TRILL Working Group
INTERNET-DRAFT
Intended Status: Informational

Samer Salam
Tissa Senevirathne
Cisco

Sam Aldrin
Donald Eastlake
Huawei

Expires: March 23, 2014

September 19, 2013

TRILL OAM Framework
draft-ietf-trill-oam-framework-03

Abstract

This document specifies a reference framework for Operations, Administration and Maintenance (OAM) in TRILL networks. The focus of the document is on the fault and performance management aspects of TRILL OAM.

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

Table of Contents

1. Introduction

This document specifies a reference framework for Operations, Administration and Maintenance (OAM, [RFC6291]) in TRILL (Transparent Interconnection of Lots of Links) networks.

TRILL [RFC6325] specifies a protocol for shortest-path frame routing in multi-hop networks with arbitrary topologies and link technologies, using the IS-IS routing protocol. TRILL capable devices are referred to as TRILL Switches or RBridges (Routing Bridges). RBridges provide an optimized and transparent Layer 2 delivery service for Ethernet unicast and multicast traffic. Some characteristics of a TRILL network that are different from IEEE 802.1 bridging are the following:

- TRILL networks support arbitrary link technology between TRILL switches. Hence, a TRILL switch port may not have a 48-bit MAC Address [802] but might, for example, have an IP address as an identifier [TRILL-IP] or no unique identifier (PPP [RFC6361]).

- TRILL networks do not enforce congruency of unicast and multicast paths between a given pair of RBridges.

- TRILL networks do not impose symmetry of the forward and reverse paths between a given pair of RBridges.

- TRILL switches terminate spanning tree protocols instead of propagating them.

In this document, we refer to the term OAM as defined in [RFC6291]. The Operations aspect involves finding problems that prevent proper functioning of the network. It also includes monitoring of the network to identify potential problems before they occur. Administration involves keeping track of network resources. Maintenance activities are focused on facilitating repairs and upgrades as well as corrective and preventive measures. [ISO/IEC 7498-4] defines 5 functional areas in the OSI model for network management, commonly referred to as FCAPS:

-Fault Management
-Configuration Management
-Accounting Management
-Performance Management
-Security Management

The focus of this document is on the first and fourth functional aspects, Fault Management and Performance Management, in TRILL networks. These primarily map to the "Operations" and "Maintenance"

part of OAM.

This draft provides a generic framework for a comprehensive solution that meets the requirements outlined in [RFC6905]. However, specific mechanisms to address these requirements are considered to be outside the scope of this document. Furthermore, another document will specify the optional reporting of errors in TRILL user traffic, such as the use of a reserved or unknown egress nickname, etc.

1.1 Terminology

The following acronyms are used in this document:

BFD - Bidirectional Forwarding Detection [RFC5880]
CFM - Connectivity Fault Management [802.1Q]
ECMP - Equal Cost Multi-Pathing
FGL - Fine Grained Label(ing) [TRILL-FGL]
IEEE - Institute for Electrical and Electronic Engineers
IP - Internet Protocol, includes both IPv4 and IPv6
LAN - Local Area Network
MAC - Media Access Control [802]
MA - Maintenance Association
ME - Maintenance Entity
MEP - Maintenance End Point
MIP - Maintenance Intermediate Point
MP - Maintenance Point (MEP or MIP)
OAM - Operations, Administration, and Maintenance [RFC6291]
PPP - Point-to-Point Protocol [RFC1661]
RBridge - Routing Bridge, a device implementing TRILL [RFC6325]
RDI - Reverse Defect Indication
TRILL - Transparent Interconnection of Lots of Links [RFC6325]
TRILL Switch - an alternate name for an RBridge
VLAN - Virtual LAN [802.1Q]

1.2 Relationship to Other OAM Work

OAM is a technology area where a wealth of prior art exists. This document leverages concepts and draws upon elements defined and/or used in the following documents:

[RFC6905] defines the requirements for TRILL OAM that serve as the basis for this framework. It also defines terminology that is used extensively in this document.

[802.1Q] specifies the Connectivity Fault Management (CFM) protocol, which defines the concepts of Maintenance Domains, Maintenance End Points, and Maintenance Intermediate Points.

[Y.1731] extends Connectivity Fault Management in the following areas: it defines fault notification and alarm suppression functions for Ethernet. It also specifies mechanisms for Ethernet performance management, including loss, delay, jitter, and throughput measurement.

[TRILL-BFD] defines a TRILL encapsulation for BFD that enables the use of the latter for network fast convergence.

2. TRILL OAM Model

2.1 OAM Layering

In the TRILL architecture, the TRILL layer is independent of the underlying Link Layer technology. Therefore, it is possible to run TRILL over any transport layer capable of carrying TRILL packets such as Ethernet [RFC6325], PPP [RFC6361], or IP [TRILL-IP]. Furthermore, TRILL provides a virtual Ethernet connectivity service that is transparent to higher layer entities (e.g. Layer 3 and above). This strict layering is observed by TRILL OAM.

Of particular interest is the layering of TRILL OAM with respect to:

- BFD, which is typically used for fast convergence

- Ethernet CFM [802.1Q] on paths from an external device, over a TRILL campus, to another external device, especially since TRILL switches are likely to be deployed where existing 802.1 bridges can be such external devices.

- Link OAM, on links interior to a TRILL campus, which is link technology specific.

Consider the example network depicted in Figure 1 below, where a TRILL network is interconnected via Ethernet links:

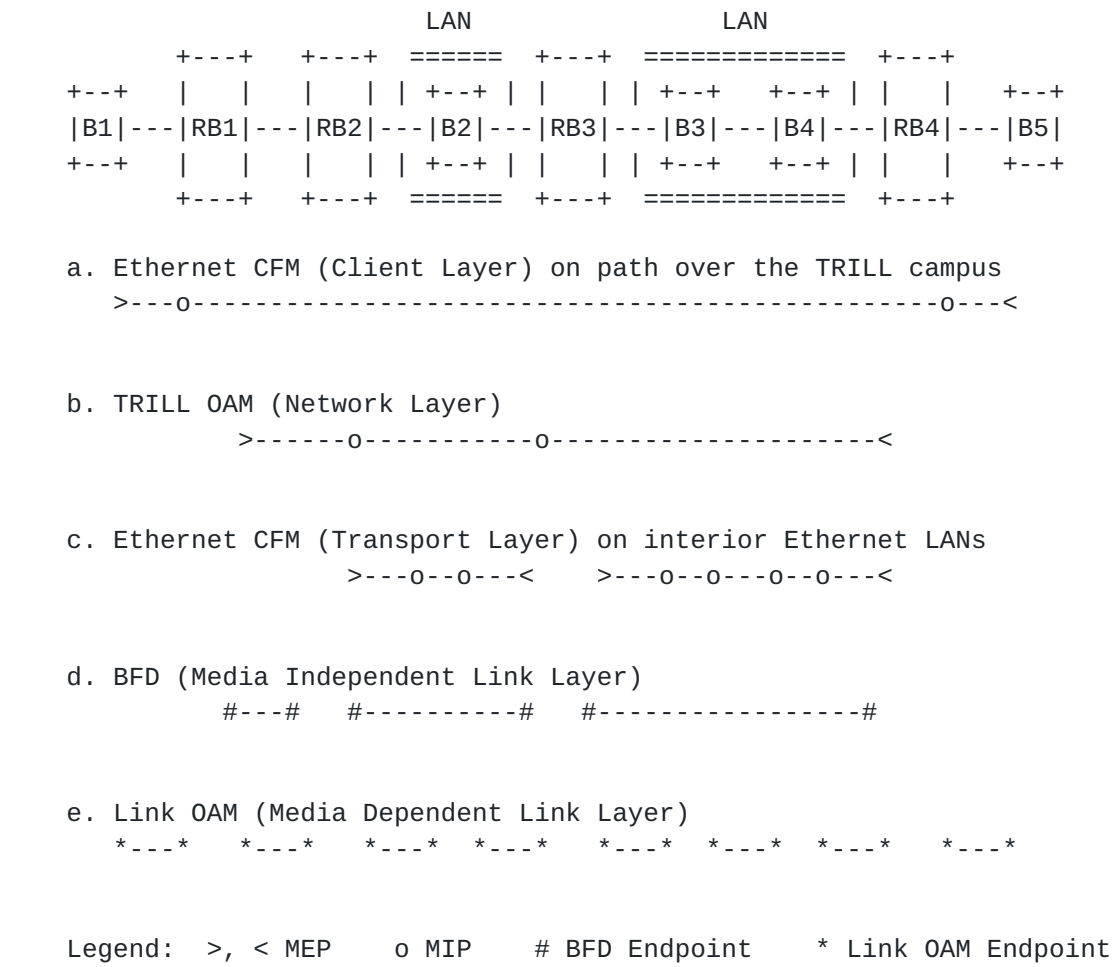

Figure 1: OAM Layering in TRILL

Where Bn and RBn (n= 1,2,3, ...) denote IEEE 802.1Q bridges and TRILL RBridges, respectively.

2.1.1 Relationship to CFM

In the context of a TRILL network, CFM can be used as either a client layer OAM or a transport layer OAM mechanism.

When acting as a client layer OAM (see Figure 1a), CFM provides fault management capabilities for the user, on an end-to-end basis over the TRILL network. Edge ports of the TRILL network may be visible to CFM operations through the optional presence of a CFM Maintenance Intermediate Point (MIP) in the TRILL switches edge Ethernet ports.

When acting as a transport layer OAM (see Figure 1c), CFM provides fault management functions for the IEEE 802.1Q bridged LANs that may interconnect RBridges. Such bridged LANs can be used as TRILL level

links between RBridges. RBridges directly connected to the intervening 802.1Q bridges may host CFM Down Maintenance End Points (MEPs).

2.1.2 Relationship to BFD

One-hop BFD (see Figure 1d) runs between adjacent RBridges and provides fast link as well as node failure detection capability [TRILL-BFD]. Note that TRILL BFD also provides some testing of the TRILL protocol stack and thus sits a layer above Link OAM, which is media specific. BFD provides fast convergence characteristics to TRILL networks. The requirements for BFD are different from those of the TRILL OAM mechanisms that are the prime focus of this document. Furthermore, BFD does not use the frame format described in section 3.1.

TRILL BFD differs from TRILL OAM in two significant ways:

1. A TRILL BFD transmitter is always bound to a specific TRILL output port.

2. TRILL BFD messages can be transmitted by the originator out a port to a neighbor RBridge when the adjacency is in the Detect or Two-Way states as well as when the adjacency is in the Report (Up) state [RFC6327].

In contrast, TRILL OAM messages are typically transmitted by appearing to have been received on a TRILL input port (refer to Section 2.2 for details). In that case, the output ports on which TRILL OAM message are sent are determined by the TRILL routing function. The TRILL routing function will only send on links that are in the Report state and have been incorporated into the local view of the campus topology.

2.1.3 Relationship to Link OAM

Link OAM (see Figure 1e) depends on the nature of the technology used in the links interconnecting RBridges. For example, for Ethernet links, [802.3] Clause 57 OAM may be used.

2.2 TRILL OAM in the RBridge Port Model

TRILL OAM processing can be represented as a layer situated between the port's TRILL encapsulation/de-capsulation function and the TRILL Forwarding Engine function, on any RBridge port. TRILL OAM requires services of the RBridge forwarding engine and utilizes information from the IS-IS control plane. Figure 2 below depicts TRILL OAM processing in the context of the RBridge port model defined in

[RFC6325]. In this figure, double lines represent flow of both frames and information.

This figure shows a conceptual model. It is to be understood that implementations need not mirror this exact model as long as the intended OAM requirements and functionality are preserved.

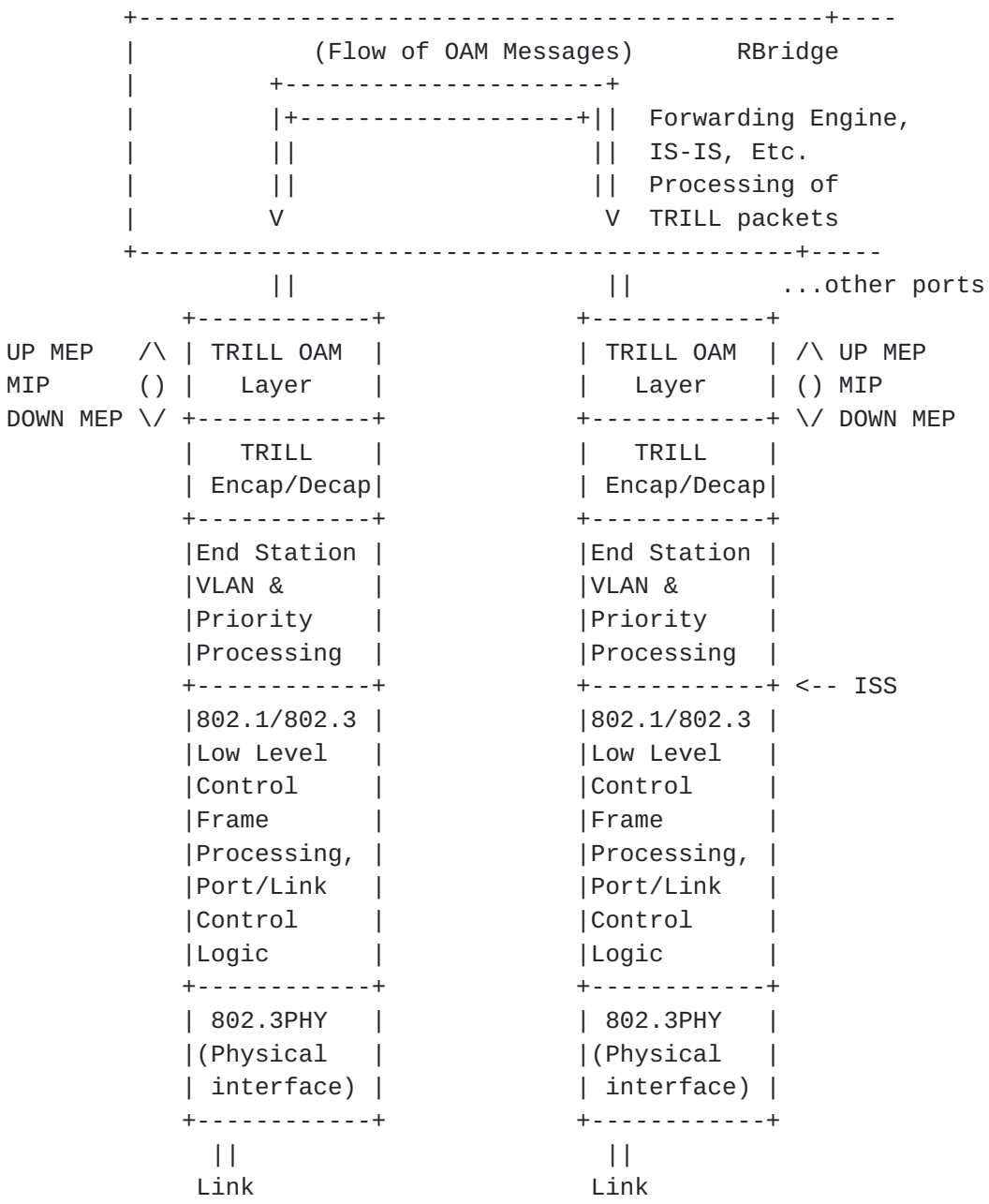

Figure 2: TRILL OAM in RBridge Port Model

Note that the terms "MEP" and "MIP" in the above figure are explained in detail in section 2.6 below.

2.3 Network, Service and Flow OAM

OAM functions in a TRILL network can be conducted at different granularity. This gives rise to 'Network', 'Service' and 'Flow' OAM,

listed in order of finer granularity.

Network OAM mechanisms provide fault and performance management functions in the context of a 'test' VLAN or fine-grained label [TRILL-FGL]. The test VLAN can be thought of as a management or diagnostics VLAN that extends to all RBridges in a TRILL network. In order to account for multipathing, Network OAM functions also make use of test flows (both unicast and multicast) to provide coverage of the various paths in the network.

Service OAM mechanisms provide fault and performance management functions in the context of the actual VLAN or fine-grained label set for which end station service is enabled. Test flows are used here, as well, to provide coverage in the case of multipathing.

Flow OAM mechanisms provide the most fine grained fault and performance management capabilities, where OAM functions are performed in the context of end station flows within VLANs or fine-grained labels. While Flow OAM provides the most granular control, it clearly poses scalability challenges if attempted on large numbers of flows.

2.4 Maintenance Domains

The concept of Maintenance Domains, or OAM Domains, is well known in the industry. IEEE [802.1Q] defines the notion of a Maintenance Domain as a collection of devices (e.g. network elements) that are grouped for administrative and/or management purposes. Maintenance domains usually delineate trust relationships, varying addressing schemes, network infrastructure capabilities, etc.

When mapped to TRILL, a Maintenance Domain is defined as a collection of RBridges in a network for which connectivity faults and performance degradation are to be managed by a single operator. All RBridges in a given Maintenance Domain are, by definition, managed by a single entity (e.g. an enterprise or a data center operator, etc.). [RFC6325] defines the operation of TRILL in a single IS-IS area, with the assumption that a single operator manages the network. In this context, a single (default) Maintenance Domain is sufficient for TRILL OAM.

However, when considering scenarios where different TRILL networks need to be interconnected, for example as discussed in [TRILL-ML], then the introduction of multiple Maintenance Domains and Maintenance Domain hierarchies becomes useful to map and enforce administrative boundaries. When considering multi-domain scenarios, the following rules must be followed: TRILL OAM domains must not partially intersect, but must either be disjoint or nest to form a hierarchy

(i.e. a higher Maintenance Domain may completely enclose a lower Domain). A Maintenance Domain is typically identified by a Domain Name and a Maintenance Level (a numeric identifier). If two domains are nested, the encompassing domain must be assigned a higher Maintenance Level number than the enclosed domain. For this reason, the encompassing domain is commonly referred to as the 'higher' domain, and the enclosed domain is referred to as the 'lower' domain. OAM functions in the lower domain are completely transparent to the higher domain. Furthermore, OAM functions in the higher domain only have visibility to the boundary of the lower domain (for example, an attempt to trace the path in the higher domain will depict the entire lower domain as a single-hop between the RBridges that constitute the boundary of that lower domain). By the same token, OAM functions in the higher domain are transparent to RBridges that are internal to the lower domain. The hierarchical nesting of domains is established through operator configuration of the RBridges.

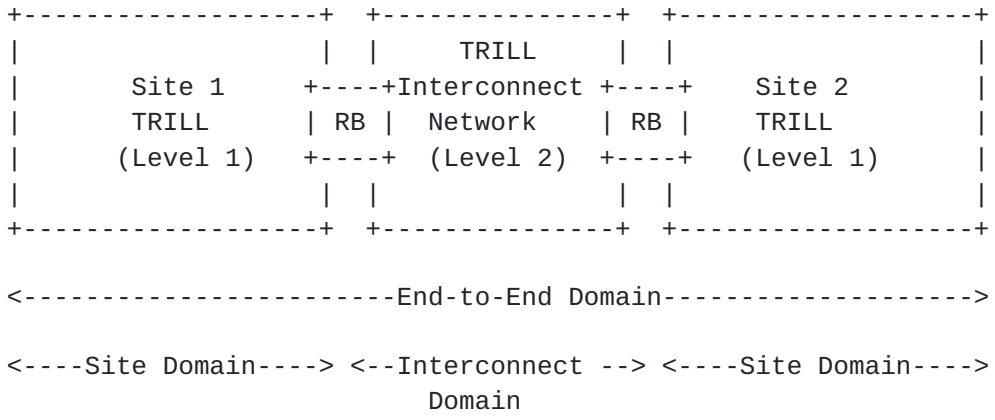

Figure 3: TRILL OAM Maintenance Domains

2.5 Maintenance Entity and Maintenance Entity Group

TRILL OAM functions are performed in the context of logical endpoint pairs referred to as Maintenance Entities (ME). A Maintenance Entity defines a relationship between two points in a TRILL network where OAM functions (e.g. monitoring operations) are applied. The two points that define a Maintenance Entity are known as Maintenance End Points (MEPs) - see section 2.6 below. The set of Maintenance End Points that belong to the same Maintenance Domain are referred to as a Maintenance Association (MA). On the network path in between MEPs, there can be zero or more intermediate points, called Maintenance Intermediate Points (MIPs). MEPs can be part of more than one ME in a given MA.

2.6 MEPs and MIPs

OAM capabilities on RBridges can be defined in terms of logical groupings of functions that can be categorized into two functional objects: Maintenance End Points (MEPs) and Maintenance Intermediate Points (MIPs). The two are collectively referred to as Maintenance Points (MPs).

MEPs are the active components of TRILL OAM: MEPs source TRILL OAM messages periodically or on-demand based on operator configuration actions. Furthermore, MEPs ensure that TRILL OAM messages do not leak outside a given Maintenance Domain, e.g. out of the TRILL network and into end stations. MIPs, on the other hand, are internal to a Maintenance Domain. They are the more passive components of TRILL OAM, primarily responsible for forwarding TRILL OAM messages and selectively responding to a subset of these messages.

The following figure shows the MEP and MIP placement for the Maintenance Domains depicted in Figure 3 above.

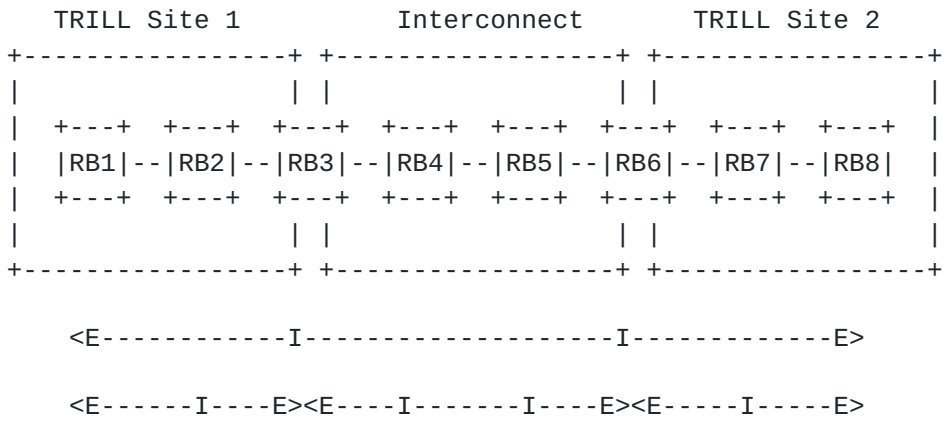

Figure 4: MEPs and MIPs

A single RBridge may host multiple MEPs of different technologies, e.g. TRILL OAM MEP(s) and [802.1Q] MEP(s). This does not mean that the protocol operation is necessarily consolidated into a single functional entity on those ports. The protocol functions for each MEP remain independent and reside in different shims in the RBridge Port model of Figure 2: the TRILL OAM MEP resides in the "TRILL OAM Processing" block whereas a CFM MEP resides in the "End Station VLAN & Priority Processing" block.

In the model of Section 2.2, a single MEP and/or MIP per MA can be

instantiated per RBridge port. A MEP is further qualified with an administratively set direction (UP or DOWN), as follows:

- An UP MEP sends and receives OAM messages through the RBridge Forwarding Engine. This means that an UP MEP effectively communicates with MEPs on other RBridges through TRILL interfaces other than the one that the MEP is configured on.

- A DOWN MEP sends and receives OAM messages through the link connected to the interface on which the MEP is configured.

In order to support TRILL OAM functions on sections, as described in [RFC6905], while maintaining the simplicity of a single TRILL OAM Maintenance Domain, the TRILL OAM Layer may be implemented on a virtual port with no physical layer (Null PHY). In this case, the Down MEP function is not supported, since the virtual port does not attach to a link; as such, a Down MEP on a virtual port would not be capable of sending or receiving OAM messages.

A TRILL OAM solution that conforms to this framework:

- must support the MIP function on TRILL ports (to support fault isolation)
- must support the UP MEP function on a TRILL virtual port (to support OAM functions on Sections, as defined in [RFC6905])
- may support the UP MEP function on TRILL ports
- may support the DOWN MEP function on TRILL ports

2.7 Maintenance Point Addressing

TRILL OAM functions must provide the capability to address a specific Maintenance Point or a set of one or more Maintenance Points in a MA. To that end, RBridges need to recognize two sets of addresses:

- Individual MP addresses

- Group MP Addresses

TRILL OAM will support the Shared MP address model, where all MPs on an RBridge share the same Individual MP address. In other words, TRILL OAM messages can be addressed to a specific RBridge but not to a specific port on an RBridge.

One cannot discern, from observing the external behavior of an RBridge, whether TRILL OAM messages are actually delivered to a certain MP or another entity within the RBridge. The Shared MP address model takes advantage of this fact by allowing MPs in different RBridge ports to share the same Individual MP address. The

MPs may still be implemented as residing on different RBridge ports and for the most part, they have distinct identities.

The Group MP addresses enable the OAM mechanism to reach all the MPs in a given MA. Certain OAM functions, e.g. pruned tree verification, require addressing a subset of the MPs in a MA. Group MP addresses are not defined for such subsets. Rather, the OAM function in question must use the Group MP addresses combined with an indication of the scope of the MP subset encoded in the OAM Message Channel. This prevents an unwieldy set of responses to Group MP addresses.

3. OAM Frame Format

3.1 Motivation

In order for TRILL OAM messages to accurately test the data-path, these messages must be transparent to transit RBridges. That is, a TRILL OAM message must be indistinguishable from a TRILL data packet through normal transit RBridge processing. Only the target RBridge, which needs to process the message, should identify and trap the packet as a control message through normal processing. Additionally methods must be provided to prevent OAM packets from being transmitted out as native frames.

The TRILL OAM packet format proposed below provides the necessary flexibility to exercise the data path as closely as possible to actual data packets.

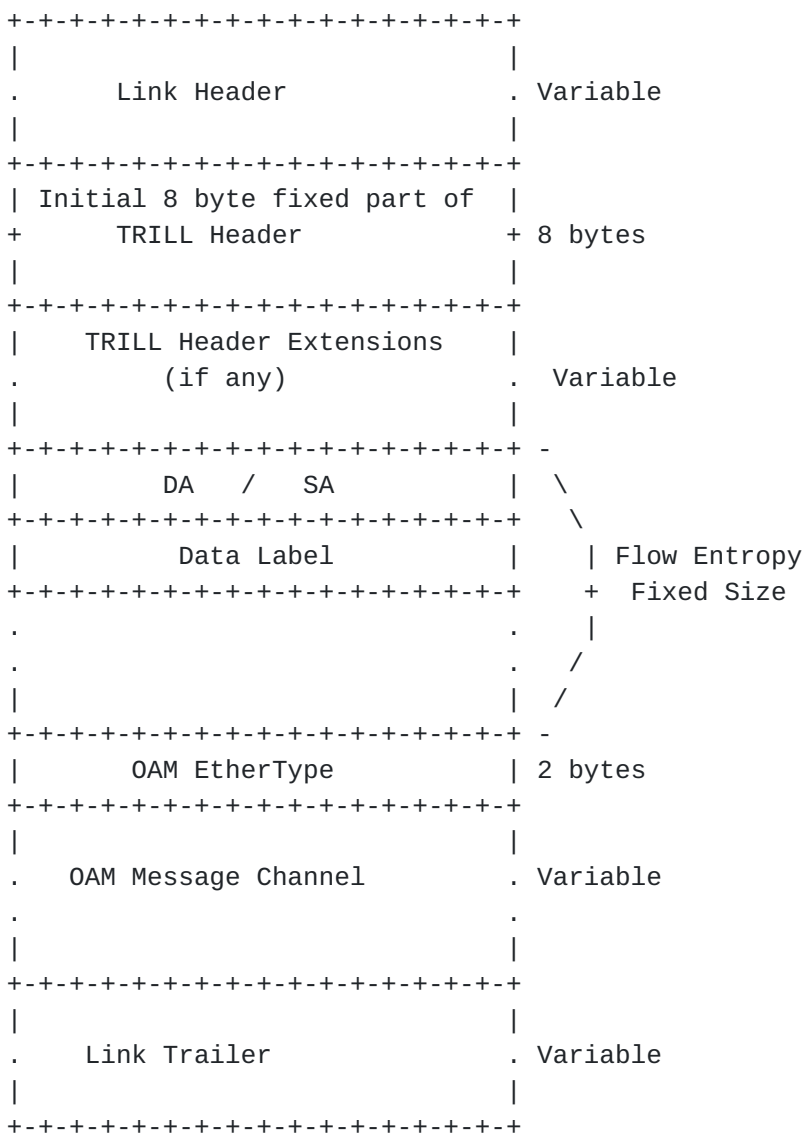

Figure 5: OAM Frame Format

The TRILL Header and the Link Header and Trailer need to be as similar as practical to the Link Header/Trailer and TRILL Header of the normal TRILL data packet corresponding to the traffic that OAM is testing.

The OAM EtherType demarcates the boundary between the Flow Entropy and the OAM Message Channel. The OAM EtherType is expected at a deterministic offset from the TRILL Header, thereby allowing applications to clearly identify the beginning of the OAM Message Channel. Additionally, it facilitates the use of the same OAM frame structure by different Ethernet technologies.

The Link Trailer is usually a checksum, such as the Ethernet Frame Check Sequence, which is examined at a low level very early in the frame input process and automatically generated as part of the low level frame output process. If the checksum fails, the frame is normally discarded with no higher level processing.

3.2 Determination of Flow Entropy

The Flow Entropy is a fixed length field that is populated with either real packet data or synthetic data that mimics the intended flow. It always start with a destination and source MAC address area followed by a Data Label area (either a VLAN or fine grain label).

For a Layer 2 flow (i.e. non-IP) the Flow Entropy must specify the desired Ethernet header, including the MAC destination and source addresses as well as a VLAN tag or fine grain label.

For a Layer 3 flow, the Flow Entropy must specify the desired Ethernet header, the IP header and UDP or TCP header fields, although the Ethernet layer header fields are also still present.

Not all fields in the Flow Entropy field need to be identical to the data flow that the OAM message is mimicking. The only requirement is for the selected flow entropy to follow the same path as the data flow that it is mimicking. In other words, the selected flow entropy must result in the same ECMP selection or multicast pruning behavior or other applicable forwarding paradigm.

When performing diagnostics on user flows, the OAM mechanisms must allow the network operator to configure the flow entropy parameters (e.g. Layer 2 and/or 3) on the RBridge from which the diagnostic operations are to be triggered.

When running OAM functions over Test Flows, the TRILL OAM may provide a mechanism for discovering the flow entropy parameters by querying the RBridges dynamically, or allow the network operator to configure the flow entropy parameters.

3.2.1 Address Learning and Flow Entropy

Edge TRILL switches, like traditional 802.1 bridges, are required to learn MAC address associations. Learning is accomplished either by snooping data packets or through other methods. The flow entropy field of TRILL OAM messages mimics real packets and may impact the address learning process of the TRILL data plane. TRILL OAM is required to provide methods to prevent any learning of addresses from the flow entropy field of OAM messages that would interfere with normal TRILL operation. This can be done, for example, by

suppressing/preventing MAC address learning from OAM messages.

3.3 OAM Message Channel

The OAM Message Channel provides methods to communicate OAM specific details between RBridges. [802.1Q] CFM and [RFC4379] have implemented OAM message channels. It is desirable to select an appropriate technology and re-use it, instead of redesigning yet another OAM channel. TRILL is a transport layer that carries Ethernet frames, so the TRILL OAM model specified earlier is based on the [802.1Q] CFM model. The use of [802.1Q] CFM encoding format for the OAM Message channel is one possible choice. [TRILL-OAM] presents a proposal on the use of [802.1Q] CFM payload as the OAM message channel.

3.4 Identification of OAM Messages

RBridges must be able to identify OAM messages that are destined to them, either individually or as a group, so as to properly process those messages.

TRILL, as defined in [RFC6325], does not specify a method to identify OAM messages. The most reliable method to identify these messages, without imposing restrictions on the Flow Entropy field, involves modifying the definition of the TRILL header to include an "Alert" flag. This flag signals that the contents of the TRILL packet is a control message as opposed to user data. The use of such a flag would not be limited to TRILL OAM, and may be leveraged by any other TRILL control protocol that require in-band behavior. The TRILL header currently has two reserved bits that are unused. One of those bits may be used as the Alert flag. In order to guarantee accurate in-band forwarding behavior, RBridges must not use the Alert flag in ECMP hashing decisions. Furthermore, to ensure that this flag remains protocol agnostic, TRILL OAM mechanisms must not rely solely on the Alert flag to identify OAM messages. Rather, these solutions must identify OAM messages based on the combination of the Alert flag and the OAM EtherType.

Since the above mechanism requires modification of the TRILL header, it is not backward compatible. TRILL OAM solutions should provide alternate methods to identify OAM messages that work on existing RBridge implementations, thereby providing backwards compatibility.

4. Fault Management

Section 4.1 below discusses proactive fault management and Section 4.2 discusses on-demand fault management.

4.1 Proactive Fault Management Functions

Proactive fault management functions are configured by the network operator to run periodically without a time bound, or are configured to trigger certain actions upon the occurrence of specific events.

4.1.1 Fault Detection (Continuity Check)

Proactive fault detection is performed by periodically monitoring the reachability between service endpoints, i.e. MEPs in a given MA, through the exchange of Continuity Check messages. The reachability between any two arbitrary MEP may be monitored for a specified path, all paths or any representative path. The fact that TRILL networks do not enforce congruency between unicast and multicast paths means that the proactive fault detection mechanism must provide procedures to monitor the unicast paths independently of the multicast paths. Furthermore, where the network has ECMP, the proactive fault detection mechanism must be capable of exercising the equal-cost paths individually.

The set of MEPs exchanging Continuity Check messages in a given domain and for a specific monitored entity (flow, network or service) must use the same transmission period. As long as the fault detection mechanism involves MEPs transmitting periodic heartbeat messages independently, then this OAM procedure is not affected by the lack of forward/reverse path symmetry in TRILL.

The proactive fault detection function must detect the following types of defects:

- Loss of continuity to one or more remote MEPs
- Unexpected connectivity between isolated VLANs or fine-grained labels (mismerge)
- Unexpected connectivity to one or more remote MEPs
- Mismatch of the Continuity Check transmission period between MEPs

4.1.2 Defect Indication

TRILL OAM must support event-driven defect indication upon the detection of a connectivity defect. Defect indications can be categorized into two types:

4.1.2.1 Forward Defect Indication

This is used to signal a failure that is detected by a lower layer OAM mechanism. Forward Defect indication is transmitted away from the direction of the failure. For example, consider a simple network comprising of four RBridges connected in series: RB1, RB2, RB3 and RB4. Both RB1 and RB4 are hosting TRILL OAM MEPs, whereas RB2 and RB3 have MIPs. If the link between RB2 and RB3 fails, then RB2 can send a

forward defect indication towards RB1 while RB3 sends a forward defect indication towards RB4.

Forward defect indication may be used for alarm suppression and/or for purpose of inter-working with other layer OAM protocols. Alarm suppression is useful when a transport/network level fault translates to multiple service or flow level faults. In such a scenario, it is enough to alert a network management station (NMS) of the single transport/network level fault in lieu of flooding that NMS with a multitude of Service or Flow granularity alarms.

4.1.2.2 Reverse Defect Indication (RDI)

RDI is used to signal that the advertising MEP has detected a loss of continuity defect. RDI is transmitted in the direction of the failure. For example, consider the same series network of the previous section (4.1.2.1). If RB1 detects that is has lost connectivity to RB4 because it is no longer receiving Continuity Check messages from the MEP on RB4, then RB1 can transmit an RDI towards RB4 to inform the latter of the failure. If the failure is unidirectional (i.e. it is affecting the direction from RB4 to RB1), then the RDI enables RB4 to become aware of the unidirectional connectivity anomaly.

In the presence of equal-cost paths between MEPs, RDI must be able to identify on which equal-cost path the failure was detected.

RDI allows single-sided management, where the network operator can examine the state of a single MEP and deduce the overall health of a monitored entity (network, flow or service).

4.2 On-Demand Fault Management Functions

On-demand fault management functions are initiated manually by the network operator either as a one-time occurrence or as an action/test that continues for a time bound period. These functions enable the operator to run diagnostics to investigate a defect condition.

4.2.1 Connectivity Verification

As specified in [RFC6905], TRILL OAM must support on-demand connectivity verification for unicast and multicast. The connectivity verification mechanism must provide a means for specifying and carrying in the messages:

- variable length payload/padding to test MTU related connectivity problems.

- test message formats as defined in [RFC2544].

4.2.1.1 Unicast

Unicast connectivity verification operation must be initiated from a MEP and may target either a MIP or another MEP. For unicast, connectivity verification can be performed at either Network or Flow granularity.

Connectivity verification at the Network granularity tests connectivity between a MEP on a source RBridge and a MIP or MEP on a target RBridge over a test VLAN or fine grain label and for a test flow. The operator must supply the source and target RBridges for the operation, and the test VLAN/flow information uses pre-set values or defaults.

Connectivity verification at the Flow granularity tests connectivity between a MEP on a source RBridge and a MIP or MEP on a target RBridge over an operator specified VLAN or fine grain label with operator specified flow parameters.

The above functions must be supported on sections, as defined in [RFC6905]. When connectivity verification is triggered over a section, and the initiating MEP does not coincide with the edge (ingress) RBridge, the MEP must use the edge RBridge nickname instead of the local RBridge nickname on the associated connectivity verification messages. The operator must supply the edge RBridge nickname as part of the operation parameters.

4.2.1.2 Multicast

For multicast, the connectivity verification function tests all branches and leaf nodes of a multi-destination distribution tree for reachability. This function should include mechanisms to prevent reply storms from overwhelming the initiating RBridge. This may be done, for example, by staggering the replies through the introduction of a random delay timer, with a preset upper bound, on the responding RBridge ([802.1Q] CFM uses similar mechanisms for Linktrace Reply messages to mitigate the load on the originating MEP). The upper bound on the timer value should be selected by the OAM solution to be long enough to accommodate large distribution trees, while allowing the connectivity verification operation to conclude within a reasonable time. To further prevent reply storms, connectivity verification operation is initiated from a MEP and must target MEPs only. MIPs are transparent to multicast connectivity verification.

Per [RFC6905], multicast connectivity verification must provide the following granularity of operation:

A. Un-pruned Tree

- Connectivity verification for un-pruned multi-destination distribution tree. The operator in this case supplies the tree identifier (root nickname) and campus wide diagnostic VLAN or fine grain label.

B. Pruned Tree

- Connectivity verification for a VLAN or fine-grain label in a given multi-destination distribution tree. The operator in this case supplies the tree identifier and VLAN or fine grain label.

- Connectivity verification for an IP multicast group in a given multi-destination distribution tree. The operator in this case supplies: the tree identifier, VLAN or fine grain label and IP (S,G) or (*,G).

4.2.2 Fault Isolation

TRILL OAM must support an on-demand connectivity fault localization function. This is the capability to trace the path of a Flow on a hop-by-hop (i.e. RBridge by RBridge) basis to isolate failures. This involves the capability to narrow down the locality of a fault to a particular port, link or node. The characteristic of forward/reverse path asymmetry, in TRILL, renders fault isolation into a direction-sensitive operation. That is, given two RBridges A and B, localization of connectivity faults between them requires running fault isolation procedures from RBridge A to RBridge B as well as from RBridge B to RBridge A. Generally speaking, single-sided fault isolation is not possible in TRILL OAM.

Furthermore, TRILL OAM should support fault isolation over distribution trees for both un-pruned as well as pruned trees. The former allows the tracing of all active branches of a tree, whereas the latter allows tracing of the active subset of branches associated with a given Flow.

5. Performance Monitoring

Performance Monitoring functions are optional in TRILL OAM, per [RFC6905]. These functions can be performed both proactively and on-demand. Proactive management involves a scheduling function, where the performance monitoring probes can be triggered on a recurring basis. Since the basic performance monitoring functions involved are the same, we make no distinction between proactive and on-demand functions in this section.

5.1 Packet Loss

Given that TRILL provides inherent support for multipoint-to-multipoint connectivity, then packet loss cannot be accurately measured by means of counting user data packets. This is because user packets can be delivered to more RBridges or more ports than are necessary (e.g. due to broadcast, un-pruned multicast or unknown unicast flooding). As such, a statistical means of approximating packet loss rate is required. This can be achieved by sending "synthetic" (i.e. TRILL OAM) packets that are counted only by those ports (MEPs) that are required to receive them. This provides a statistical approximation of the number of data frames lost, even with multipoint-to-multipoint connectivity. TRILL OAM mechanisms for synthetic packet loss measurement should follow the statistical considerations specified in [MEF35], especially with regards to the volume/frequency of synthetic traffic generation and associated impact on packet loss count accuracy.

Packet loss probes must be initiated from a MEP and must target a MEP. This function should be supported on sections, as defined in [RFC6905]. When packet loss is measured over a section, and the initiating MEP does not coincide with the edge (ingress) RBridge, the MEP must use the edge RBridge nickname instead of the local RBridge nickname on the associated loss measurement messages. The user must supply the edge RBridge nickname as part of the operation parameters.

TRILL OAM mechanisms should support one-way and two-way packet loss monitoring. In one-way monitoring, a source RBridge triggers packet loss monitoring messages to a target RBridge, and the latter is responsible for calculating the loss in the direction from the source RBridge towards the target RBridge. In two-way monitoring, a source RBridge triggers packet loss monitoring messages to a target RBridge, and the latter replies to the source with response messages. The source RBridge can then monitor packet loss in both directions (source to target and target to source).

5.2 Packet Delay

Packet delay is measured by inserting time-stamps in TRILL OAM packets. In order to ensure high accuracy of measurement, TRILL OAM must specify the time-stamp location at fixed offsets within the OAM packet in order to facilitate hardware-based time-stamping. Hardware implementations must implement the time-stamping function as close to the wire as practical in order to maintain high accuracy.

TRILL OAM mechanisms should support one-way and two-way packet delay monitoring. In one-way monitoring, a source RBridge triggers packet delay-monitoring messages to a target RBridge, and the latter is

responsible for calculating the delay in the direction from the source RBridge towards the target RBridge. This requires synchronization of the clocks between the two RBridges. In two-way monitoring, a source RBridge triggers packet delay monitoring messages to a target RBridge, and the latter replies to the source with response messages. The source RBridge can then monitor packet delay in both directions (source to target and target to source) as well as the cumulative round-trip delay. In this case as well, monitoring the delay in a single direction requires clock synchronization between the two RBridges. Whereas monitoring the round-trip delay does not require clock synchronization. Mechanisms for clock synchronization between RBridges are outside the scope of this document.

6. Operational and Manageability Considerations

6.1 TRILL OAM Configuration

RBridges may be configured to enable TRILL OAM functions via the device Command Line Interface (CLI) or through one of the defined management protocols, such as SNMP [RFC3410] or NETCONF [RFC6241].

In order to maintain the plug-and-play characteristics of TRILL, the number of parameters that need to be configured on RBridges, in order to activate TRILL OAM, should be kept to a minimum. To that end, TRILL OAM mechanisms should rely on default values and auto-discovery mechanisms (e.g. leveraging IS-IS) where applicable. The following is a non-exhaustive list of configuration parameters that apply to TRILL OAM.

6.1.1 Maintenance Domain Parameters

- Maintenance Domain Name
  An alphanumeric name for the Maintenance Domain. The recommended default value is the character string "DEFAULT".

- Maintenance Domain Level
  An integer in the range 0 to 7 indicating the Level at which the Maintenance Domain is to be created. Default value is 0.

6.1.2 Maintenance Association Parameters

- MA Name
  An alphanumeric name that uniquely identifies the Maintenance Association. This is an IETF [RFC2579] DisplayString, with the exception that character codes 0-31 (decimal) are not used. The recommended default value is a character string set to the value of the VLAN or fine grain label as "vl" or "fgl" concatenated with the

VLAN ID or FGL ID as an unsigned decimal integer.

- List of MEP Identifiers
  A list of the identifiers of the MEPs that belong to the MA. This is optional, and required only if the operator wants to detect missing MEPs as part of the Continuity Check function.

6.1.3 Maintenance Endpoint Parameters

- MEP Identifier
  An integer, unique over a given Maintenance Association, identifying a specific MEP. [802.1Q] CFM limits this to the range 1 to 8191. This document recommends expanding the range from 1 to 65535 so that the RBridge Nickname can be used as a default value. This will help keep TRILL OAM low-touch in terms of configuration overhead.

- Direction
  Indicates whether this is an UP MEP or DOWN MEP.

- Associated Interface
  Specifies the interface on which the MEP is configured.

- MA context
  Specifies the Maintenance Association to which the MEP belongs.

6.1.4 Continuity Check Parameters (applicable per MA)

- Transmission interval
  Indicates the interval at which Continuity Check messages are sent by a MEP.

- Loss threshold
  Indicates the number of consecutive Continuity Check messages that a MEP must not receive from any one of the other MEPs in its MA before indicating either a MEP failure or a network failure. Recommended default value is 3.

- VLAN / Fine grain label / Flow parameters
  The VLAN or fine grain label and flow parameters to be used in the Continuity Check messages.

- Hop Count
  The Hop Count to be used in the Continuity Check messages.

6.1.5 Connectivity Verification Parameters (applicable per operation)

- MA context
  Specifies the Maintenance Association in which the Connectivity Verification operation is to be performed.

- Target RBridge Nickname (unicast), Tree Identifier (Multicast) and IP multicast group
  For unicast, the Nickname of the RBridge that is the target of the Connectivity Verification operation. For multicast, the target Tree Identifier for un-pruned tree verification or the Tree Identifier and IP multicast group (S, G) or (*, G) for pruned tree verification.

- VLAN / Fine grain label / Flow parameters
  The VLAN or fine grain label and flow parameters to be used in the Connectivity Verification message.

- Operation timeout value
  The timeout on the initiating MEP before the Connectivity Verification operation is declared to have failed. The recommended default value is 5 seconds.

- Repeat Count
  The number of Connectivity Verification messages that must be transmitted per operation. The recommended default value is 1.

- Hop Count
  The Hop Count to be used in the Connectivity Verification messages.

- Reply Mode
  Indicates whether the response to the Connectivity Verification operation should be sent in-band or out-of-band.

- Scope List (Multicast)
  List of MEP Identifiers that must respond to the message.

6.1.6 Fault Isolation Parameters (applicable per operation)

- MA context
  Specifies the Maintenance Association in which the Fault Isolation operation is to be performed.

- Target RBridge Nickname (unicast), Tree Identifier (Multicast) and IP multicast group
  For unicast, the Nickname of the RBridge that is the target of the Fault Isolation operation. For multicast, the target Tree Identifier for un-pruned tree tracing or the Tree Identifier and IP multicast group (S, G) or (*, G) for pruned tree tracing.

- VLAN / Fine grain label / Flow parameters
  The VLAN or fine grain label and flow parameters to be used in the Fault Isolation messages.

- Operation timeout value
  The timeout on the initiating MEP before the Fault Isolation operation is declared to have failed. The recommended default value is 5 seconds.

- Hop Count
  The Hop Count to be used in the Fault Isolation messages.

- Reply Mode
  Indicates whether the response to the Fault Isolation operation should be sent in-band or out-of-band.

- Scope List (Multicast)
  List of MEP Identifiers that must respond to the message.

6.1.7 Packet Loss Monitoring

- MA context
  Specifies the Maintenance Association in which the Packet Loss Monitoring operation is to be performed.

- Target RBridge Nickname
  The Nickname of the RBridge that is the target of the Packet Loss Monitoring operation.

- VLAN / Fine grain label / Flow parameters
  The VLAN or fine grain label and flow parameters to be used in the Packet Loss monitoring messages.

- Transmission Rate
  The transmission rate at which the Packet Loss monitoring messages are to be sent.

- Monitoring Interval
  The total duration of time for which a single Packet Loss monitoring probe is to continue.

- Repeat Count
  The number of probe operations to be performed. For on-demand monitoring, this is typically set to 1. For proactive monitoring this may be set to allow for infinite monitoring.

- Hop Count
  The Hop Count to be used in the Packet Loss monitoring messages.

- Mode
  Indicates whether one-way or two-way loss measurement is required.

6.1.8 Packet Delay Monitoring

- MA context
  Specifies the Maintenance Association in which the Packet Delay monitoring operation is to be performed

- Target RBridge Nickname
  The Nickname of the RBridge that is the target of the Packet Delay monitoring operation.

- VLAN / Fine grain label / Flow parameters
  The VLAN or fine grain label and flow parameters to be used in the Packet Delay monitoring messages.

- Transmission Rate
  The transmission rate at which the Packet Delay monitoring messages are to be sent.

- Monitoring Interval
  The total duration of time for which a single Packet Delay monitoring probe is to continue.

- Repeat Count
  The number of probe operations to be performed. For on-demand monitoring, this is typically set to 1. For proactive monitoring this may be set to allow for infinite monitoring.

- Hop Count
  The Hop Count to be used in the Packet Delay monitoring messages.

- Mode
  Indicates whether one-way or two-way delay measurement is required.

6.2 TRILL OAM Notifications

TRILL OAM mechanisms should trigger notifications to alert operators to certain conditions. Such conditions include but are not limited to:

- Faults detected by proactive mechanisms.
- Reception of event-driven defect indications.
- Logged security incidents pertaining to the OAM message channel.
- Protocol errors (e.g. caused by mis-configuration).

Notifications generated by TRILL OAM mechanisms may be via SNMP,

Syslog messages [RFC5424] or any other standard management protocol that supports asynchronous notifications.

6.3 Collecting Performance Monitoring Metrics

When performing the optional TRILL OAM Performance Monitoring functions, two RBridge designations are involved: a source RBridge and a target RBridge. The source RBridge is the one from which the Performance Monitoring probe is initiated, and the target RBridge is the destination of the probe. The goal being to monitor performance characteristics between the two RBridges. The RBridge from which the network operator can extract the results of the probe (i.e. the Performance Monitoring metrics) depends on whether one-way or two-way performance monitoring functions are performed:

In the case of one-way performance monitoring functions, the metrics will be available at the target RBridge.

In the case of two-way performance monitoring functions, all the metrics will be available at the source RBridge, and a subset will be available at the target RBridge. More specifically, metrics in the direction from source to target as well as the direction from target to source will be available at the source RBridge. Whereas, metrics in the direction from source to target will be available at the target RBridge.

7. Security Considerations

TRILL OAM must provide mechanisms for:

- Preventing denial of service attacks caused by exploitation of the OAM message channel, where a rogue device may overload the RBridges and the network with OAM messages. This could lead to interruption of the OAM services and in the extreme case disrupt network connectivity. Mechanisms such as control-plane policing combined with shaping or rate limiting of OAM messaging can be employed to mitigate this.

- Optionally authenticate at communicating endpoints (MEPs and MIPs) that an OAM message has originated at an appropriate communicating endpoint.

- Preventing TRILL OAM packets from leaking outside of the TRILL network or outside their corresponding Maintenance Domain. This can be done by having MEPs implement a filtering function based on the Maintenance Level associated with received OAM packets.

For general TRILL Security Considerations, see [RFC6325].

8. IANA Considerations

This document requires no IANA Actions. RFC Editor: Please delete this section before publication.

9. Acknowledgements

We thank Gayle Noble, Dan Romascanu, Olen Stokes, Susan Hares, Ali Karimi and Prabhu Raj for their thorough review of this work and their comments.

10. References

10.1 Normative References

[RFC6905] Senevirathne, et al., "Requirements for Operations, Administration and Maintenance (OAM) in Transparent Interconnection of Lots of Links (TRILL)", RFC 6905, March 2013.

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

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for

Network Interconnect Devices", RFC 2544, March 1999.

[RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, April 1999.

[RFC6291] Andersson et al., BCP 161 "Guidelines for the Use of the "OAM" Acronym in the IETF", June 2011.

[RFC6327] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Dutt, D., and V. Manral, "Routing Bridges (RBridges): Adjacency", RFC 6327, July 2011.

[TRILL-FGL] D. Eastlake et al., "TRILL Fine-Grained Labeling", draft-ietf-trill-fine-labeling, work in progress.

[802.1Q] "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Bridges and Virtual Bridge Local Area Networks", IEEE Std 802.1Q-2011, 31 August 2011.

[802] "IEEE Standard for Local and Metropolitan Area Networks - Overview and Architecture", IEEE Std 802-2001, 8 March 2002.

10.2 Informative References

[Y.1731] "ITU-T Recommendation Y.1731 (02/08) - OAM functions and mechanisms for Ethernet based networks", February 2008.

[ISO/IEC 7498-4] "Information processing systems -- Open Systems Interconnection -- Basic Reference Model -- Part 4: Management framework", ISO/IEC, 1989.

[TRILL-BFD] V. Manral, et al., "TRILL (Transparent Interconnetion of Lots of Links): Bidirectional Forwarding Detection (BFD) Support", draft-ietf-trill-rbridge-bfd-07, work in progress, July 2012.

[TRILL-OAM] T. Senevirathne, et al., "TRILL Fault Management", draft-tissa-trill-oam-fm-01, work in progress, February 2013.

[TRILL-IP] M. Wasserman, et al., "Transparent Interconnection of Lots of Links (TRILL) over IP", draft-mrw-trill-over-ip-02, work in progress, September 2012.

[RFC1661] Simpson, W., Ed., "The Point-to-Point Protocol (PPP)", STD

51, RFC 1661, July 1994.

[RFC6361] Carlson & Eastlake, "PPP Transparent Interconnection of Lots of Links (TRILL) Protocol Control Protocol", RFC 6361, August 2011.

[RFC5880] Katz & Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

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

[TRILL-ML] Perlman, et al., "Alternatives for Multilevel TRILL", draft-perlman-trill-rbridge-multilevel-06, work in progress, July 2013.

[RFC3410] Case, et al., "Introduction and Applicability Statements for Internet-Standard Management Framework", RFC 3410, December 2002.

[RFC6241] Enns, et al., "Network Configuration Protocol (NETCONF)", RFC 6241, June 2011.

[RFC5424] Gerhards, R., "The Syslog Protocol", RFC 5424, March 2009.

[MEF35] "MEF 35 - Service OAM Performance Monitoring Implementation Agreement", Metro Ethernet Forum, April 2012.

Authors' Addresses

Samer Salam
Cisco
595 Burrard Street, Suite 2123
Vancouver, BC V7X 1J1, Canada
Email: ssalam@cisco.com

Tissa Senevirathne
Cisco
375 East Tasman Drive
San Jose, CA 95134, USA
Email: tsenevir@cisco.com

Sam Aldrin
Huawei Technologies
2330 Central Expressway
Santa Clara, CA 95050, USA
Email: sam.aldrin@gmail.com

Donald Eastlake
Huawei Technologies
155 Beaver Street
Milford, MA 01757, USA
Tel: 1-508-333-2270
Email: d3e3e3@gmail.com