Abstract

This document discusses the requirement of on-path telemetry for multicast traffic. The existing solutions are examined and their issues are addressed with new modifications that adapt to the multicast scenario.

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.

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 January 2, 2020.

Copyright Notice

Copyright (c) 2019 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
1. Introduction

Multicast traffic is an important traffic type in today’s Internet. Multicast provides services that are often real time (e.g., online meeting) or have strict QoS requirements (e.g., IPTV, Market Data). Multicast packet drop and delay can severely affect the application performance and user experience.

It is important to monitor the performance of the multicast traffic. Existing OAM techniques cannot gain direct and accurate information about the multicast traffic. New on-path telemetry techniques such as In-situ OAM [I-D.brockners-inband-oam-data] and Postcard-based Telemetry [I-D.song-ippm-postcard-based-telemetry] provide promising means to directly monitor the network experience of multicast traffic. However, multicast traffic has some unique characteristics which pose some challenges on efficiently applying such techniques.

This document describes the requirement for multicast traffic telemetry and shows the issues of the existing on-path telemetry.
techniques. We then propose modifications to make these techniques adapt to the multicast application.

2. Requirements for Multicast Traffic Telemetry

Multicast traffic is forwarded through a multicast tree. With PIM and P2MP (MLDP, RSVP-TE) the forwarding tree is established and maintained by the multicast routing protocol. With BIER, no state is created in the network to establish a forwarding tree, instead, a bier header provides the necessary information for each packet to know the egress points. Multicast packets are only replicated at each tree branch node for efficiency.

There are several requirements for multicast traffic telemetry:

- Reconstruct and visualize the multicast tree through data plane monitoring.
- Gather the multicast packet delay and jitter performance.
- Find the multicast packet drop location and reason.
- Gather the VPN state and tunnel information in case of P2MP multicast.

All of these requirements need to directly monitor the multicast traffic and derive data from the multicast packets. The conventional OAM mechanisms such as multicast ping and trace are insufficient to meet these requirements.

3. Issues of Existing Techniques

On-path Telemetry techniques that directly retrieve data from multicast traffic’s live network experience are ideal to address the above mentioned requirements. The representative techniques include In-situ OAM (IOAM) [I-D.brockners-inband-oam-data] and Postcard-based Telemetry (PBT) [I-D.song-ippm-postcard-based-telemetry]. However, unlike unicast, multicast poses some unique challenges to applying these techniques.

Multicast packets are replicated at each branch node in the corresponding multicast tree. Therefore, there are multiple copies of a packets in the network.

If IOAM is used for on-path data collection, the partial trace data will also be replicated into multiple copies. The end result is that each copy of the multicast packet has a complete trace. Most of the data is redundant. Data redundancy introduces unnecessary header
overhead, wastes network bandwidth, and complicates the data processing. In case the multicast tree is large and the path is long, the redundancy problem becomes severe.

PBT can be used to eliminate such data redundancy, because each node on the tree only sends a postcard covering local data. However, PBT cannot track the tree branches properly so it can bring confusion about the multicast tree topology. For example, Node A has two branches, one to Node B and one to node D, and Node B leads to Node C and Node D leads to Node E. From the received postcard, one cannot tell whether or not Node C(E) is the next hop of Node B(D).

The fundamental reason for this problem is that there is not an identifier (either implicit or explicit) to correlate the data on each branch.

4. Proposed Modifications to Existing Techniques

We propose two solutions to address the above issues. One is built on PBT and requires augmentation to the instruction header of PBT-I; the other combines the IOAM trace mode and PBT for an optimized solution.

4.1. Per-hop postcard

The straightforward way to mitigate the PBT’s drawback is to augment it with a branch identifier field. Note that this only works for the PBT-I variation where an instruction header is present. To make the branch identifier globally unique, the branch node ID plus an index is used. For example, if Node A has two branches, to Nodes B and C, Node A will use [A, 0] as the branch identifier for the branch to B, and [A, 1] for the branch to C. The identifier is unchanged and carried with the multicast packet until the next branch node. Each postcard needs to include the branch identifier in the export data. The branch identifier and the other fields such as flow ID and sequence number are sufficient for the data analyzer to reconstruct the topology of the multicast tree.

Figure 1 shows an example of this solution. "P" stands for the postcard packet. The square brackets contains the branch identifier. The curly brace contains the telemetry data about a specific node.
4.2. Per-section postcard

The second solution is a combination of the IOAM trace mode and PBT. To avoid the data redundancy, at each branch node, the trace data accumulated so far is exported by a postcard before the packet is replicated. In this case, each branch still needs to keep some identifier to help correlate the postcards for each tree section. The natural way is to simply carry the branch node’s data (including its ID) in the trace of each branch. This is also necessary because each replicated multicast packet can have different telemetry data pertaining to this particular copy (e.g., node delay, egress timestamp, and egress interface). As a consequence of, the local data exported by each branch node can only contain partial data (e.g., ingress interface and ingress timestamp).

Figure 2 shows an example in a part of a multicast tree. Node B and D are two branch nodes so they will export a postcard covering the trace data for the previous section. The end node of each path need to export the data of the last section as a postcard as well.
5. Considerations for Different Multicast Protocols

MTRACEv2 [RFC8487] provides an active probing approach for the tracing of an IP multicast routing path. Mtrace can also provide information such as the packet rates and losses, as well as other diagnostic information. New on-path telemetry techniques will enhance Mtrace, and other existing OAM solutions, with more granular and realtime network status data through direct measurements. There are various multicast protocols that are used to forward the multicast data. Each will require their own unique on-path telemetry solution.

5.1. Application in PIM

PIM-SM [RFC7761] is the most widely used multicast routing protocol deployed today. Of the various PIM modes (PIM-SM, PIM-DM, BIDIR-PIM, PIM-SSM), PIM-SSM is the preferred method due to its simplicity and removal of network source discovery complexity. With all PIM modes, control plane state is established in the network in order to forward multicast UDP data packets. But with PIM-SSM, the discovery of multicast sources is performed outside of the network via HTTP, SDN, etc. IP Multicast packets fall within the range of 224.0.0.0 through 239.255.255.255.
239.255.255.255. The telemetry solution will need to work within this address range and provide telemetry data for this UDP traffic.

The proposed solutions for encapsulating the telemetry instruction header and metadata in IPv4/IPv6 UDP packets are described in [I-D.herbert-ipv4-udpenccap-eh] and [I-D.ioametal-ippm-6man-ioam-ipv6-deployment].

5.2. Application in P2MP

Multicast Label Distribution Protocol (MLDP) and P2MP RSVP-TE are commonly used within a Multicast VPN (MVPN) environment. MLDP provides extensions to LDP to establish point-to-multipoint (P2MP) and multipoint-to-multipoint (MP2MP) label switched paths (LSPs) in MPLS networks. P2MP RSVP-TE provides extensions to RSVP-TE for establish traffic-engineered P2MP LSPs in MPLS networks. The telemetry solution will need to be able to follow these P2MP paths. The telemetry instruction header and data should be encapsulated into MPLS packets on P2MP paths. A corresponding proposal is described in [I-D.song-mpls-extension-header].

5.3. Application in BIER

BIER [RFC8279] adds a new header to multicast packets and allows the multicast packets to be forwarded according to the header only. By eliminating the requirement of maintaining per multicast group state, BIER is more scalable than the traditional multicast solutions.

OAM Requirements for BIER [I-D.ietf-bier-oam-requirements] lists many of the requirements for OAM at the BIER layer which will help in the forming of on-path telemetry requirements as well.

There is also current work to provide solutions for BIER forwarding in ipv6 networks. For instance, a solution, BIER in Non-MPLS IPv6 Networks [I-D.xie-bier-ipv6-encapsulation], proposes a new bier Option Type codepoint from the "Destination Options and Hop-by-Hop Options" IPv6 sub-registry. This is similar to what IOAM proposes for IPv6 transport.

Depending on how the BIER header is encapsulated into packets with different transport protocols, the method to encapsulate the telemetry instruction header and metadata also varies. It is also possible to make the instruction header and metadata a part of the BIER header itself, such as in a TLV.
6. Security Considerations

No new security issues are identified other than those discovered by the IOAM and PBT drafts.

7. IANA Considerations

The document makes no request of IANA.

8. Contributors

TBD

9. Acknowledgments

TBD

10. References

10.1. Normative References


informative references

10.2. Informative References


[I-D.xie-bier-ipv6-encapsulation]


Authors' Addresses

Haoyu Song (editor)
Futurewei Technologies
2330 Central Expressway
Santa Clara
USA

Email: hsong@futurewei.com

Mike McBride
Futurewei Technologies
2330 Central Expressway
Santa Clara
USA

Email: mmcbride@futurewei.com