic in the network by incorporating IOAM data fields into in-flight data packets.

IOAM data may be represented in one of four possible IOAM options: Pre-allocated Trace Option, Incremental Trace Option, Proof of Transit (POT) Option, and Edge-to-Edge Option. This document defines three new flags in the Pre-allocated and Incremental Trace options: the Loopback, Active, and Immediate Export flags.

2. Conventions

2.1. Requirement Language

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 [RFC2119].

2.2. Terminology

Abbreviations used in this document:

IOAM: In-situ Operations, Administration, and Maintenance

OAM: Operations, Administration, and Maintenance

3. New IOAM Trace Option Flags

This document defines three new flags in the Pre-allocated and Incremental Trace options:

Bit 1 "Loopback" (L-bit). Loopback mode is used to send a copy of a packet back towards the source, as further described in Section 4.

Bit 2 "Active" (A-bit). When set, this indicates that this is an active IOAM packet, where "active" is used in the sense defined in [RFC7799], rather than a data packet. The packet may be an IOAM probe packet, or a replicated data packet (the second and third use cases of Section 5).

Bit 3 "Immediate Export" (I-bit). Immediate export mode is used to export IOAM data fields immediately at every IOAM supported network node, instead of adding the IOAM data fields to the packet traversing the network. Further details are provided in Section 6.

4. Loopback in IOAM

Loopback is used for trigerring each transit device along the path to loop back a copy of the data packet. Loopback mode assumes that a return path from transit nodes and destination nodes towards the source exists. The encapsulating node decides (e.g., using a filter) which packets loopback mode is enabled for by setting the loopback bit. The encapsulating node also needs to ensure that sufficient space is available in the IOAM header for loopback operation, which includes intermediate nodes adding trace data on the original path and then again on the return path. A loopback bit that is set indicates to the transit nodes processing this option that they are

to create a copy of the received packet and send the copy back to the source of the packet. The copy has its metadata added after being copied in order to allow any egress-dependent information to be set based on the egress of the copy rather than the original. The original packet continues towards its destination. The source address of the original packet is used as the destination address in the copied packet. The address of the node performing the copy operation is used as the source address. The L-bit MUST be cleared in the copy of the packet that a node sends back towards the source. On its way back towards the source, the copied packet is processed like any other packet with IOAM information, including adding any requested data at each transit node (assuming there is sufficient space). Once the return packet reaches the IOAM domain boundary, IOAM decapsulation occurs as with any other packet containing IOAM information. Because any intermediate node receiving such a packet would not know how to process the original packet, and because there would be a risk of the original packet leaking past the initiator of the IOAM loopback, the initiator of an IOAM loopback MUST be the initiator of the packet. Once a loopback packet is received back at the initiator, it is a local matter how it is recognized as a loopback packet.

5. Active Measurement with IOAM

Active measurement methods [RFC7799] make use of synthetically generated packets in order to facilitate the measurement. This section presents use cases of active measurement using the IOAM Active flag.

The active flag indicates that a packet is used for active measurement. An IOAM decapsulating node that receives a packet with the Active flag set in one of its Trace options must terminate the packet.

An example of an IOAM deployment scenario is illustrated in Figure 1. The figure depicts two endpoints, a source and a destination. The data traffic from the source to the destination is forwarded through a set of network devices, including an IOAM encapsulating node, which incorporates one or more IOAM option, a decapsulating node, which removes the IOAM options, optionally one or more transit nodes. The IOAM options are encapsulated in one of the IOAM encapsulation types, e.g., [I-D.ietf-sfc-ioam-nsh], or [I-D.ioametal-ippm-6man-ioam-ipv6-options].

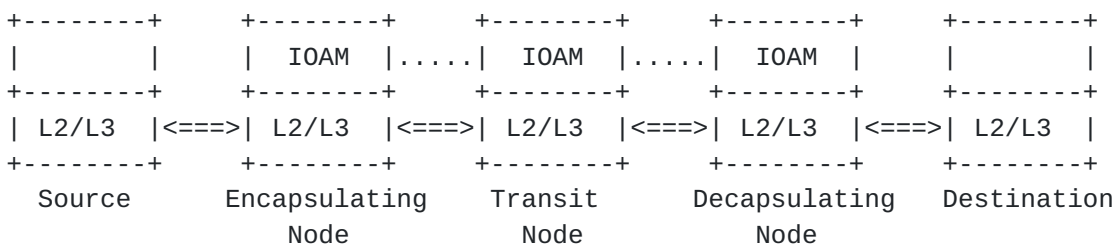

Figure 1: Network using IOAM.

This draft focuses on three possible use cases of active measurement using IOAM. These use cases are described using the example of Figure 1.

- Endpoint active measurement: synthetic probe packets are sent between the source and destination, traversing the IOAM domain. Since the probe packets are sent between the endpoints, these packets are treated as data packets by the IOAM domain, and do not require special treatment at the IOAM layer.

- IOAM active measurement using probe packets: probe packets are generated and transmitted by the IOAM encapsulating node, and are expected to be terminated by the decapsulating node. IOAM data related to probe packets may be exported by one or more nodes along its path, by an exporting protocol that is outside the scope of this document (e.g., [I-D.spiegel-ippm-ioam-rawexport]). Probe packets include a Trace Option which has its Active flag set, indicating that the decapsulating node must terminate them.

- IOAM active measurement using replicated data packets: probe packets are created by the encapsulating node by selecting some or all of the en route data packets and replicating them. A selected data packet that is replicated, and its (possibly truncated) copy is forwarded with one or more IOAM option, while the original packet is forwarded normally, without IOAM options. To the extent possible, the original data packet and its replica are forwarded through the same path. The replica includes a Trace Option that has its Active flag set, indicating that the decapsulating node should terminate it.

6. Immediate Exporting

Immediate exporting can be used to export IOAM data to a collector instead of incorporating this data into en route data packets. The various types of IOAM nodes MUST process packets with the I-bit set as follows:

1. An encapsulating IOAM node configured to set the I-bit encapsulates the packet with the IOAM header and sets the I-bit, leaving the IOAM header without locally collected data, and exports the requested IOAM data immediately. The encapsulating IOAM node is the only type of node allowed to set the I-bit.

2. A transit node that processes a packet with the I-bit set is expected to export the requested IOAM data, and not incorporate it into the IOAM header.

3. A decapsulating IOAM node that processes a packet with the I-bit set is expected to export the requested IOAM data, and decapsulate the IOAM header.

Note that in case of "Immediate Export" being employed, no IOAM trace data is added to the packets traversing the network. As a means to support correlation of exported IOAM data different nodes in the network, a deployment could consider attaching an IOAM E2E option in addition to the trace option, that includes a sequence number. See Bit 1 in the IOAM-E2E-Types. Please refer to [I-D.ietf-ippm-ioam-data] for a discussion of IOAM data export and associated formats.

7. IANA Considerations

IANA is requested to allocate the following bits in the "IOAM Trace Flags Registry" as follows:

Bit 1 "Loopback" (L-bit)

Bit 2 "Active" (A-bit)

Bit 3 "Immediate Export" (I-bit)

Note that bit 0 is the most significant bit in the Flags Registry.

8. Performance Considerations

Each of the three flags that are defined in this document may have performance implications. When using the loopback mechanism a copy of the data packet is sent back to the sender, thus generating more traffic than originally sent by the endpoints. Using active measurement with the active flag requires the use of synthetic (overhead) traffic. The Immediate Export flag triggers exported packets to be exported to a collector, which in some cases may impact the collector's performance, or the performance along the paths leading to the collector.

Each of the three mechanisms has a cost in terms of the network bandwidth, and may potentially load the node that analyzes the data. Therefore, rate limiting may be enabled so as to ensure that the three mechanisms are used at a rate that does not significantly affect the network bandwidth, and does not overload the collector (or the source node in the case of loopback). It should be possible to use each of the three mechanisms on a subset of the data traffic.

9. Security Considerations

The security considerations of IOAM in general are discussed in [I-D.ietf-ippm-ioam-data]. Specifically, an attacker may try to use the functionality that is defined in this document to attack the network.

An attacker may attempt to overload network devices by injecting synthetic packets that include an IOAM Trace Option with one or more of the flags defined in this document. Similarly, an on-path attacker may maliciously set one or more of the flags of transit packets.

- Loopback flag: an attacker that sets this flag, either in synthetic packets or transit packet, can potentially cause an amplification, since each device along the path creates a copy of the data packet and sends it back to the source. The attacker can potentially leverage the loopback flag for a Distributed Denial of Service (DDoS) attack, as multiple devices send looped-back copies of a packet to a single source.

- Active flag: the impact of synthetic packets with the active flag is no worse than synthetic data packets in which the Active flag is not set. By setting the active flag in en route packets an attacker can prevent these packets from reaching their destination, since the packet is terminated by the decapsulating device; however, note that an on-path attacker may achieve the same goal by changing the destination address of a packet. Another potential threat is amplification; if an attacker causes transit switches to replicate more packets than they are intended to replicate, either by setting the Active flag or by sending synthetic packets, then traffic is amplified, causing bandwidth degredation.

- Immediate Export flag: setting this flag, either in synthetic packets or in transit packets may overload the collector or analyzer devices. Since this flag affects multiple devices along the network path, it potentially amplifies the effect on the network bandwidth and on the collector's load.

In order to mitigate the attacks described above, it should be possible for IOAM-enabled devices to limit each of the three mechanisms to a configurable rate; Network devices should be able to limit the rate of: (i) looped-back traffic, (ii) replicated active packets, and (iii) exported packets.

IOAM is assumed to be deployed in a restricted administrative domain, thus limiting the scope of the threats above and their affect. This is a fundamental assumtion with respect to the security aspects of IOAM, as further discussed in [I-D.ietf-ippm-ioam-data].

10. References

10.1. 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-05 (work in progress), March 2019.

[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>.

10.2. Informative References

[I-D.ietf-sfc-ioam-nsh]
Brockners, F., Bhandari, S., Govindan, V., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., and R. Chang, "Network Service Header (NSH) Encapsulation for In-situ OAM (IOAM) Data", draft-ietf-sfc-ioam-nsh-01 (work in progress), March 2019.

[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., Krishnan, S., and R. Asati, "In-situ OAM IPv6 Options", draft-ioametal-ippm-6man-ioam-ipv6-options-02 (work in progress), March 2019.

[I-D.spiegel-ippm-ioam-rawexport]
Spiegel, M., Brockners, F., Bhandari, S., and R. Sivakolundu, "In-situ OAM raw data export with IPFIX", draft-spiegel-ippm-ioam-rawexport-01 (work in progress), October 2018.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

Authors' Addresses

Tal Mizrahi
Huawei Network.IO Innovation Lab
Israel

Email: tal.mizrahi.phd@gmail.com

Frank Brockners
Cisco Systems, Inc.
Hansaallee 249, 3rd Floor
DUESSELDORF, NORDRHEIN-WESTFALEN 40549
Germany

Email: fbrockne@cisco.com

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

Email: shwethab@cisco.com

Ramesh Sivakolundu
Cisco Systems, Inc.
170 West Tasman Dr.
SAN JOSE, CA 95134
U.S.A.

Email: sramesh@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States

Email: cpignata@cisco.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
4750 Patrick Henry Drive
Santa Clara, CA 95054
US

Email: mspiegel@barefootnetworks.com

John Lemon
Broadcom
270 Innovation Drive
San Jose, CA 95134
US

Email: john.lemon@broadcom.com