TRILL Working Group
Internet Draft
Intended status: Standard Track

Tissa Senevirathne
Norman Finn
Samer Salam
Deepak Kumar
CISCO

Donald Eastlake
Sam Aldrin
YiZhou Li
Huawei

February 17, 2013

Expires: August 2013

TRILL Fault Management
draft-tissa-trill-oam-fm-01.txt

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), 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/ietf/1id-abstracts.txt

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

This Internet-Draft will expire on August 17, 2013.

Copyright Notice

Copyright (c) 2013 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.

Abstract

TRILL OAM Fault Management solution is presented in this document. Methods proposed in this document follow the IEEE 802.1 CFM framework and reuse OAM tools where possible. Additional messages and TLVs are defined for TRILL specific applications or where different set of information is required than IEEE 802.1 CFM.

Table of Contents

1. Introduction....4
2. Conventions used in this document....4
3. General Format of TRILL OAM frames....5
3.1. Identification of TRILL OAM frames....7
3.2. Use of TRILL OAM Flag....7
3.2.1. Handling of TRILL frames with "A" Flag....8
3.3. Backwards Compatibility Method....8
3.4. OAM Capability Announcement....9
4. TRILL OAM Layering vs. IEEE Layering....10
4.1. Processing at ISS Layer....11
4.1.1. Receive Processing....11
4.1.2. Transmit Processing....11
4.2. End Station VLAN and Priority Processing....11
4.2.1. Receive Processing....11
4.2.2. Transmit Procession....11
4.3. TRLL Encapsulation and De-capsulation Layer....11
4.3.1. Receive Processing for Unicast packets....11
4.3.2. Transmit Processing for unicast packets....12
4.3.3. Receive Processing for Multicast packets....12
4.3.4. Transmit Processing of Multicast packets....13
4.4. TRILL OAM Layer Processing....14
5. Maintenance Associations (MA) in TRILL....15
6. MEP Addressing....16
6.1. Use of MIP in TRILL....19
7. Approach for Backwards Compatibility....21
8. Continuity Check Message (CCM)....22
9. TRILL OAM Message Channel....24

9.1. TRILL OAM Message header.....24

9.2. TRILL OAM Opcodes.....25
9.3. Format of TRILL OAM TLV.....25
9.4. TRILL OAM TLVs.....26
9.4.1. Common TLVs between 802.1ag and TRILL.....26
9.4.2. TRILL OAM Specific TLVs.....26
9.4.2.1. TRILL OAM Application Identifier TLV.....27
9.4.3. Out Of Band Reply Address TLV.....28
9.4.3.1. Diagnostics Label TLV.....29
9.4.3.2. Original Data Payload TLV.....30
9.4.3.3. RBridge scope TLV.....30
9.4.3.4. Previous RBridge nickname TLV.....31
9.4.3.5. Next Hop RBridge List TLV.....31
9.4.3.6. Multicast Receiver Port count TLV.....32
9.4.4. Flow Identifier (flow-id) TLV.....33
10. Loopback Message.....34
10.1.1. Loopback OAM Message format.....34
10.1.2. Theory of Operation.....34
10.1.2.1. Originator RBridge.....34
10.1.2.2. Intermediate RBridge.....35
10.1.2.3. Destination RBridge.....35
11. Path Trace Message.....36
11.1.1. Theory of Operation.....36
11.1.1.1. Originator RBridge.....36
11.1.1.2. Intermediate RBridge.....37
11.1.1.3. Destination RBridge.....38
12. Multi-Destination Tree Verification (MTV) Message.....38
12.1. Multi-Destination Tree Verification (MTV) OAM Message Format.....39
12.2. Theory of Operation.....39
12.2.1. Originator RBridge.....39
12.2.2. Receiving RBridge.....40
12.2.3. In scope RBridges.....40
13. Application of Continuity Check Message (CCM) in TRILL.....41
13.1. CCM Error notification - Method-1.....42
13.2. CCM Error Notification Method-2.....43
13.3. Theory of Operation.....44
13.3.1. Originator RBridge.....44
13.3.2. Intermediate RBridge.....45
13.3.3. Destination RBridge.....45
14. Multiple Fragment Reply.....45
15. Security Considerations.....46
16. Allocation Considerations.....46
16.1. IEEE Allocation Considerations.....46
16.2. IANA Considerations.....46
17. References.....47
17.1. Normative References.....47
17.2. Informative References.....47
18. Acknowledgments.....48

1. Introduction

The general structure of TRILL OAM messages is presented in [TRILLOAMFM]. According to [TRILLOAMFM], TRILL OAM messages consist of five parts: link header, TRILL header, flow entropy, OAM message channel, and link trailer.

The OAM message channel allows defining various control information and carrying OAM related data between TRILL switches, also known as RBridges or Routing Bridges.

The OAM message channel, if defined properly, can be shared between different technologies. A common OAM channel allows a uniform user experience for the customers, savings on operator training, re-use of software code base and faster time to market.

This document uses the message format defined in IEEE 802.1ag Connectivity Fault Management (CFM) [8021Q] as the basis for the TRILL OAM message channel.

The ITU-T Y.1731 standard utilizes the same messaging format as [8021Q] and OAM messages where applicable. In this document, we take a similar stance and propose reusing [8021Q] in TRILL OAM. We assume readers are familiar with [8021Q] and Y1731. Readers who are not familiar with these documents are encouraged to review [8021Q] and Y1731.

2. Conventions used in this document

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].

Acronyms used in the document include the following:

MP - Maintenance Point [TRILLOAMFM]

MEP - Maintenance End Point [TRILLOAMFM] [8021Q]

MIP - Maintenance Intermediate Point [TRILLOAMFM] [8021Q]

MA - Maintenance Association [8021Q] [TRILLOAMFM]

CCM - Continuity Check Message [8021Q]

LBM - Loop Back Message [8021Q]

PTM - Path Trace Message

MTV - Multi-destination Tree Verification Message

OAM - Operations, Administration, and Maintenance [RFC6291]

TRILL - Transparent Interconnection of Lots of Links [RFC6325]

FGL - Fine Grained Label [RFCfgl]

3. General Format of TRILL OAM frames

The TRILL forwarding paradigm allows an implementation to select a path from a set of equal cost paths to forward a packet. Selection of the path of choice is implementation dependent. However, it is a common practice to utilize Layer 2 through Layer 4 information in the frame payload for path selection.

For accurate monitoring and/or diagnostics, OAM Messages are required to follow the same path as corresponding data packets. [TRILLOAMFM] proposes a high-level format of the OAM messages. The details of the TRILL OAM frame format are defined in this document.

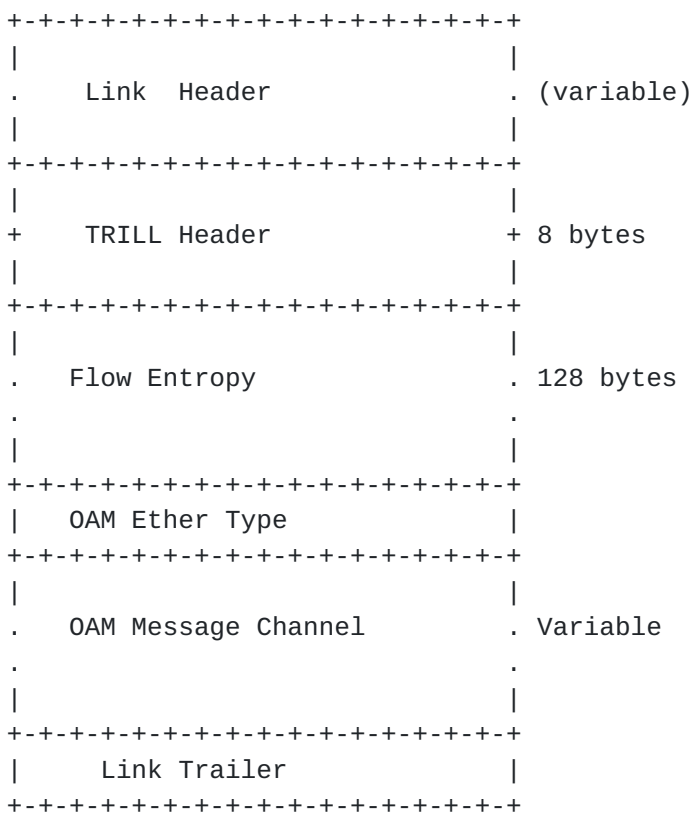

Figure 1 Format of TRILL OAM Messages

Link Header: Media-dependent header. For Ethernet, this includes Destination MAC, Source MAC, VLAN (optional) and EtherType fields.

TRILL Header: Minimum of 8 bytes when the Extended Header is not included [RFC6325]

Flow Entropy: This is a 128-byte fixed size opaque field. The least significant bits of the field MUST be padded with zeros, up to 128 bytes, when the flow entropy is less than 128 bytes. Flow entropy enables emulation of the forwarding behavior of the desired data packets.

OAM Ether Type: OAM Ether Type is 16-bit EtherType that identifies the OAM Message channel which follows. This document specifies using the EtherType allocated for 802.1ag for this purpose. Identifying the OAM Message Channel with a dedicated EtherType allows the easy identification of the beginning of the OAM message channel across multiple standards.

OAM Message Channel: This is a variable size section that carries OAM related information. We propose reusing the message format defined in [8021Q] for this purpose.

Link Trailer: Media-dependent trailer. For Ethernet, this is the FCS (Frame Check Sequence).

3.1. Identification of TRILL OAM frames

TRILL, as originally specified in [RFC6325], did not have a specific flag or a method to identify OAM frames. This document updates RFC6325 to include specific methods to identify TRILL OAM frames. Section 3.2. below explains the details of the method. However, it is important, for backwards compatibility reasons, to define methods of identifying TRILL OAM frames without using these extensions. Section 3.3. presents a set of possible methods for identifying OAM frames without using the proposed extensions of section 3.2. The methods defined in section 3.3. impose limitations on the construction of the flow entropy field of the OAM frames and SHOULD be used for backwards compatibility scenarios only.

3.2. Use of TRILL OAM Flag

The TRILL Header, as defined in [RFC6325], has two reserved bits that are currently unused. RBridges are currently required to ignore these fields. This document specifies use of the reserved bit next to Version field in the TRILL header as the Alert flag. Alert flag will be denoted by 'A. (TISSA: Move to A)

Implementations that follow the extension of using the "A" flag to identify frames MUST exclusively use that flag and methods specified in section 3.2.1. The "A" flag MUST NOT be utilized for forwarding decisions such as the selection of ECMP paths, etc.

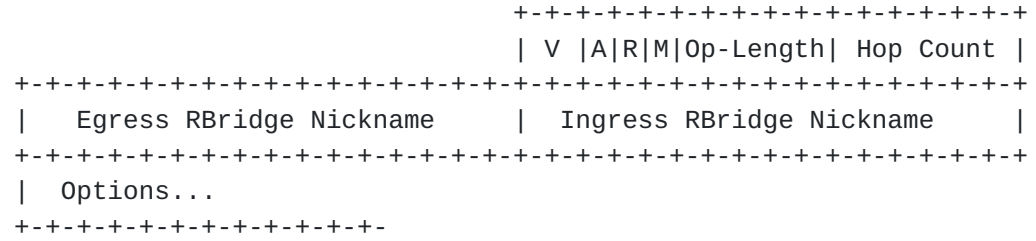

Figure 2 TRILL Header

A (1 bit) - Indicates this is a possible OAM frame and is subject to specific handling as specified in this document.

All other fields carry the same meaning as defined in RFC6325.

3.2.1. Handling of TRILL frames with "A" Flag

Value "1" in the A flag indicates TRILL frames that may qualify as OAM frames. Implementations are further required to validate such frames by comparing the value at the OAM Ether Type (Figure 1) location with the CFM EtherType "0x8902" [8021Q]. If the value matches, such frames are identified as TRILL OAM frames and SHOULD be processed as discussed in Section 4.

3.3. Backwards Compatibility Method

For unicast frames, TRILL OAM packets are identified by its TRILL egress nickname and the presence of either Reserved Inner.MacSA (TBD) or OAM Ether Type 0x8902 [8021Q].

For multicast frames, TRILL OAM packets are identified by either OAM EtherType 0x8902 [8021Q] or Reserved Inner.MacSA (TBD) .

The following table summarizes the identification of different OAM frames from data frames.

Flow Entropy	Inner MacSA	OAM Ether Type	Egress nickname
unicast L2	N/A	Match	Match
Multicast L2	N/A	Match	N/A
Unicast IP	Match	N/A	Match
Multicast IP	Match	N/A	N/A
Notification	N/A	Match	Match

Figure 3 Identification of TRILL OAM Frames

3.4. OAM Capability Announcement

Any given TRILL RBridge can be one of: OAM incapable OR OAM capable with new extensions OR OAM capable with backwards-compatible method. The OAM request originator, prior to origination of the request is required to identify the OAM capability of the target and generate the appropriate OAM message.

We propose to utilize the capability flags defined in TRILL version sub-TLV (TRILL-VER) [rfc6326bis]. The following Flags are defined:

O - OAM Capable

B - Backwards Compatible.

A capability announcement, with O Flag set to 1 and B flag set to 1, indicates that the implementation is OAM capable but utilize backwards compatible method defined in section 3.3. A capability announcement, with O Flag set to 1 and B flag set to 0, indicates that the implementation is OAM capable and utilizes the method specified in section 3.2.

When O Flag is set to 0, the announcing implementation is considered not capable of OAM and B flag is ignored on the receiving side.

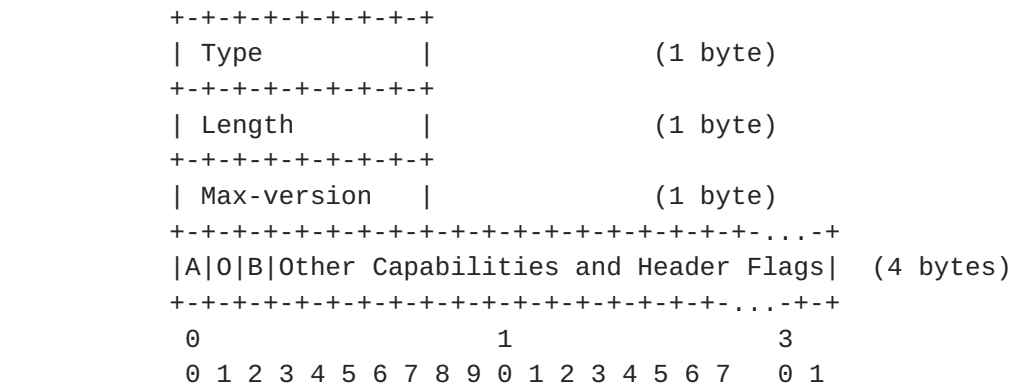

Figure 4 TRILL-VER sub-TLV [rfc6326bis] with O and B flags

NOTE: Bit position of O and B flags in the TRILL-VER sub-TLV are presented above as an example. Actual positions of the flags will be determined by TRILL WG and IANA and future revision of this document will be updated to include the allocations.

4. TRILL OAM Layering vs. IEEE Layering

In this section we present the placement of the TRILL OAM shim within the IEEE 802.1 layers. The processing of both the Transmit and Receive directions is explained.

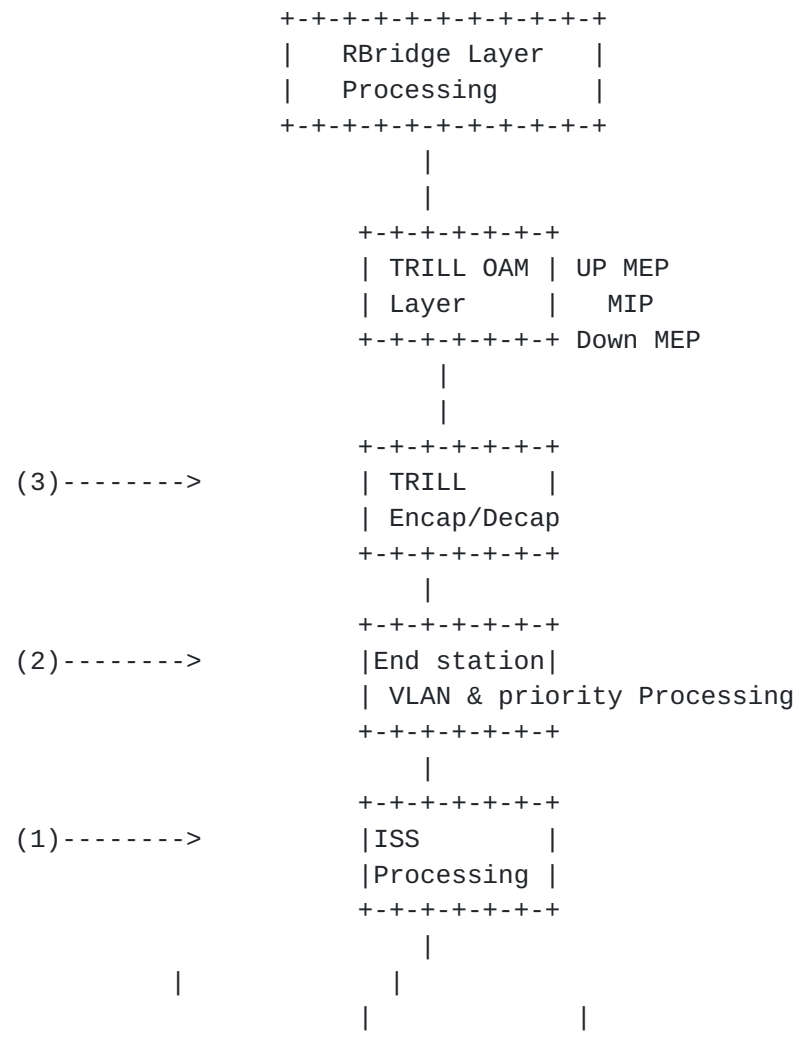

Figure 5 Placement of TRILL MP within IEEE 802.1

[RFC6325] Section 4.6 provides a detail explanation of frame processing. Please refer to [RFC6325] for processing scenarios not covered herein.

4.1. Processing at ISS Layer

4.1.1. Receive Processing

The ISS Layer receives an indication from the port. It extracts DA, SA and marks the remainder of the payload as M1. ISS Layer passes on (DA,SA,M1) as an indication to the higher layer.

For TRILL frames, this is Outer DA and Outer SA. M1 is the remainder of the packet from the VLAN EtherType onwards.

4.1.2. Transmit Processing.

ISS layer receives indication from the higher layer that contains (DA,SA,M1). It constructs an Ethernet frame and passes down to the port.

4.2. End Station VLAN and Priority Processing

4.2.1. Receive Processing

Receives (DA,SA,M1) indication from ISS Layer. Extracts the VLAN from the M1 part of the received indication and construct (DA,SA,VLAN,PRI,M2). VLAN+PRI+M2 map to M1 in the received indication. Pass (DA,SA,VLAN,PRI,M2) to the TRILL encap/decap procession layer.

4.2.2. Transmit Procession

Receive (DA,SA,VLAN+PRI,M2) indication from TRILL encap/decap processing layer. Merge VLAN, M2 to from M1. Pass down (DA,SA,M1) to the ISS processing Layer.

4.3. TRLL Encapsulation and De-capsulation Layer

4.3.1. Receive Processing for Unicast packets

Receive indication (DA,SA,VLAN, PRI, M2) from End Station VLAN and Priority Processing Layer.

- o If DA matches Local DA and Frame is of TRILL EtherType
  - . Discard DA, SA, VLAN, PRI. From M2, derive (TRILL-HDR, iDA, iSA, i-VL, M3)
  - . If TRILL nickname is Local and TRILL-OAM Flag is set

    Pass on to OAM processing

. Else pass on (TRILL-HDR, iDA, iSA, i-VL, M3) to RBridge Layer

o If DA matches local DA and EtherType is not TRILL type

. Discard frame

o If DA does not match and port is Appointed Forwarder and EtherType is not TRILL

. Insert TRILL-Hdr and send (TRILL-HDR, iDA,iSA,i-VL, M3) indication to RBridge Layer <- This is the edge function

4.3.2. Transmit Processing for unicast packets

o Receive indication (TRILL-HDR, iDA, iSA, iVL, M3) from RBridge Layer

o If egress TRILL nickname is local

o If port is Appointed Forwarder and (TRILL Alert Flag set and OAM EtherType present) then

. Strip TRILL-HDR and construct (DA, SA, VLAN, M2)

o Else

. Set discard flag

o If egress TRILL nickname is not local

o Insert Outer DA, Outer SA, Outer VLAN, TRILL EtherType and construct (DA,SA,VLAN,M2). Where M2 is (TRILL-HDR, iDA, iSA, iVL, M)

o Else set the discard flag

o If discard flag is false forward (DA,SA,V,M2) to the VLAN End Station processing Layer. Otherwise, discard the packet.

4.3.3. Receive Processing for Multicast packets

o Receive (DA,SA,V,M2) from VLAN end station processing layer

o If the DA matches the Well-known TRILL multicast MAC address and Ethertype of the frame is TRILL

o Strip DA,SA and V. From M2, construct (TRILL-HDR, iDA, iSA, iVL and M3).

- o If TRILL OAM Flag is set and Ether Type OAM is present at the end of Flow entropy
  - . Perform OAM Processing
- o Else extract the TRILL header, inner MAC addresses and inner VLAN and pass indication (TRILL-HDR, iDA, iSA, iVL and M3) to TRILL RBridge Layer

- o If the DA matches the well-known TRILL multicast MAC address but Ethertype is not TRILL
  - o Discard the packet
- o If the DA does not match the well-known TRILL multicast MAC address and Ether Type is not TRILL type
  - o Insert TRILL-HDR and construct (TRILL-HDR, iDA, iSA, IVL, M3)
  - o Pass the (TRILL-HDR, iDA, iSA, IVL, M3) to RBridge Layer
- o Else
  - o Discard the packet

4.3.4. Transmit Processing of Multicast packets

- o Receive indication (TRILL-HDR, iDA, iSA, iVL, M3) from RBridge layer.
- o If TRILL-HDR multicast flag set and TRILL-HDR Alert flag set and OAM EtherType present then:
  - o (DA,SA,V,M2) by inserting TRILL ODA, OSA, O-VL and TRILL ether type. M2 here is (Ethertype TRILL, TRILL-HDR, iDA, iSA, iVL, M)

    NOTE: Second copy of native format is not made.
- o Else If TRILL-HDR multicast flag set and Alert flag not set
  - o If the port is appointed Forwarder Strip TRILL-HDR, iSA, iDA, iVL and construct (DA,SA,V,M2) for native format.
  - o Make a second copy (DA,SA,V,M2) by inserting TRILL ODA, OSA, O-VL and TRILL ether type. M2 here is (Ethertype TRILL, TRILL-HDR, iDA, iSA, iVL, M)

- o Else unicast packets as defined in section 4.3.2.

- o Pass the indication (DA,SA,V,M2) to End Station VLAN processing layer.

4.4. TRILL OAM Layer Processing

TRILL OAM Processing Layer is located between the TRILL Encapsulation and De-capsulation layer and RBridge Layer. It performs 1. Identification of OAM frames that need local processing 2. Perform OAM processing or redirect to the CPU for OAM processing.

- o Receive indication (TRILL-HDR, iDA, iSA, iVL, M3) from RBridge layer.

- o If the TRILL Multicast Flag is set and TRILL Alert Flag is set and TRILL OAM EtherType is present then
  - o If MEP or MIP is configured on the inner VLAN of the packet then
    - . discard packets that have MD-LEVEL Less than that of the MEP or packets that does not have MD-LEVEL present (e.g due to packet truncation).
    - . If MD-LEVEL matches MD-LEVEL of the MEP then
      - . Re-direct to OAM Processing (Do not forward further)
    - . If MD-LEVEL matches MD-LEVEL of MIP then
      - . Make a Copy for OAM processing and continue

- o Else if TRILL Alert Flag is set and TRILL OAM EtherType is present then
  - o If MEP or MIP is configured on the inner VLAN of the packet then
    - . discard packets that have MD-LEVEL not present or MD-LEVEL is Less than the that of the MEP.
    - . If MD-LEVEL matches MD-LEVEL of the MEP then
      - . Re-direct to OAM Processing (Do not forward further)
    - . If MD-LEVEL matches MD-LEVEL of MIP then
    - . Make a Copy for OAM processing and continue

- o Else // Non OAM l Packet
  - o Continue

- o Pass the indication (DA,SA,V,M2) to End Station VLAN processing layer.

NOTE: In the Received path, processing above compares against Down MEP and MIP Half functions. In the transmit processing it compares against Up MEP and MIP Half functions.

Appointed Forwarder is a Functionality that TRILL Encap/De-Cap layer performs. TRILL Encap/De-cap Layer is responsible for prevention of leaking of OAM packets as native frames.

5. Maintenance Associations (MA) in TRILL

[8021Q] defines a maintenance association as a logical relationship between a group of nodes. Each Maintenance Association (MA) is identified with a unique MAID of 48 bytes [8021Q]. CCM and other related OAM functions operate within the scope of an MA. The definition of MA is technology independent. Similarly it is encoded within the OAM message, not on the technology dependent portion of the packet. Hence we propose to utilize the MAID as defined in [8021Q]. This also allows us to utilize CCM and LBM messages defined in [8021Q], as is.

In TRILL, an MA may contain two or more RBridges (MEPs). For unicast, it is likely that the MA contains exactly two MEPs that are the two end-points of the flow. For multicast, the MA may contain two or more MEPs.

For TRILL, in addition to all of the standard 802.ag MIB definitions, each MEP's MIB contains one or more flow entropy definitions corresponding to the set of flows that the MEP monitors.

We propose to augment the [8021Q] MIB to add the TRILL specific information. Figure 6, below depicts the augmentation of the CFM MIB to add the TRILL specific Flow Entropy.

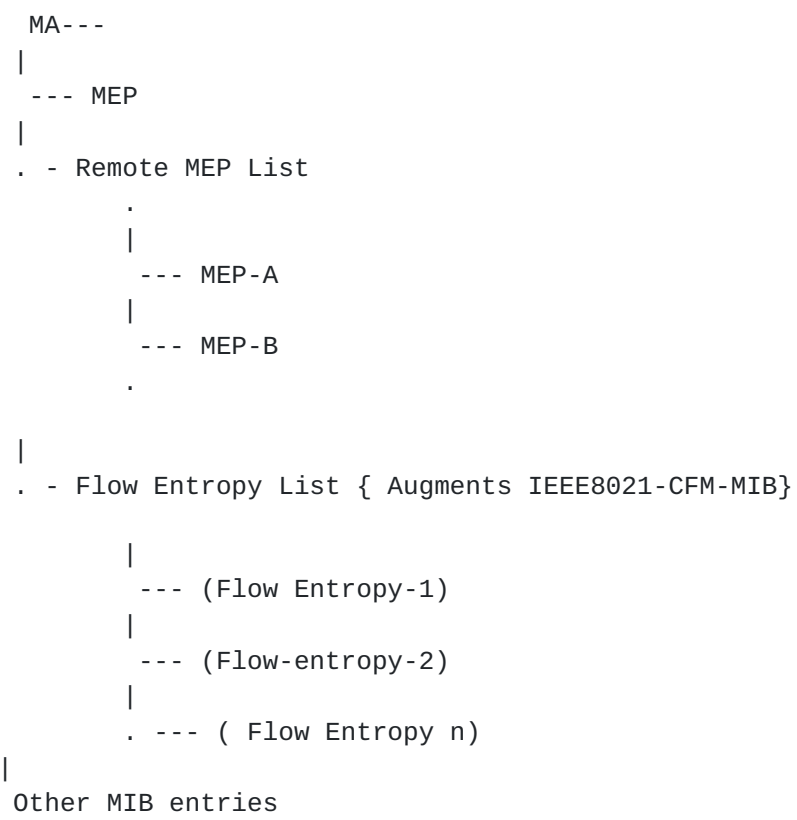

Figure 6 Correlation of TRILL augmented MIB

6. MEP Addressing

In IEEE 802.1ag [8021Q], OAM messages address the target MEP by utilizing a unique MAC address. In TRILL, for qualifying OAM packets, we propose to use a combination of the egress RBridge nickname and Inner VLAN/FGL to address the MEP.

At the MEP, OAM packets go through a hierarchy of op-code de-multiplexers. The op-code de-multiplexers channel the incoming OAM packets to the appropriate message processor (e.g. LBM) The reader may refer to Figure 7 below for a visual depiction of these different de-multiplexers.

1. Identify the packets that need OAM processing at the Local RBridge Section 4.

a. Identify the MEP that is associated with the Inner VLAN.

2. MEP first validate the MD-LEVEL and then

    a. Redirect to MD-LEVEL De-multiplexer

3. MD-LEVEL de-multiplexer compares the MD-Level of the packet against the MD level of the local MEPs of a given MD-Level on the port (Note: there can be more than one MEP at the same MD-Level but belonging to different MAs)

    a. If the packet MD-LEVEL is equal to the configured MD-LEVEL of the MEP, then pass to the Opcode de-multiplexer

    b. If the packet MD-LEVEL is less, then the configured MD-LEVEL of the MEP discard the packet

    c. If the packer MD-LEVEL is greater, then the configured MD-LEVEL of the MEP pass on to the next higher MD-LEVEL de-multiplexer, if available. Otherwise, if no such higher MD-LEVEL de-multiplexer exists then forward the packet as normal data.

4. Opcode De-multiplexer compares the opcode in the packet with supported opcodes

    a. If Op-code is CCM, LBM,LBR, PTM,PTR, MTVM, MTVR, then pass on to the correct Processor

    b. If Op-code is Unknown, then discard.

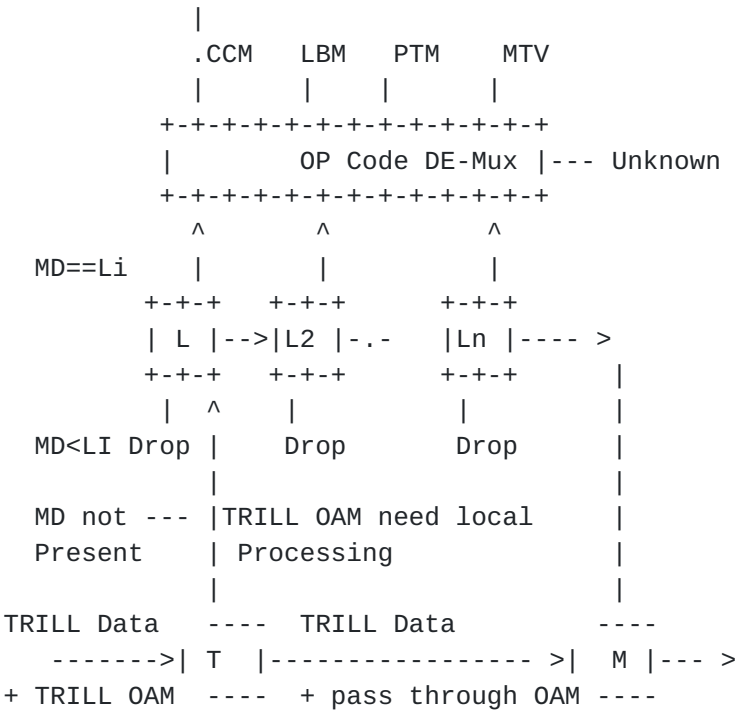

Figure 7 OAM De-Multiplexers at MEP for active SAP

T : Denotes Tap, that identifies OAM frames that need local processing. These are the packets with OAM flag set AND OAM Ether type is present after the flow entropy of the packet

M : Is the post processing merge, merges data and OAM messages that are pass through. Additionally, Merge component ensure, as explained earlier, OAM packets are not forwarded out as native frames.

L : Denotes MD-Level processing. Packets with MD-Level less than the Level will be dropped. Packets with equal MD-Level are passed on to the opcode de-multiplexer. Others are passed on to the next level MD processors or eventually to the merge point (M).

NOTE: LBM, MTV and PT are not subject to MA de-multiplexers. These packets do not have an MA encoded in the packet. Adequate response can be generated to these packets, without loss of functionality, by any of the MEP present on that interface or an entity within the RBridge.

6.1. Use of MIP in TRILL

Maintenance Intermediate Points (MIP) are mainly used for fault isolation. Link Trace Messages in [8021Q] utilize a well-known multicast MAC address and MIPs generate responses to Link Trace messages. Response to Link Trace messages or lack thereof can be used for fault isolation in TRILL.

As explained in section 11. , we propose to use a hop-count expiry approach for fault isolation and path tracing. The approach is very similar to the well-known IP trace-route approach. Hence, explicit addressing of MIPs is not required for the purpose of fault isolation.

Any given RBridge can have multiple MIPs located within a interface. As such, a mechanism is required to identify which MIP should respond or to an incoming OAM message.

We propose to use the same approach as presented above for MEPs with some variations. It is important to note that "M", merge block of MIP does not prevent OAM packets leaking out as native frames. On edge interfaces, MEPs MUST be configured to prevent the leaking of TRILL OAM packets out of the TRILL Campus.

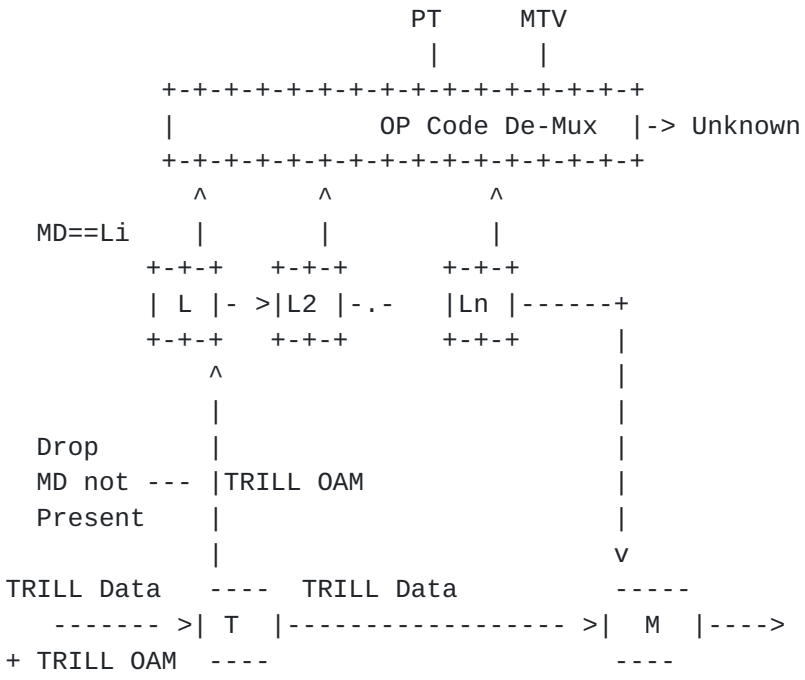

Figure 8 OAM De-Multiplexers at MIP for active SAP

T: TAP processing for MIP. All packets with OAM flag set are captured.

L : MD Level Processing, Packet with matching MD Level are "copied" to the Opcode de-multiplexer and original packet is passed on to the next MD level processor. Other packets are simply passed on to the next MD level processor, without copying to the OP code de-multiplexer.

M : Merge processor, merge OAM packets to be forwarded along with the data flow.

Packets that carry Path Trace (PTM) or Multi-destination Tree Verification (MTV) OpCode are passed on to the respective processors.

Packets with unknown OpCodes are counted and discarded.

7. Approach for Backwards Compatibility

Methodology presented in this document is in-line with the [8021Q] framework or provide fault management coverage. However, in practice, some platforms may not have the required capabilities to support some of the proposed techniques. In this section, we present a method that allows RBridges, which do not have the required hardware capabilities, to participate in the proposed OAM solution.

For backwards compatibility, we propose to locate MEPs and MIPs in the CPU. This will be referred to as the "central brain" model as opposed to "port brain" model.

In the "central brain" model, an RBridge using either ACLs or some other method forwards qualifying OAM messages to the CPU. The CPU then performs the required processing and multiplexing to the correct MP (Maintenance Point).

Additionally, RBridges MUST have the capability to prevent the leaking of OAM packets, as specified in [TRILLOAMREQ] and in the Transmission processing in Figure 9.

```
Receiver Processing:

If (M==1 && F==1) then
   Copy to CPU and Forward normally as defined in RFC 6325
Else if (M==0 && F==1 && egress nickname is the processing RBridge)
then
   Forward to CPU BUT DO NOT forward along the data plane

Else
   Forward as defined in [RFC6325]
End;

Transmit Processing:

If (F==1) then
  Forward as defined in [RFC6325] BUT Do not de-capsulate and
forward as a native frame
Else
  Forward as defined in [RFC6325]
```

Figure 9 Pseudo code for Backward compatible Processing

[8021Q] requires that the MEP filters or pass through OAM messages based on the MD-Level. The MD-Level is embedded deep in the OAM message. Hence, conventional methods of frame filtering may not be able to filter frames based on the MD-Level. As a result, OAM messages, that must be dropped due to MD level mismatch, may leak in to a TRILL domain with different MD-Level.

This leaking may not cause any functionality loss. Receiving MEP/MIP is required to validate the MD-level prior to acting on the message. Any frames received with an incorrect MD-Level will be dropped.

Generally, TRILL campuses are managed by a single operator, hence there is no risk of security exposure. However, in the event of multi operator deployments, operators should be aware of possible exposure of device specific information and appropriate measures must be taken.

It is also important to note that the MPLS OAM [RFC4379] framework does not include the concept of domains and OAM filtering based on operators. It is our opinion that the lack of OAM frame filtering based on domains does not introduce significant functional deficiency or security risk.

8. Continuity Check Message (CCM)

CCM are used to monitor connectivity and configuration errors. [8021Q] monitors connectivity by listening to periodic CCM messages received from its remote MEP partners in the MA. An [8021Q] MEP identifies cross-connect errors by comparing the MAID in the received CCM message with the MEP's local MAID. The MAID [8021Q] is a 48 byte field that is technology independent. Similarly, the MEPID is a 2 byte field that is independent of the technology. Given this generic definition of CCM fields, CCM as defined in [8021Q] can be utilized in TRILL with no changes. TRILL specific information may be carried in CCMs when encoded using TRILL specific TLVs or sub-TLVs. This is possible since CCMs may carry optional TLVs.

Unlike classical Ethernet environments, TRILL contains multipath forwarding. The path taken by a packet depends on the payload of the packet. The Maintenance Association identifies the interested end-points (MEPs) of a given monitored path. For unicast there are only two MEPs per MA. For multicast there can be two or more MEPs in the MA. Within the MA, we propose to define the entropy values of the monitored flows. CCM transmit logic will utilize these flow entropy values when constructing the CCM packets. Please see section 13. later in the document for the theory of operation of CCM.

We propose to augment the MIB of [8021Q] with definition of flow-

entropy. Please see [TRILLOAMMIB] for definition of these and other

TRILL related OAM MIB definitions. Below Figure depicts the correlation between MA, CCM and proposed flow-entropy.

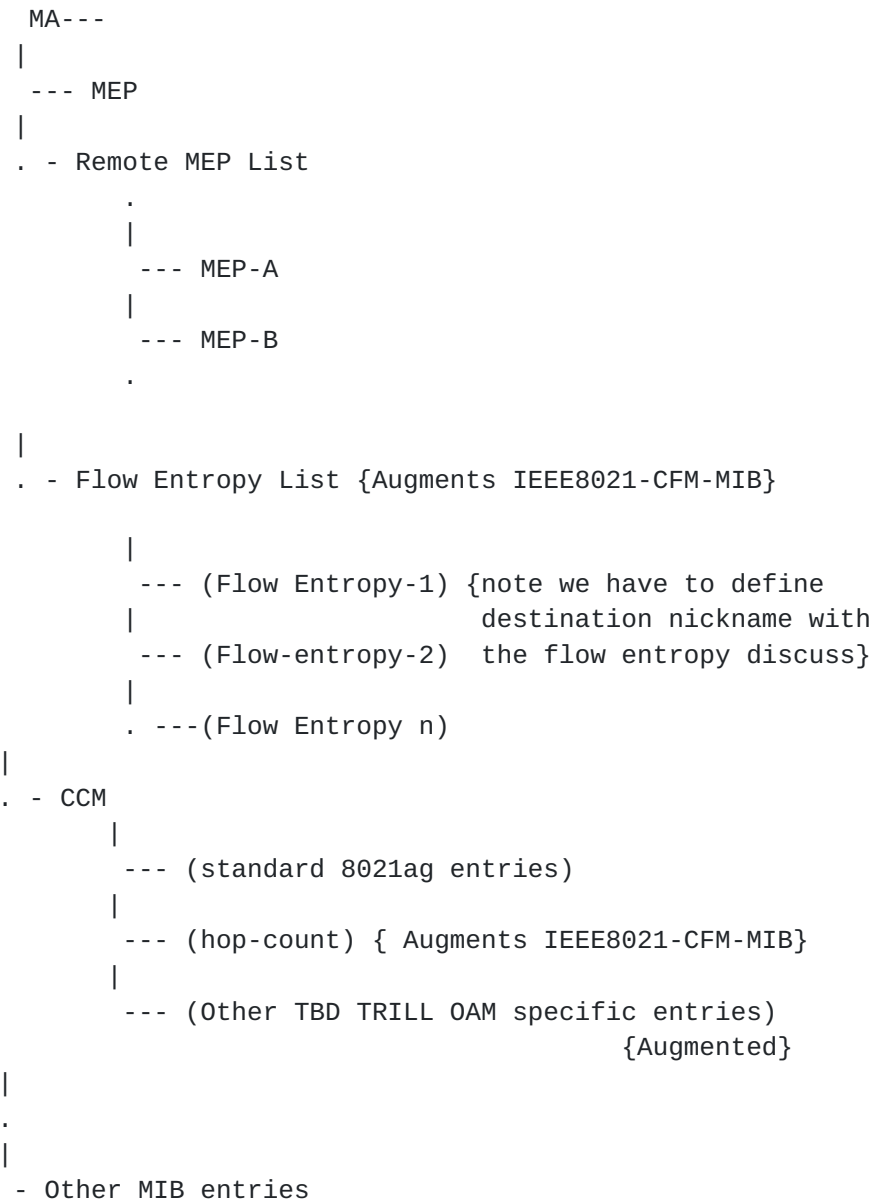

Figure 10 Augmentation of CCM MIB in TRILL

In a multi-pathing environment, a Flow - by definition - is

unidirectional. A question may arise as to what flow entropy to be

used in the response. CCMs are unidirectional and have no explicit reply; as such, the issue of the response flow entropy does not arise. In the transmitted CCM, each MEP reports local status using the Remote Defect Indication (RDI) flag. Additionally, a MEP may raise SNMP TRAPs [TRLLOAMMIB] as Alarms when a connectivity failure occurs.

9. TRILL OAM Message Channel

The TRILL OAM Message Channel can be divided into two parts: TRILL OAM Message header and TRILL OAM Message TLVs. Every OAM Message MUST contain a single TRILL OAM message header and a set of one or more specified OAM Message TLVs.

9.1. TRILL OAM Message header

As discussed earlier, we propose to use the message format defined in IEEE 802.1ag. We believe a common messaging framework between [8021Q], TRILL and other similar standards such as Y.1731 can be accomplished by re-using the OAM message header defined in [8021Q].

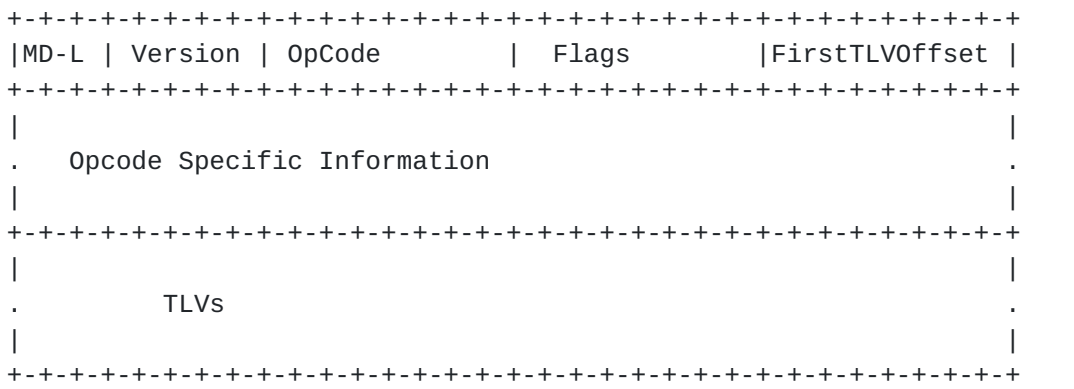

Figure 11 OAM Message Format

- MD-L: Maintenance Domain Level (3 bits). Identifies the maintenance domain level. For TRILL, this MAY be always set to zero. However, in multilevel TRILL, backbone MAY be of a different MD-LEVEL. (Please refer to [8021Q] for the definition of MD-Level)

- Version: Indicates the version (5 bits). As specified in [8021Q].

- o Flags: Includes operational flags (1 byte). The definition of flags is Opcode-specific and is covered in the applicable sections.

- o FirstTLVOffset: Defines the location of the first TLV, in bytes, starting from the end of the FirstTLVOffset field (1 byte). (Refer to [8021Q] for the definition of the FirstTLVOffset.)

MD-L, Version, Opcode, Flags and FirstTLVOffset fields collectively are referred to as the OAM Message Header.

The Opcode specific information section of the OAM Message may contain Session Identification number, time-stamp, etc.

9.2. TRILL OAM Opcodes

The following Opcodes are defined for TRILL. Each of the Opcodes defines a separate TRILL OAM message. Details of the messages are presented in the related sections.

TRILL OAM Message Opcodes:

TBD-64 : Path Trace Reply
TBD-65 : Path Trace Message
TBD-66 : Notification Message
TBD-67 : Multicast Tree Verification Reply
TBD-68 : Multicast Tree Verification Message

9.3. Format of TRILL OAM TLV

We propose to use the same TLV format as defined in section 21.5.1 of [8021Q]. The following figure depicts the general format of a TRILL OAM TLV:

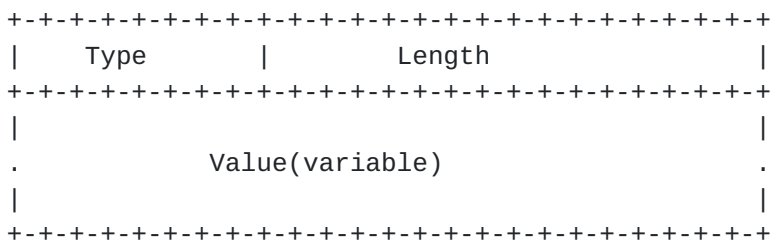

Figure 12 TRILL OAM TLV

Type (1 octet) : Specifies the Type of the TLV (see sections 9.4. for TLV types).

Length (2 octets) : Specifies the length of the 'Value' field in octets. Length of the 'Value' field can be either zero or more octets.

Value (variable): The length and the content of this field depend on the type of the TLV. Please refer to applicable TLV definitions for the details.

Semantics and usage of Type values allocated for TRILL OAM purpose are defined by this document and other future related documents.

9.4. TRILL OAM TLVs

In this section we define TRILL related TLVs. We propose to re-use [8021Q] defined TLVs where applicable. Types 32-63 are reserved for ITU-T Y.1731. We propose to reserve Types 64-95 for TRILL OAM TLVs.

9.4.1. Common TLVs between 802.1ag and TRILL

The following TLVs are defined in [8021Q]. We propose to re-use them where applicable. The format and semantics of the TLVs are as defined in [8021Q]. NOTE: Presented within brackets is the corresponding Type defined in [8021Q].

1. End TLV (0)
2. Sender ID TLV (1)
3. Port Status TLV (2)
4. Data TLV (3)
5. Interface Status TLV (4)
6. Reply Ingress TLV (5)
7. Reply Egress TLV (6)
8. LTM Egress Identifier TLV (7)
9. LTR Egress Identifier TLV (8)
10. Reserved (9-30)
11. Organization specific TLV (31)

9.4.2. TRILL OAM Specific TLVs

As indicated above, Types 64-95 will be requested to be reserved for TRILL OAM purposes. Listed below is a summary of TRILL OAM TLVs and their corresponding codes. Format and semantics of TRILL OAM TLVs are defined in subsequent sections.

1. TRILL OAM Application Identifier (TBD-TLV-64)
2. Out of Band IP Address (TBD_TLV-65)

3. Diagnostic VLAN (TBD-TLV-66)

4. RBridge Scope (TBD-TLV-67)
5. Original Payload (TBD-TLV-68)
6. Previous RBridge Nickname(TBD-TLV-69)
7. TRILL Next Hop RBridge List (ECMP) (TBD-TLV-70)
8. Multicast Receiver Availability (TBD-TLV-71)
9. Flow Identifier (TBD-TLV-72)
10. Reserved (TBD-TLV-72 to TBD-TLV-95)

9.4.2.1. TRILL OAM Application Identifier TLV

TRILL OAM Application Identifier TLV carries TRILL OAM application specific information. The TRILL OAM Application Identifier TLV MUST always be present and MUST be the first TLV in TRILL OAM messages. Messages that do not include the TRILL OAM Application Identifier TLV as the first TLV MUST be discarded by an RBridge, unless that RBridge is running Ethernet CFM.

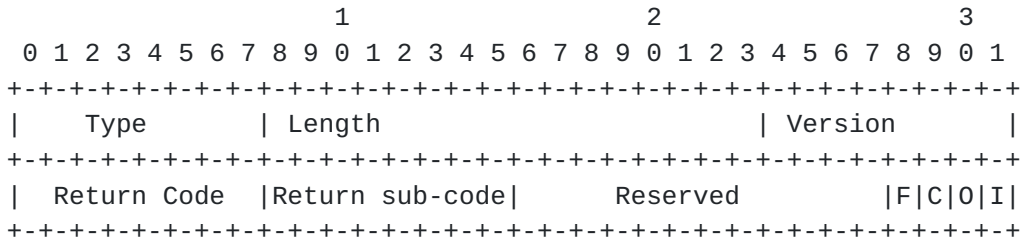

Figure 13 TRILL OAM Message TLV

Type (1 octet) = 64 indicate that this is the TRILL OAM Version

Length (2 octets) = 6

TRILL OAM Version (1 Octet), currently set to zero. Indicates the TRILL OAM version. TRILL OAM version can be different than the [8021Q] version.

Return Code (1 Octet): Set to zero on requests. Set to an appropriate value in response or notification messages.

Return sub-code (1 Octet): Return sub-code is set to zero on transmission of request message. Return sub-code identifies categories within a specific Return code. Return sub-code MUST be interpreted within a Return code.

Reserved: set to zero on transmission and ignored on reception.

F (1 bit) : Final flag, when set, indicates this is the last response.

C (1 bit): Label error (VLAN/Label mapping error), if set indicates that the label (VLAN/FGL) in the flow entropy is different than the label included in the diagnostic TLV. This field is ignored in request messages and MUST only be interpreted in response messages.

O (1 bit) : If set, indicates, OAM out-of-band response requested.

I (1 bit) : If set, indicates, OAM in-band response requested.

NOTE: When both O and I bits are set to zero, indicates that no response is required (silent mode). User MAY specify both O and I or one of them or none.

9.4.3. Out Of Band Reply Address TLV

Out of Band Reply Address TLV specifies the address to which an out of band OAM reply message MUST be sent. When O bit in the TRILL Version TLV is not set, Out of Band Reply Address TLV is ignored.

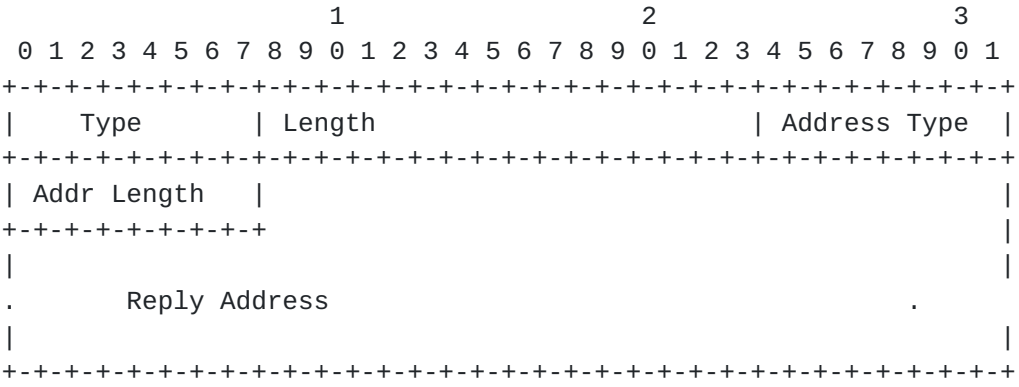

Figure 14 Out of Band IP Address TLV

Type (1 octet) = 64

Length (2 octets) = Variable. Minimum length is 2.

Address Type (1 Octet): 0 - IPv4. 1 - IPv6. 2- TRILL RBridge nickname. All other values reserved.

Addr Length (1 Octet). 4 - IPv4. 16 - IPv6, 2 - TRILL RBRidge nickname.

Reply Address (variable): Address where the reply needed to be sent. Length depends on the address specification.

9.4.3.1. Diagnostics Label TLV

Diagnostic label specifies the data label (VLAN or FGL) in which the OAM messages are generated. Receiving RBridge MUST compare the data label of the Flow entropy to the data label specified in the Diagnostic Label TLV. Label Error Flag in the response (TRILL OAM Message Version TLV) MUST be set when the two VLANs do not match.

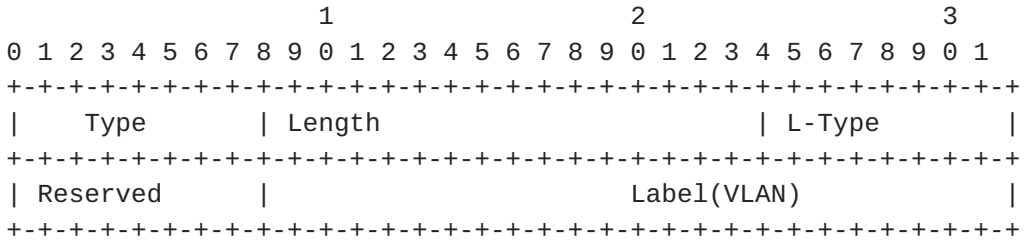

Figure 15 Diagnostic VLAN TLV

Type (1 octet) = 65 indicates that this is the TRILL Diagnostic VLAN TLV

Length (2 octets) = 5

L-Type (Label type, 1 octet)

0- indicate 802.1Q 12 bit VLAN.

1 - indicate TRILL 24 bit fine grain label

Label (24 bits): Either 12 bit VLAN or 24 bit fine grain label.

RBridges do not perform Label error checking when Label TLV is not included in the OAM message. In certain deployment intermediate devices may perform label (VLAN) translation. In such scenarios, originator should not include the diagnostic Label TLV in OAM messages. Inclusion of diagnostic TLV will generated unwanted label error notifications.

9.4.3.2. Original Data Payload TLV

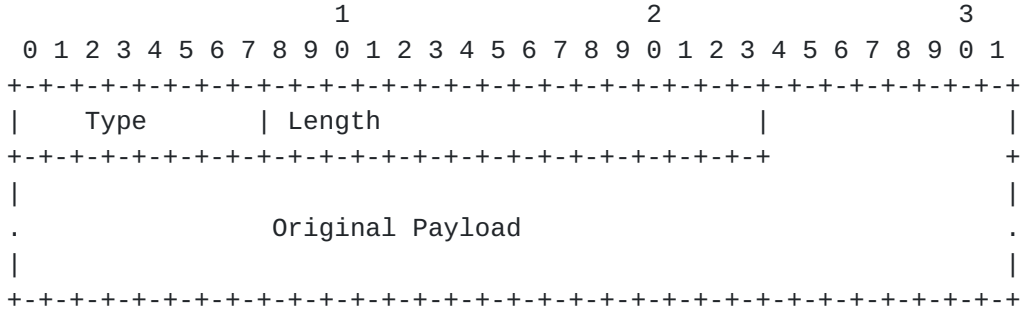

Figure 16 Out of Band IP Address TLV

Length (2 octets) = variable

9.4.3.3. RBridge scope TLV

RBridge scope TLV identifies nicknames of RBridges from which a response is required. RBridge scope TLV is only applicable to Multicast Tree Verification messages. This TLV SHOULD NOT be included in other messages. Receiving RBridges MUST ignore this TLV on messages other than Multicast Verification Message.

Each TLV can contain up to 255 nicknames of in scope RBridges. A Multicast Verification Message may contain multiple "RBridge scope TLVs", in the event that more than 255 in scope RBridges need to be specified.

Absence of the "RBridge scope TLV" indicates that a response is needed from all the RBridges. Please see section 12. for details.

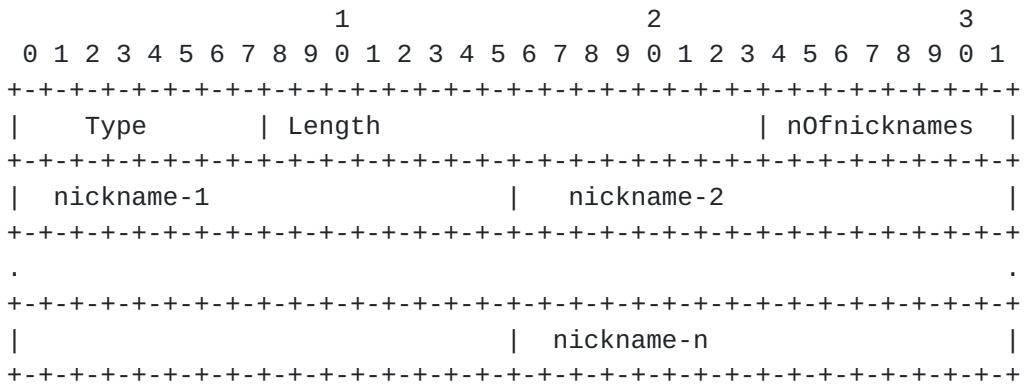

Figure 17 RBridge Scope TLV

Type (1 octet) = 67 indicates that this is the "RBridge scope TLV"

Length (2 octets) = variable. Minimum value is 2.

Nickname (2 octets) = 16 bit RBridge nickname.

9.4.3.4. Previous RBridge nickname TLV

"Previous RBridge nickname TLV" identifies the nickname or nicknames of the upstream RBridge. [RFC6325] allows a given RBridge to hold multiple nicknames.

"Upstream RBridge nickname TLV" is an optional TLV. Multiple instances of this TLV MAY be included when an upstream RBridge is represented by more than 255 nicknames (highly unlikely).

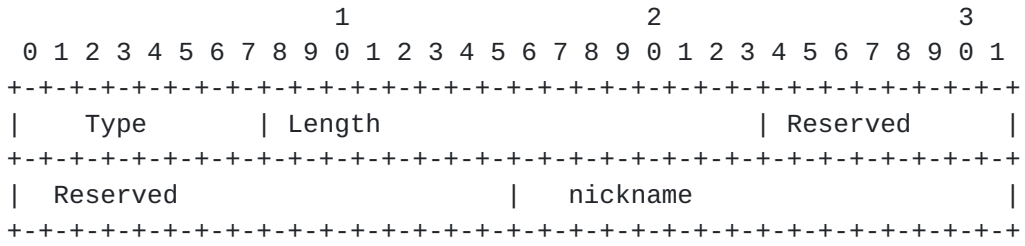

Figure 18 Upstream RBridge nickname TLV

Type (1 octet) = 69 indicates that this is the "Upstream RBridge nickname"

Length (2 octets) = 4.

Nickname (2 octets) = 16 bit RBridge nickname.

9.4.3.5. Next Hop RBridge List TLV

"Next Hop RBridge List TLV" identifies the nickname or nicknames of the downstream next hop RBridges. [RFC6325] allows a given RBridge to have multiple Equal Cost Paths to a specified destination. Each next hop RBridge is represented by one of its nicknames.

"Next Hop RBridge List TLV" is an optional TLV. Multiple instances of this TLV MAY be included when there are more than 255 Equal Cost Paths to the destination.

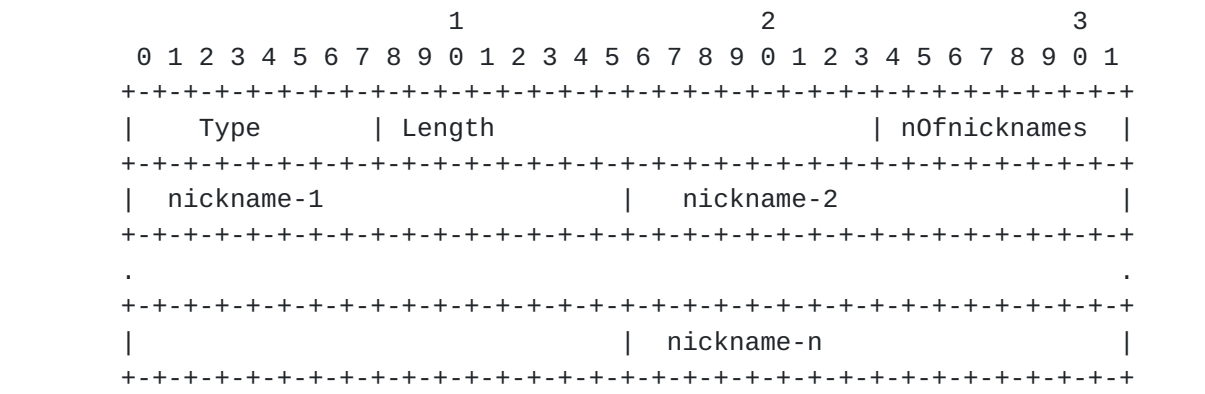

Figure 19 Next Hop RBridge List TLV

Type (1 octet) = 70 indicates that this is the "Next nickname"

Length (2 octets) = variable. Minimum value is 2.

Nickname (2 octets) = 16 bit RBridge nickname.

9.4.3.6. Multicast Receiver Port count TLV

"Multicast Receiver Port Count TLV" identifies the number of ports interested in receiving the specified multicast stream within the responding RBridge on the VLAN specified by the Diagnostic VLAN TLV.

Multicast Receiver Port countis an Optional TLV.

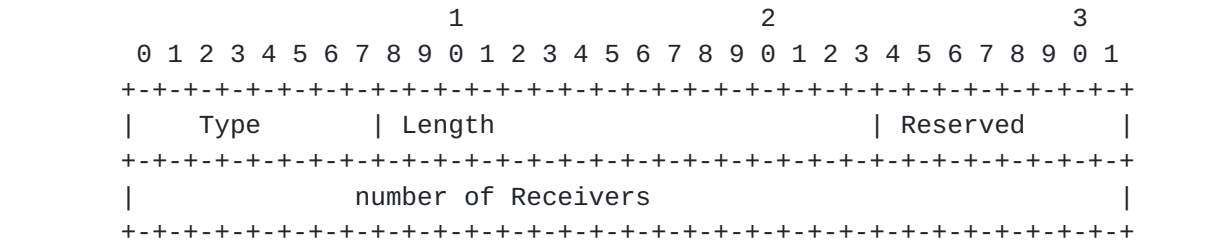

Figure 20 Multicast Receiver Availability TLV

Type (1 octet) = 71 indicates that this is the "Multicast Availability TLV"

Length (2 octets) = 5.

Number if Receivers (4 octets) = Indicates the number of Multicast receivers available on the responding RBridge on the VLAN specified by the diagnostic VLAN.

9.4.4. Flow Identifier (flow-id) TLV

Flow Identifier (flow-id) uniquely identifies a specific flow. The flow-id value is unique per MEP and needed to be interpreted as such.

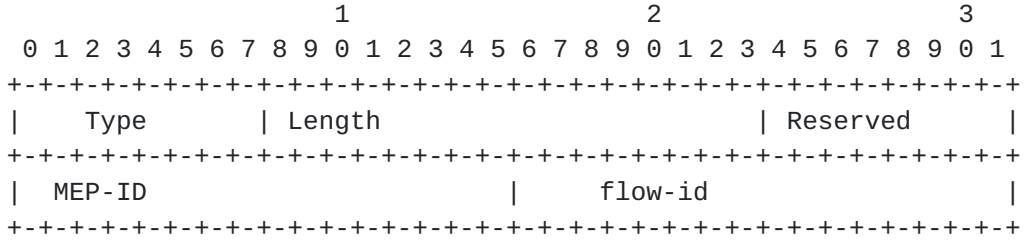

Figure 21 Out of Band IP Address TLV

Type (1 octet) = 72

Length (2 octets) = 5.

Reserved (1 octet) set to 0 on transmission and ignored on reception.

MEP-ID (2 octets) = MEP-ID of the originator [8021Q].

Flow-id (2 octets) = uniquely identifies the flow per MEP. Different MEP may allocate the same flow-id value. The {MEP-ID,flow-id} pair is globally unique.

Inclusion of the MEP-ID in the flow-id TLV allows inclusion of MEP-ID for messages that does not contain MEP-ID in OAM header. Applications may use MEP-ID information for different purposes of troubleshooting.

10. Loopback Message

10.1.1. Loopback OAM Message format

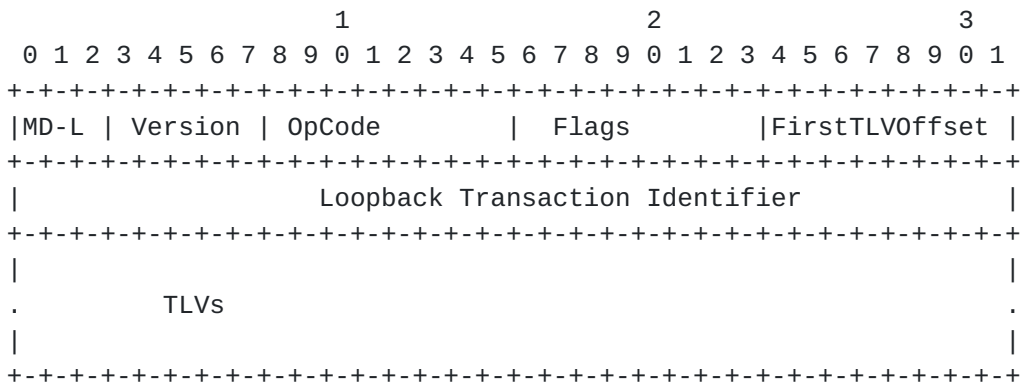

Figure 22 Loopback OAM Message Format

The above figure depicts the format of the Loopback Request and response messages as defined in [8021Q]. The Opcode for Loopback Message is set to 65 and the Opcode for the Reply Message is set to 64. Session Identification Number is a 32-bit integer that allows the requesting RBridge to uniquely identify the corresponding session. Responding RBridges, MUST echo the received "Loopback Transaction Identifier" number without modification.

10.1.2. Theory of Operation

10.1.2.1. Originator RBridge

Originator RBridge Identifies the destination RBridge nickname based on user specification or based on location of the specified destination inner MAC address.

Constructs the flow entropy based on user specified parameters or implementation specific default parameters.

Constructs the TRILL OAM header: Set the opcode to Loopback message type (3). Assign applicable Loopback Transaction Identifier number for the request.

TRILL OAM Version TLV MUST be included and with the flags set to applicable values.

Include following OAM TLVs, where applicable

- Out-of-band Reply address TLV

- o Diagnostic Label TLV

- o Sender ID TLV

Specify the Hop count of the TRILL data frame per user specification or utilize an applicable Hop count value.

Dispatch the OAM frame for transmission.

RBridge may continue to retransmit the request at periodic intervals, until a response is received or the re-transmission count expires. At each transmission Session Identification number MUST be incremented.

10.1.2.2. Intermediate RBridge

Intermediate RBridges forward the frame as a normal data frame and no special handling is required.

10.1.2.3. Destination RBridge

If the Loopback message is addressed to the local RBridge and satisfies the OAM identification criteria specified in section 3.1. then, the RBridge data plane forwards the message to the CPU for further processing.

TRILL OAM application layer further validates the received OAM frame by examining the presence of OAM-Ethertype at the end of the flow entropy and the MD Level. Frames that do not contain OAM-Ethertype at the end of the flow entropy MUST be discarded.

Construction of the TRILL OAM response:

TRILL OAM application encodes the received TRILL header and flow entropy in the Original payload TLV and includes it in the OAM message.

Set the Return Code and Return sub code to applicable values. Update the TRILL OAM opcode to 2 (Loopback Message Reply)

Optionally, if the VLAN/FGL identifier value of the received flow entropy differs from the value specified in the diagnostic Label, set the Label Error Flag on TRILL OAM Application Identifier TLV.

Include the sender ID TLV (1)

If in-band response was requested, dispatch the frame to the TRILL data plane with request-originator RBRidge nickname as the egress

RBridge nickname.

If out-of-band response was requested, dispatch the frame to the IP forwarding process.

11. Path Trace Message

The primary use of the Path Trace Message is for fault isolation. It may also be used for plotting the path taken from a given RBridge to another RBridge.

[8021Q] accomplishes the objectives of the TRILL Path Trace Message using Link Trace Messages. Link Trace Messages utilize a well-known multicast MAC address. This works for [8021Q], because for 802.1 both the unicast and multicast paths are congruent. However, TRILL is multicast and unicast incongruent. Hence, we propose TRILL OAM to utilize a new message format: the Path Trace message.

The Path Trace Message has the same format as Loopback Message. Opcode for Path Trace Reply Message is 65 and Request 64

Operation of Path Trace message is identical to Loopback message except, that it is first transmitted with a TRILL Hop count field value of 1. Sending RBridge expects a Time Expiry Return-Code from the next hop or a successful response. If a Time Expiry Return-code is received as the response, the originator RBridge records the information received from intermediate node that generated the Time Expiry message and resends the message by incrementing the previous Hop count value by 1. This process is continued until, a response is received from the destination RBridge or Path Trace process timeout occur or Hop count reaches a configured maximum value.

11.1.1. Theory of Operation

11.1.1.1. Originator RBridge

Identify the destination RBridge based on user specification or based on location of the specified MAC address.

Construct the flow entropy based on user specified parameters or implementation specific default parameters.

Construct the TRILL OAM header: Set the opcode to Path Trace Request message type (65). Assign applicable Session Identification number for the request. Return-code and sub-code MUST be set to zero.

TRILL OAM Application Identifier TLV MUST be included and set the flags to applicable values.

Include following OAM TLVs, where applicable

- o Out-of-band IP address TLV
- o Diagnostic Label TLV
- o Include the Sender ID TLV

Specify the Hop count of the TRILL data frame as 1 for the first request.

Dispatch the OAM frame to the TRILL data plane for transmission.

An RBridge may continue to retransmit the request at periodic intervals, until a response is received or the re-transmission count expires. At each new re-transmission, the Session Identification number MUST be incremented. Additionally, for responses received from intermediate RBridges, the RBridge nickname and interface information MUST be recorded.

11.1.1.2. Intermediate RBridge

Path Trace Messages transit through Intermediate RBridges transparently, unless Hop-count has expired.

TRILL OAM application layer further validates the received OAM frame by examining the presence of TRILL OAM Flag and OAM-Ethertype at the end of the flow entropy and by examining the MD Level. Frames that do not contain OAM-Ethertype at the end of the flow entropy MUST be discarded.

Construction of the TRILL OAM response:

TRILL OAM application encodes the received TRILL header and flow entropy in the Original payload TLV and include it in the OAM message.

Set the Return Code to (2) "Time Expired" and Return sub code to zero (0). Update the TRILL OAM opcode to 64 (Path Trace Message Reply).

If the VLAN/FGL identifier value of the received flow entropy differs from the value specified in the diagnostic Label, set the Label Error Flag on TRILL OAM Application Identifier TLV.

Include following TLVs

Upstream RBridge nickname TLV (69)

Reply Ingress TLV (5)

Reply Egress TLV (6)

Interface Status TLV (4)

TRILL Next Hop RBridge (Repeat for each ECMP) (70)

Sender ID TLV (1)

If Label error detected, set C flag (Label error detected) in the version.

If in-band response was requested, dispatch the frame to the TRILL data plane with request-originator RBRidge nickname as the egress RBridge nickname.

If out-of-band response was requested, dispatch the frame to the standard IP forwarding process.

11.1.1.3. Destination RBridge

Processing is identical to section 11.1.1.2. With the exception that TRILL OAM Opcode is set to Path Trace Reply (64).

12. Multi-Destination Tree Verification (MTV) Message

Multi-Destination Tree Verification messages allow verifying TRILL distribution tree integrity and pruning. TRILL VLAN/FGL and multicast pruning are described in [RFC6325] [RFCclcorrect] and [RFCfgl]. Multi-destination tree verification and Multicast group verification messages are designed to detect pruning defects. Additionally, these tools can be used for plotting a given multicast tree within the TRILL campus.

Multi-Destination tree verification OAM frames are copied to the CPU of every intermediate RBridge that is part of the distribution tree being verified. The originator of the Multi-destination Tree verification message, specifies the scope of RBridges from which a response is required. Only, the RBridges listed in the scope field respond to the request. Other RBridges silently discard the request. Inclusion of scope parameter is required to prevent receiving a large number of responses. Typical scenario of distribution tree verification or group verification involves verifying multicast connectivity to selected set of end-nodes as opposed to the entire network. Availability of the scope facilitates narrowing down the

focus to only the interested RBridges.

Implementations MAY choose to rate-limit CPU bound multicast traffic. As a result of rate-limiting or due to other congestion conditions, MTV messages may be discarded from time to time by the intermediate RBRidges and the requester may be required to retransmit the request. Implementations SHOULD narrow the embedded scope of retransmission request only to RBRidges that have failed to respond.

12.1. Multi-Destination Tree Verification (MTV) OAM Message Format

Format of MTV OAM Message format is identical to that of Loopback Message format defined in section 10. with the exception that the Loopback Transaction Identifier, in section 10.1.1. , is replaced with the Session Identifier.

12.2. Theory of Operation

12.2.1. Originator RBridge

User is required at minimum to specify either the distriubiton trees that need to be verified, or Multicast MAC address and VLAN/FGL, or VLAN/FGL and Multicast destination IP address. Alternatively, for more specific multicast flow verification, the user MAY specify more information e.g. source MAC address, VLAN/FGL, Destination and Source IP addresses. Implementations, at a minimum, must allow the user to specify a choice of distribution trees, Destination Multicast MAC address and VLAN/FGL that needed to be verified. Although, it is not mandatory, it is highly desired to provide an option to specify the scope. It should be noted that the source MAC address and some other parameters may not be specified if the Backwards Compatibility Method of section 3.2 is used to identify the OAM frames.

Default parameters MUST be used for unspecified parameters. Flow entropy is constructed based on user specified parameters and/or default parameters.

Based on user specified parameters, the originating RBridge identifies the nickname that represent the multicast tree.

Obtain the applicable Hop count value for the selected multicast tree.

Construct TRILL OAM message header and include Session Identification number. Session Identification number facilitate the originator to map the response to the correct request.

TRILL OAM Application Identifier TLV MUST be included.

Op-Code MUST be specified as Multicast Tree Verification Message (70)

Include RBridge scope TLV (67)

Optionally, include following TLV, where applicable

- Out-of-band IP address
- Diagnostic Label
- Sender ID TLV (1)

Specify the Hop count of the TRILL data frame per user specification. Or utilize the applicable Hop count value, if TRILL Hop count is not being specified by the user.

Dispatch the OAM frame to the TRILL data plane to be ingressed for transmission.

RBridge may continue to retransmit the request at a periodic interval, until a response is received or the re-transmission count expires. At each new re-transmission, the Session Identification number MUST be incremented. At each re-transmission, the RBridge may further reduce the scope to the RBridges that it has not received a response from.

12.2.2. Receiving RBridge

Receiving RBridges identify multicast verification frames per the procedure explained in sections 3.2.

CPU of the RBridge validates the frame and analyzes the scope RBridge list. If the RBridge scope TLV is present and the local RBridge nickname is not specified in the scope list, it will silently discard the frame. If the local RBridge is specified in the scope list OR RBridge scope TLV is absent, the receiving RBridge proceeds with further processing as defined in section 12.2.3.

12.2.3. In scope RBridges

Construction of the TRILL OAM response:

TRILL OAM application encodes the received TRILL header and flow entropy in the Original payload TLV and include in the OAM message.

Set the Return Code to (0) and Return sub code to zero (0). Update the TRILL OAM opcode to 67 (Multicast Tree Verification Reply).

Include following TLVs

Upstream RBridge nickname TLV (69)

Reply Ingress TLV (5)

Interface Status TLV (4)

TRILL Next Hop RBridge (Repeat for each downstream RBridge) (70)

Sender ID TLV (1)

Multicast Receiver Availability TLV (71)

If VLAN cross connect error detected, set C flag (Cross connect error detected) in the version.

If in-band response was requested, dispatch the frame to the TRILL data plane with request-originator RBRidge nickname as the egress RBridge nickname.

If out-of-band response was requested, dispatch the frame to the standard IP forwarding process.

13. Application of Continuity Check Message (CCM) in TRILL

Section 8. provides an overview of CCM Messages defined in [8021Q] and how they can be used within the TRILL OAM. In this section, we present the application and Theory of Operations of CCM within the TRILL OAM framework. Readers are referred to [8021Q] for CCM message format and applicable TLV definitions and usages. Only the TRILL specific aspects are explained below.

In TRILL, between any two given MEPs there can be multiple potential paths. Whereas in [8021Q], there is always a single path between any two MEPs, at any given time. [TRILLOAMREQ] requires solutions to have the ability to monitor continuity over one or more paths.

CCM Messages are uni-directional, such that there is no explicit response to a received CCM message. Connectivity status is indicated by setting the applicable flags (e.g. RDI) of the CCM messages transmitted by an MEP.

It is important that the proposed solution accomplishes the requirements specified in [TRILLOAMREQ] within the framework of [8021Q] in a straightforward manner and with minimum changes.

Section 8, above proposed to define multiple flows within the CCM

object, each corresponding to a flow that a given MEP wishes to monitor.

Receiving MEPs do not cross check whether a received CCM belongs to a specific flow from the originating RBridge. Any attempt to track status of individual flows may explode the amount of state information that any given RBridge has to maintain.

Obvious question arises is, how does the originating RBridge knows which flow or flows are at fault?

13.1. CCM Error notification - Method-1

This is accomplished with a combination of RDI flag in the CCM header and SNMP Notifications (Traps).

Each MEP transmits 4 CCM messages per each flow. ([8021Q] detects CCM fault when 3 consecutive CCM messages are lost). Each CCM Message has a unique sequence number.

When an MEP notice a CCM timeout from a remote MEP (MEP-A), it sets the RDI flag on next CCM message it generates. Additionally, it logs and sends SNMP notification that contain the remote MEP Identification, Sequence Number of the last CCM message it received and if available the Sequence Number of the first CCM message it received after the failure. CCM Messages generated by MEP-A has monotonically increasing Sequence Numbers; hence operator can easily identify flows that correspond to specific Sequence Numbers.

Following example illustrate the above.

Assume there are two MEPs, MEP-A and MEP-B.

Assume there are 3 flows between MEP-A and MEP-B.

Lets assume MEP-A allocates sequence numbers as follows

Flow-1 {1,2,3,4,13,14,15,16,.. }

Flow-2 {5,6,7,8,17,18,19,20,.. }

Flow-3 {9,10,12,11,21,22,23,24,.. }

Lets Assume Flow-2 is at fault.

MEP-B, receives CCM from MEP-A with sequence numbers 1,2,3,4, but did not receive 5,6,7,8. CCM timeout is set to 3 CCM intervals in [8021Q]. Hence MEP-B detects the error at 8'th CCM message. At this

time the sequence number of the last good CCM message MEP-B has

received from MEP-A is 4. Hence MEP-B will generate an CCM error SNMP notification with MEP-A and Last good sequence number 4.

When MEP-A switch to flow-3 after transmitting flow-2, MEP-B will start receiving CCM messages, in this example it will be CCM message with Sequence Numbers 9,21 and so on. When receipt of a new CCM message from a specific MEP, after a CCM timeout, TRILL OAM will generate SNMP Notification of CCM resume witn remote MEP-ID and the first valid Sequence number after the CCM timeout. In the foregoing example, it is MEP-A and Sequence Number 9.

We propose to augment remote MEP list under CCM MIB Object to contain "Last Sequence Number" and "CCM Timeout" variables. Last Sequence Number is updated every time a CCM is received from remote MEP. CCM Timeout variable is set when a CCM timeout has occurred and cleared when a CCM is received. Combination of the two new MIB variables and use of monotonically increasing sequence numbers allow TRILL OAM to clearly identify specific flow or flows at fault.

13.2. CCM Error Notification Method-2

This is accomplished with a combination of RDI flag in the CCM header, flow-id TLV and SNMP Notifications (Traps).

Each MEP transmits 4 CCM messages per each flow. ([8021Q] detects CCM fault when 3 consecutive CCM messages are lost). Each CCM Message has a unique sequence number and unique flow-identifier. The flow identifier is included in the OAM message via flow-id TLV.

When an MEP notice a CCM timeout from a remote MEP (MEP-A), it sets the RDI flag on next CCM message it generates. Additionally, it logs and sends SNMP notification that contain the remote MEP Identification, flow-id and the Sequence Number of the last CCM message it received and if available, the flow-id and the Sequence Number of the first CCM message it received after the failure. Each MEP maintain a unique flow-id per each flow, hence operator can easily identify flows that correspond to the specific flow-id.

Following example illustrate the above.

Assume there are two MEP, MEP-A and MEP-B.

Assume there are 3 flows between MEP-A and MEP-B.

Lets assume MEP-A allocates sequence numbers as follows

Flow-1 Sequence={1,2,3,4,13,14,15,16,.. } flow-id=(1)

Flow-2 Sequence={5,6,7,8,17,18,19,20,.. } flow-id=(2)

Flow-3 Sequence={9,10,12,11,21,22,23,24,.. } flow-id=(3)

Lets Assume Flow-2 is at fault.

MEP-B, receives CCM from MEP-A with sequence numbers 1,2,3,4, but did not receive 5,6,7,8. CCM timeout is set to 3 CCM intervals in [8021Q]. Hence MEP-B detects the error at 8'th CCM message. At this time the sequence number of the last good CCM message MEP-B has received from MEP-A is 4 and flow-id of the last good CCM Message is (1). Hence MEP-B will generate a CCM error SNMP notification with MEP-A and Last good flow-id (1) and sequence number 4.

When MEP-A switch to flow-3 after transmitting flow-2, MEP-B will start receiving CCM messages, in this example it will be CCM message with Sequence Numbers 9,10,11,12,21 and so on. When receipt of a new CCM message from a specific MEP, after a CCM timeout, TRILL OAM will generate SNMP Notification of CCM resume with remote MEP-ID and the first valid flow-id and the Sequence number after the CCM timeout. In the foregoing example, it is MEP-A, flow-id (1) and Sequence Number 9.

We propose to augment remote MEP list under CCM MIB Object to contain "Last Sequence Number", flow-id and "CCM Timeout" variables. Last Sequence Number and flow-id are updated every time a CCM is received from a remote MEP. CCM Timeout variable is set when CCM timeout is occurred and cleared when CCM is received.

13.3. Theory of Operation

13.3.1. Originator RBridge

Derive the flow entropy based on flow entropy specified in the CCM Management object.

Construct the TRILL CCM OAM header as specified in [8021Q].

TRILL OAM Version TLV MUST be included as the first TLV and set the flags to applicable values.

Include other TLV specified in [8021Q]

Include following optional TRILL OAM TLVs, where applicable

- Sender ID TLV

Specify the Hop count of the TRILL data frame per user specification or utilize an applicable Hop count value.

Dispatch the OAM frame to the TRILL data plane for transmission.

RBridge transmits a total of 4 requests, each at CCM retransmission interval. At each transmission Session Identification number MUST be incremented by one.

At the 5 retransmission interval, flow entropy of the CCM packet is updated to the next flow entropy specified in the CCM Management Object. If current flow entropy is the last flow entropy specified, move to the first flow entropy specified and continue the process.

13.3.2. Intermediate RBridge

Intermediate RBridges forward the frame as a normal data frame and no special handling is required.

13.3.3. Destination RBridge

If the CCM Message is addressed to the local RBridge or multicast and satisfies OAM identification methods specified in sections 3.2. then the RBridge data plane forwards the message to the CPU for further processing.

TRILL OAM application layer further validates the received OAM frame by examining the presence of OAM-Ethertype at the end of the flow entropy. Frames that do not contain OAM-Ethertype at the end of the flow entropy MUST be discarded.

Validate the MD-LEVEL and pass the packet to the Opcode de-multiplexer. Opcode de-multiplexer delivers CCM packets to the CCM process.

CCM Process performs processing specified in [8021Q].

Additionally CCM process updates the CCM Management Object with the sequence number of the received CCM packet. Note: Last received CCM sequence number and CCM timeout is tracked per each remote MEP.

If CCM timeout is true for the sending remote MEP, then clear the CCM timeout in the CCM Management object and generate SNMP notification as specified above.

14. Multiple Fragment Reply

Response Message as described in 4.4.2.1 allows Multiple Fragment Reply with use of Final Flag. In case of Multiple Fragment Reply, due to response exceeding MTU size, all messages MUST follow the procedure defined in this section.

All Reply Messages MUST be encoded as described in this document.

Same session Identification Number MUST be included in all related fragments of the same message.

TRILL OAM Application Identifier TLV MUST BE included with the appropriate Final Flag field. Final Flag, MUST, only be set on the final fragment of the reply.

15. Security Considerations

For general TRILL related security considerations, please refer to [RFC6325]. Specific security considerations related methods presented in this document are currently under investigation.

16. Allocation Considerations

16.1. IEEE Allocation Considerations

The IEEE 802.1 Working Group is requested to allocate a separate opcode and TLV space within 802.1QCFM messages for TRILL purpose.

16.2. IANA Considerations

- IANA is requested to allocate a multicast MAC address from the block assigned to TRILL

- Set up sub-registry within the TRILL Parameters registry for block of TRILL OAM OpCodes -

- Set up sub-registry within the TRILL Parameters registry for TRILL OAM TLV Types -

- Set up sub-registry within the TRILL Parameters registry for TRILL OAM return code and return sub codes -

- Request a unicast MAC addressed, reserved for identification of OAM packets discussed in backward compatibility method (section 3.3.)

17. References

17.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC6325] Perlman, R., et.al., "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.

[RFCfgl] D. Eastlake, M. Zhang, P. Agarwal, R. Perlman, D. Dutt, "TRILL: Fine-Grained Labeling", draft-ietf-trill-fine-labeling, work in progress.

17.2. Informative References

[RFC6291] Andersson, L., et.al., "Guidelines for the use of the "OAM" Acronym in the IETF" RFC 6291, June 2011.

[TRILLOAMMIB] "TRILL OAM MIB", To be published.

[RFC4379] Kompella, K. et.al, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[TRILLOAMREQ] Senevirathne, T., et.al., "Requirements for Operations, Administration and Maintenance (OAM) in TRILL", draft-ietf-trill-oam-req, Work in Progress, November, 2012.

[TRILLOAMFM] Salam, S., et.al., "TRILL OAM Framework", draft-ietf-trill-oam-framework, Work in Progress, November, 2012.

[RFCclcorrect] Eastlake, Donald, et.al. "TRILL: Clarifications, Corrections, and Updates, draft-ietf-trill-clear-correct, July 2012.

[8021Q] IEEE, "Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks", IEEE Std 802.1Q-2011, August, 2011.

18. Acknowledgments

Work in this document was largely inspired by the directions provided by Stewart Bryant in finding a common OAM solution between SDO.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Tissa Senevirathne
CISCO Systems
375 East Tasman Drive.
San Jose, CA 95134
USA.

Phone: +1 408-853-2291
Email: tsenevir@cisco.com

Samer Salam
CISCO Systems
595 Burrard St. Suite 2123
Vancouver, BC V7X 1J1, Canada

Email: ssalam@cisco.com

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

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

Donald Eastlake
Huawei Technologies
155 Beaver Street
Milford, MA 01757

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

Sam Aldrin
Huawei Technologies
2330 Central Express Way
Santa Clara, CA 95951
USA

Email: aldrin.ietf@gmail.com

Yizhou Li
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Phone: +86-25-56625375
Email: liyizhou@huawei.com