

TRILL Working Group
INTERNET-DRAFT
Intended Status: Standard Track

Y. Li
D. Eastlake
H. Chen
Huawei Technologies
D. Kumar
Cisco
S. Gupta
IP Infusion
July 6, 2015

Expires: January 6, 2016

TRILL: Traceable OAM
draft-yizhou-trill-traceable-oam-00

Abstract

TRILL fault management [[RFC7455](#)] specifies the messages and operations for OAM in TRILL network. The sender collects the replies for the OAM-relevant request it sent and uses the replies as the indication of the network faults. In certain circumstances the sender needs to collect multiple replies to isolate the fault, e.g. repetitively sending Path Trace Messages (PTM) with increasing value of hop count and collecting the replies on them to figure out the fault point of certain path.

With the increasing deployment of Software Defined Network (SDN), a centralized management server can be used to help with fault management. The server then is responsible to collect OAM messages and analyze them to either isolate the network fault or compile overall OAM information. It naturally uses SDN structure to alleviate the effort of the requester node and provide a centralized solution to produce the operation and management feedback of the network.

This document specifies the extensions of TRILL OAM message and the operations about the network nodes and the centralized management server to trace and collect OAM relevant messages for further analysis.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as

Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/1id-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2015 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 (<http://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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- [1. Introduction](#) [4](#)
- [2. Terminology Used in This Document](#) [5](#)
- [3. Traceable Flag](#) [5](#)
- [4. Operation Theory](#) [6](#)
 - [4.1 Path Trace Message \(PTM\) with Traceable Flag](#) [6](#)
 - [4.1.1 Actions by Originator RBridge](#) [7](#)
 - [4.1.2 Intermediate RBridge](#) [8](#)
 - [4.1.3 Destination RBridge](#) [8](#)
 - [4.1.4 Centralized Management Server](#) [8](#)
 - [4.2 Multi-Destination Tree Verification Message \(MTVM\) with Traceable Flag](#) [8](#)
 - [4.2.1 Actions by Originator RBridge](#) [9](#)
 - [4.2.2 Receiving RBridge](#) [10](#)
 - [4.2.3 In-Scope RBridges](#) [10](#)
 - [4.2.4 Centralized Management Server](#) [11](#)

[5. Security Considerations](#) [11](#)
[6. IANA Considerations](#) [11](#)
[7. References](#) [11](#)
 [7.1 Normative References](#) [11](#)
 [7.2 Informative References](#) [12](#)
[8. Acknowledgments](#) [12](#)
Authors' Addresses [12](#)

1. Introduction

TRILL fault management [RFC7455] specifies OAM messages, TLVs and operations for fault detection and isolation in TRILL network. Requester collects the replies of the OAM requests it sent and tries to analyze the potential faults from them. For path tracing, the sending RBridge utilizes a Hop Count Expiry approach [RFC7455] which is similar as IP trace-route. The sending RBridge sends multiple requests in sequence with incrementing value of Hop Count field and collects the returning responses to construct the path and isolate the fault.

With the deployment of centralized management server in TRILL network, new requirements and approaches for fault isolation and path tracing are emerging. Figure 1 shows a TRILL network with a centralized management server.

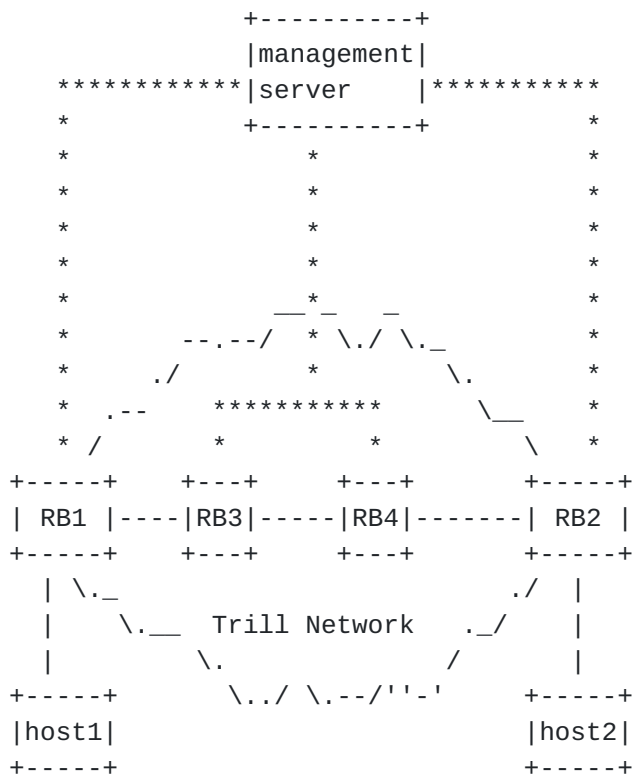


Figure 1. TRILL network with a centralized management server

The centralized management server is capable to construct the OAM messages of a specified flow and feed them into an ingress RBridge, say RB1. When the message is delivered to the egress RBridge RB2, the intermediate RBridges RB3 and RB4 are able to replicate the messages

to the management server. The server is responsible to do all the analysis to trace the path and isolate the fault. Such approach is easily deployable in a network with a controller. For instance, if the management server is an Openflow [Openflow] controller, RBridges may use Packet-in message to send the packets to the Openflow controller and the controller may use Packet-out message to feed the constructed OAM messages into the ingress RB at the beginning.

The document defines the flags and TLVs to help the RBridges to identify the received OAM messages destined for a centralized management server and provides the server with sufficient information for further analysis.

2. Terminology Used in This Document

This document uses the terminology from [[RFC6325](#)], [[RFC7174](#)] and [[RFC7455](#)]. Some additional terms listed below:

campus: Name for a TRILL network, like "bridged LAN" is a name for a bridged network. It does not have any academic implication.

Data Label: VLAN or FGL.

ECMP: Equal Cost Multi-Path [[RFC6325](#)].

FGL: Fine Grained Label [[RFC7172](#)].

RBridge: An alternative name for a TRILL switch.

TRILL switch: A device implementing the TRILL protocol. Sometimes called an RBridge.

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 [RFC-2119](#) [[RFC2119](#)].

3. Traceable Flag

A new flag 'T' is defined in TRILL OAM message header [[RFC7455](#)] as an indicator for traceable message in figure 2. T flag is applicable to Path Trace Message (PTM) and Multi-Destination Tree Verification Message (MTVM). Loopback message and Continuity Check Message SHOULD not set T flag to 1.

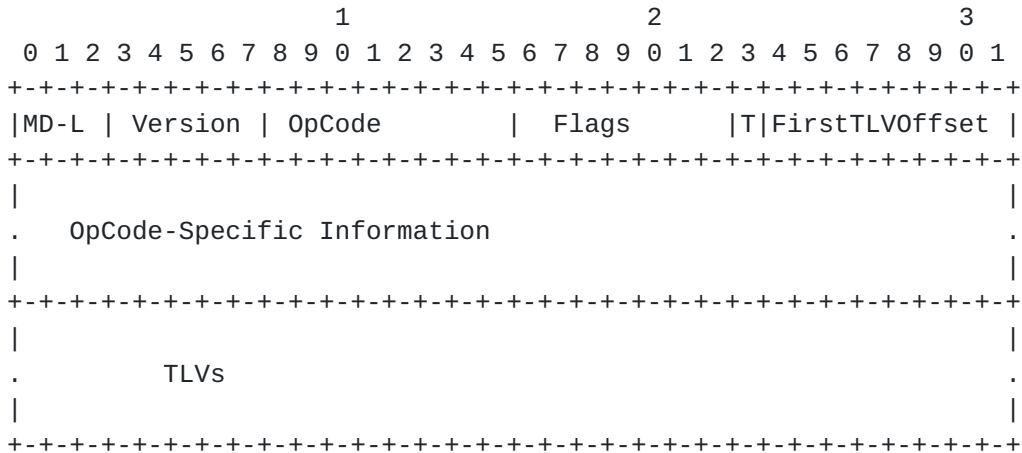


Figure 2. T Flag in TRILL OAM Message Header

o T (1 bit): Traceable flag. When set, indicates no response should be sent back to the requester and the entire TRILL frame should be sent to a centralized management server for tracing.

Basically the traceable flag implies three functions in the trill campus:

1. To indicate the intermediate RBs to capture the frames and replicate it to CPU.
2. To guide the intermediate RB to perform certain operations which may be different from the traditional OAM operations. For example, as we can use packet-in to send the whole packet to openflow controller, it is not necessary to add Original Data Payload TLV to the packet.
3. To make sure the sender will not expect any response and turn off certain mechanisms like time out.

4. Operation Theory

OAM message with Traceable flag is most useful in functionalities requiring tracing, e.g., trace route like behaviors.

4.1 Path Trace Message (PTM) with Traceable Flag

TRILL fault management [[RFC7455](#)] adopts an IP trace-route like approach which relies on the hop count expiry to send the PTM message to RBridge for further handling. The sender needs to repetitively send the requests with increasing value of hop count. With traceable flag on, the centralized management server may collect the replicated frame along the path and check the hop count value in TRILL header

directly. By sorting the hop count value decreasingly, it is easy to plot the path taken for a specific flow or figure out the break point for fault isolation.

As a centralized management server normally has more memory space than an RBridge, the server may choose to record the flow entropy to path mapping information. When a fault is suspected between two RBridges, the sever may optimally choose minimum number of flow entropies from the records it saved to feed into the ingress RBridge to spread over the paths.

4.1.1 Actions by Originator RBridge

The originator RBridge takes the following actions:

- o Identifies the destination RBridge based on user specification or based on location of the specified destination MAC address.
- o Constructs the Flow Entropy based on user-specified parameters or implementation-specific default parameters.
- o Specifies the Hop Count of the TRILL Data frame to be larger than the expected number of hops.
- o Constructs the TRILL OAM header: set the OpCode to Path Trace Message type (65). Assign an applicable Session Identification number for the request. Return Code and Return Sub-code MUST be set to zero. Set Traceable flags to 1.
- o Includes the following OAM TLVs, where applicable:
 - Out-of-Band Reply Address TLV: When Traceable flag is set, Out-of-Band Reply Address TLV needs to be set to the address that the traced message should be sent. It is normally the IP address of the centralized management server. This address may be absent if the default centralized management server address has been configured on every RBridge.
 - Diagnostic Label TLV
 - Sender ID TLV
- o Dispatches the OAM frame to the TRILL data plane for transmission.

The originator RBridge SHOULD not expect the replies for the Path Trace Message with Traceable Flag set it sent.

4.1.2 Intermediate RBridge

The intermediate RBridges need to be configured properly as MIP for VLAN/FGL based MA. The TRILL OAM application layer validates the received OAM frame by examining the presence of the TRILL Alert flag, OAM Ethertype at the end of the Flow Entropy, the OpCode being PTM and Traceable Flag set, the intermediate RBridges take the following actions:

- o Optionally include the following TLVs:

- Previous RBridge Nickname TLV (69)
- Interface Status TLV (4)
- Next-Hop RBridge List TLV (70)
- Sender ID TLV (1)

- o Forward the received message including the TRILL header, the payload and the appended TLVs (if any) to the address specified in Out-of-Band Reply Address TLV. If Out-of-Band Reply Address TLV is not present, either forward it to a system default centralized management server or discard it.

4.1.3 Destination RBridge

Processing is identical to that in [Section 4.1.2](#). The Destination RBridge should not further forward the message in order to prevent leaking of the packet out of the TRILL campus

4.1.4 Centralized Management Server

The centralized management server is normally served as an SDN controller, e.g. an Openflow controller. It is up to the implementation how to deal with the collected packets of PTM with traceable flag from RBridges. The common logic is the centralized management server compares the Session Identification Number and hop count value in TRILL header to trace the path the packet taken.

4.2 Multi-Destination Tree Verification Message (MTVM) with Traceable Flag

MTVM can be used by the OAM tools for plotting the entire or VLAN/FGL pruned distribution tree, reachability verification for set of

RBridges on a given tree or trace along a specified tree to a set of RBridges.

A new TLV is defined as shown in figure 3.

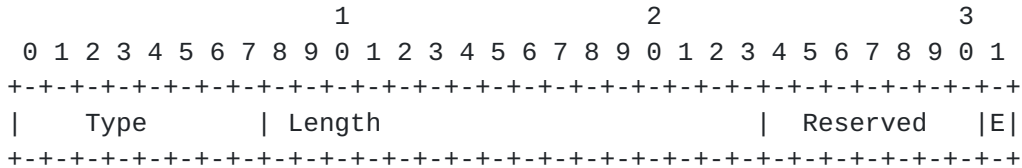


Figure 3. Tree Trace Mode TLV

- o Type (1 octet): 75 (TBD), Tree Trace Mode TLV
- o Length (2 octets): 1
- o E (1 bit): Egress Flag. When RBridge Scope TLV is not present and E flag is 1, trace the receiving RBridge which are egress RBridges on the tree of the specified VLAN/FGL or multicast group; otherwise, ignore this flag.

4.2.1 Actions by Originator RBridge

The originator RBridge takes the following actions:

- o Identifies the nickname of distribution tree to be traced.
- o Constructs the Flow Entropy based on user-specified parameters or implementation-specific default parameters.
- o Specifies the applicable Hop Count value.
- o Constructs the TRILL OAM header: set the OpCode to Multicast Tree Verification Message type (67). Assign an applicable Session Identification number for the request. Return Code and Return Sub-code MUST be set to zero. Set Traceable flags to 1.
- o Includes the following OAM TLVs, where applicable:
 - Out-of-Band Reply Address TLV: When Traceable flag is set, Out-of-Band Reply Address TLV needs to be set to the address that the traced message should be sent. It is normally the address of the centralized management server
 - RBridge Scope TLV

- Tree Trace Mode TLV: When RBridge Scope TLV is present, E flag of this TLV SHOULD not be set to 1.
 - Diagnostic Label TLV
 - Sender ID TLV
- o Dispatches the OAM frame to the TRILL data plane for transmission.

The originator RBridge SHOULD not expect the replies for the Multicast Tree Verification Message with Traceable Flag it sent.

4.2.2 Receiving RBridge

The TRILL OAM application layer validates the received OAM frame by examining the presence of the TRILL Alert flag and OAM Ethertype at the end of the Flow Entropy. If Traceable Flag is set to 1 in MTVM, the RBridge validates the frame and analyzes if it is an in-scope RBridge.

If the RBridge Scope TLV is present and the local RBridge nickname is specified in the scope list, the receiving RBridge proceeds with further processing as defined in [Section 4.1.3](#).

If the RBridge Scope TLV is absent, the receiving RBridge SHOULD check the Tree Trace Mode TLV. If E flag is 0, the receiving RBridge proceeds with further processing as defined in [Section 4.1.3](#). If E flag is 1 and the receiving RBridge is an egress BBridge for the specified VLAN/FGL or multicast group, the receiving RBridge proceeds with further processing as defined in [Section 4.1.3](#).

For other cases, the receiving RBridge is not considered as in-scope RBridge and should not perform as per [section 4.2.3](#).

4.2.3 In-Scope RBridges

In-Scope RBridges refers to those should tentatively take actions for MTVM request. They are part of receiving RBridges as described in last sub-section.

- o Optionally include the following TLVs:
 - Previous RBridge Nickname TLV (69)
 - Interface Status TLV (4)
 - Next-Hop RBridge List TLV (70)

- Sender ID TLV (1)
- Multicast Receiver Port Count TLV (71)

o Forward the received message including the TRILL header, the payload and the appended TLVs (if any) to the address specified in Out-of-Band Reply Address TLV. If Out-of-Band Reply Address TLV is not present, either forward it to a system default centralized management server or discard it.

4.2.4 Centralized Management Server

The centralized management server is normally served as an SDN controller, e.g. an Openflow controller. It is up to the implementation how to deal with the collected packets of MTVM with traceable flag from RBridges. The common logic is the centralized management server compares the Session Identification Number and hop count value in TRILL header to trace the path the packet taken along a distribution tree. It can be used to plot the entire tree or pruned tree or to find out who are the edge RBridges connecting users for a specified VLAN/FGL.

5. Security Considerations

For general TRILL fault management security considerations, please refer to [[RFC7455](#)].

6. IANA Considerations

One TLV Type is required to be assigned from the "CFM OAM IETF TLV Types" sub-registry as follows:

Value	Assignment
-----	-----
75	Tree Trace Mode TLV

7. References

7.1 Normative References

- [RFC6325] Perlman, R., et.al. "RBridge: Base Protocol Specification", [RFC 6325](#), July 2011.
- [RFC6439] Eastlake, D. et.al., "RBridge: Appointed Forwarder", [RFC 6439](#), November 2011.

- [RFC6905] Senevirathne, T., Bond, D., Aldrin, S., Li, Y., and R. Watve, "Requirements for Operations, Administration, and Maintenance (OAM) in Transparent Interconnection of Lots of Links (TRILL)", [RFC 6905](#), March 2013.
- [RFC7172] Eastlake 3rd, D., Zhang, M., Agarwal, P., Perlman, R., and D. Dutt, "Transparent Interconnection of Lots of Links (TRILL): Fine-Grained Labeling", [RFC 7172](#), May 2014.
- [RFC7174] Salam, S., Senevirathne, T., Aldrin, S., and D. Eastlake 3rd, "Transparent Interconnection of Lots of Links (TRILL) Operations, Administration, and Maintenance (OAM) Framework", [RFC 7174](#), May 2014,
- [RFC7180] Eastlake 3rd, D., Zhang, M., Ghanwani, A., Manral, V., and A. Banerjee, "Transparent Interconnection of Lots of Links (TRILL): Clarifications, Corrections, and Updates", [RFC 7180](#), May 2014.
- [RFC7455] Senevirathne, T., Finn, N., Salam, S., Kumar, D., Eastlake 3rd, D., Aldrin, S., and Y. Li, "Transparent Interconnection of Lots of Links (TRILL): Fault Management", [RFC 7455](#), March 2015.

7.2 Informative References

- [OpenFlow] OpenFlow Switch Specification Version, March 26, 2015. (<https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.5.1.pdf>)

8. Acknowledgments

TBD

Authors' Addresses

Yizhou Li
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Phone: +86-25-56624629
Email: liyizhou@huawei.com

Donald Eastlake
Huawei R&D USA
155 Beaver Street
Milford, MA 01757 USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Hao Chen
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Email: philips.chenhao@huawei.com

Deepak Kumar
CISCO Systems
510 McCarthy Blvd,
Milpitas, CA 95035, USA

Phone : +1 408-853-9760
Email: dekumar@cisco.com

Sujay Gupta
IP Infusion
RMZ Centennial
Mahadevapura Post
Bangalore - 560048
India

Email: sujay.gupta@ipinfusion.com

