TRILL Working Group
INTERNET-DRAFT
Intended Status: Informational

Samer Salam
Tissa Senevirathne
Cisco

Sam Aldrin
Donald Eastlake
Huawei

Expires: April 22, 2013
October 19, 2012

TRILL OAM Framework
draft-salam-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

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 Ethernet 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 supports multipathing of unicast as well as multicast traffic.

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, namely Fault Management and Performance Management, in TRILL networks. These primarily map to the "Operations" and "Maintenance"

part of OAM.

The draft provides a generic framework for a comprehensive solution that meets the requirements outlined in [TRILL-OAM-REQ]. However, specific mechanisms to address these requirements are considered to be outside the scope of this document.

1.1 Terminology

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].

In addition, the following acronyms are used:

BFD - Bidirectional Forwarding Detection [RFC5880]
CFM - Connectivity Fault Management [802.1Q]
FGL - Fine Grained Label(ing) [TRILL-FGL]
IEEE - Institute for Electrical and Electronic Engineers
IP - Internet Protocol, includes both IPv4 and IPv6
L2VPN - Layer 2 Virtual Private Network
LAN - Local Area Network
MEG - Maintenance Entity Group

MEP - Maintenance End Point
MIP - Maintenance Intermediate Point
MP - Maintenance Point (MEP or MIP)
OAM - Operations, Administration, and Maintenance [RFC6291]
RBridge - Routing Bridge, a device implementing TRILL [RFC6325]
TRILL - Transparent Interconnection of Lots of Links [RFC6325]
TRILL Switch - an alternate name for an RBridge
VLAN - Virtual LAN

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:

[TRILL-OAM-REQ] defines the requirements for TRILL OAM which serve as the basis for this framework.

[802.1Q] specifies the Connectivity Fault Management 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.

[RFC6136] specifies a reference model for OAM as it relates to L2VPN services, pseudowires and associated Public Switched Network tunnels. The document also specifies OAM requirements for L2VPN services.

[RFC6371] describes a framework to support a comprehensive set of OAM procedures that fulfill the MPLS-TP OAM requirements for fault, performance, and protection-switching management and that do not rely on the presence of a control plane.

[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 frames such as Ethernet [RFC6325], PPP [RFC6361], or MPLS. 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:


```
                           LAN                LAN
           +---+   +---+  ======  +---+  =============  +---+
    +--+   |   |   |   | | +--+ | |   | | +--+   +--+ | |   |   +--+
    |B1|---|RB1|---|RB2|---|B2|---|RB3|---|B3|---|B4|---|RB4|---|B5|
    +--+   |   |   |   | | +--+ | |   | | +--+   +--+ | |   |   +--+
           +---+   +---+  ======  +---+  =============  +---+
```

a. Ethernet CFM (Client Layer) on path over the TRILL campus

```
       >---o------------------------------------------------o---<


    b. TRILL OAM (Network Layer)
               >------o-----------o---------------------<


    c. Ethernet CFM (Transport Layer) on interior Ethernet LANs
                      >---o--o---<    >---o--o---o--o---<


    d. BFD (Media Independent Link Layer)
               #---#   #----------#   #-----------------#


    e. Link OAM (Media Dependent Link Layer)
       *---*   *---*   *---*  *---*   *---*  *---*  *---*   *---*


    Legend:  > MEP    o MIP    # BFD Endpoint    * Link OAM Endpoint
```

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 BFD sits a layer above Link OAM, which is media specific. BFD provides fast convergence characteristics to TRILL networks. It is worth noting that 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 bound to a specific TRILL output port as explained below.

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 Up state [RFC6327].

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

For example, assume there are five parallel equal cost links between RB1 and RB2 that have not been aggregated. (Links that are aggregated with [802.1AX] appear to TRILL to be a single link accessible through a single TRILL port.) However, RB1 is only capable of doing up to 4-way ECMP. TRILL OAM messages, as dispatched by the TRILL Routing function, will use 4 of the 5 links. But it is desirable to be able to monitor the fifth link to be sure it is available for failover. TRILL BFD messages sent by RB1 will use the output port to which their session is bound. RB1 can easily monitor all 5 links to RB2 by using a TRILL BFD session bound to each of the 5 output ports.

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 e.g., 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 modeled as a shim situated between the Extended Internal Sublayer Service (EISS) in [802.1Q] and the RBridge Forwarding Engine function, on a virtual port with no physical layer (Null PHY). 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 whereas single lines represent flow of information only.

While 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.

```
           +-----------------------------------------------+----
           |              (Flow of OAM Messages)       RBridge
           |             +----------------------+
           |             |+-------------------+||  Forwarding Engine,
           |             ||                    ||  IS-IS, Etc.
           |             ||                    ||  Processing of native
           |             V                      V  and TRILL frames
           +---------------------------------------------+-----
            ||          ||                     ||  ...other trunk ports
            ||    +------------+             +------------+
            ||    | TRILL OAM  |             |            |
            ||    | Processing |             | Port VLAN  |
            ||    +------------+             | Processing |
            ||                               |            |
            |+-------------------------+     |            |
            +-------------------------+|     +------------+ <-- ISS
                                      ||     |802.1/802.3 |
   +-------------------------------+  ||     |Low Level   |
   |             MAC Relay         |  ||     |Control     |
   +----||------------------||-----+  ||     |Frame       |
   +----||------+ ... +-----||-----+  ||     |Processing, |
   |            |     |            |  ||     |Port/Link   |
   | Port VLAN  |     | Port VLAN  |  ||     |Control     |
   | Processing |     | Processing |  ||     |Logic       |
   +------------+     +------------+  ||     +------------+
   |802.1/802.3 |     |802.1/802.3 |  ||     | 802.3PHY   |
   |Low Level   |     |Low Level   |  ||     +------------+
   |Control     |     |Control     |  ||
   |Frame       |     |Frame       |  ||
   |Processing, |     |Processing, |  ||
   |Port/Link   |     |Port/Link   |  ||
   |Control     |     |Control     |  ||
   |Logic       |     |Logic       |  ||
   +------------+     +------------+  ||
   | 802.3PHY   | ... | 802.3PHY   |  ||
   +------------+     +------------+  ||
     Access/Shared            ||        ||
       Ports                  |+--------+|
                              +----------+
```

Figure 2: TRILL OAM in RBridge Port Model

Note that there is a single virtual interface, per RBridge, that hosts the TRILL OAM shim. The rationale for this model is discussed in section 2.6 "MEPs and MIPs".

2.3 Network, Service and Flow OAM

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

Network OAM mechanisms provide fault and performance management functions in the context of a representative 'test' VLAN or fine grained label [TRILL-FGL]. The test VLAN can be thought of as a management or diagnostics VLAN which 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 granular fault and performance management capabilities, where OAM functions are performed in the context of end station service VLANs or fine grained labels and user flows. 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], [RFC6136], [RFC5654], etc... all define 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 faults in connectivity or performance 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 the network is managed by a single operator. 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 e.g. as discussed in [TRILLML], then the introduction of multiple Maintenance Domains and Maintenance Domain hierarchies becomes useful to map and contain administrative

boundaries. When considering multi-domain scenarios, the following rules must be followed: TRILL OAM domains MUST NOT overlap, but MUST either be disjoint or nest to form a hierarchy (i.e. a higher Maintenance Domain MAY completely engulf a lower Domain). A Maintenance Domain is typically identified by a Domain Name and a Maintenance Level (a numeric identifier). The larger the Domain, the higher the Level number.

```
+--------------------+  +----------------+  +--------------------+
|                   |  |      TRILL      |  |                    |
|       Site 1     +-----+Interconnect +-----+    Site 2          |
|       TRILL      | RB |  Network     | RB |    TRILL           |
|      (Level 1)   +-----+  (Level 2)   +-----+   (Level 1)       |
|                   |  |                |  |                     |
+--------------------+  +----------------+  +--------------------+

<-------------------------End-to-End Domain-------------------->

<----Site Domain----> <--Interconnect --> <----Site Domain---->
                          Domain
```

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 which define a Maintenance Entity are known as Maintenance End Points (MEPs) - see section 2.6 below. The set of Maintenance Entities that belong to the same Maintenance Domain are referred to as a Maintenance Entity Group (MEG). On the network path in between MEPs, there can be zero or more intermediate points, called Maintenance Intermediate Points (MIPs). MEPs and MIPs are associated with the MEG and can be part of more than one ME in a given MEG.

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 proactively or on-demand based on operator invocation. 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.

```
   TRILL Site 1          Interconnect       TRILL Site 2
+------------------+ +-------------------+ +------------------+
|                  | |                   | |                  |
|  +----+  +----+  +----+  +----+  +----+  +----+  +----+  +----+  |
|  |RB1|--|RB2|--|RB3|--|RB4|--|RB5|--|RB6|--|RB7|--|RB8|  |
|  +----+  +----+  +----+  +----+  +----+  +----+  +----+  +----+  |
|                  | |                   | |                  |
+------------------+ +-------------------+ +------------------+
```

```
<E------------I-------------------I-------------E>

<E------I----E><E----I-------I----E><E-----I-----E>


Legend E: MEP      I: MIP
```

Figure 4: MEPs and MIPs

It is worth noting that 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 "802.1Q Port VLAN Processing" block.

The model of Section 2.2 implies that a single MEP and/or MIP per MEG can be instantiated per RBridge. This simplifies implementations and enables TRILL OAM to perform management functions on sections, as specified in [TRILL-OAM-REQ], while maintaining the simplicity of a single TRILL OAM Maintenance Domain. We do not distinguish between Up MPs and Down MPs (as defined in [802.1Q]) in this framework. Given that the MPs always reside on a special virtual port with no PHY layer, MP directionality is irrelevant.

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 MEG. 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 MEG. Certain OAM functions, e.g. pruned tree verification, require addressing a subset of the MPs in a MEG. 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 the unwieldy response 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 frame 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 frame format proposed below provides the necessary flexibility to exercise the data path as closely as possible to actual data packets.


```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
.      Link Header              . Variable
|                               |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
+      TRILL Header             + 8 bytes
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
.   Flow Entropy                . Fixed Size
.                               .
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       OAM EtherType           | 2 bytes
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
.   OAM Message Channel         . Variable
.                               .
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
.    Link Trailer               . Variable
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: OAM Frame Format

The TRILL Header is as specified in [RFC6325] and the Link Header and Trailer are as specified for the link technology. (Link types standardized so far are [RFC6325] for Ethernet and [RFC6361] for PPP). These fields need to be as similar as practical to the Link Header/Trailer and TRILL Header of the normal TRILL data frame 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.

For a Layer 2 flow (i.e. non-IP) the Flow Entropy must specify the 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 Ethernet header, the IP header and UDP or TCP header fields.

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 should provide a mechanism for discovering the flow entropy parameters by querying the RBridges dynamically.

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 e.g., 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.

It may be possible to use a combination of one of the unused fields or bits in the TRILL Header and the OAM EtherType to identify TRILL OAM messages.

[RFC6325] does not specify any method of identifying OAM messages. Hence, for backwards compatibility reasons, TRILL OAM solutions must provide methods to identify OAM messages through the use of well-known patterns in the Flow Entropy field; for e.g., by using a reserved MAC address as the inner MAC SA.

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 MEG, 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 (LoC) to one or more remote MEPs
- Unexpected connectivity between isolated VLANs (mismerge)
- Unexpected connectivity to one or more remote MEPs
- Period mis-configuration

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 e.g., consider a simple network comprising of four RBridges connected in tandem: 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 (LoC) defect. RDI is transmitted in the direction of the failure. For e.g., consider the same tandem network of the previous section. 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.

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 and continue 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 [TRILL-OAM-REQ], 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 traffic patterns 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 representative test VLAN 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 [TRILL-OAM-REQ]. 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 multidestination distribution tree for reachability. This function should include mechanisms to prevent reply storms from overwhelming the initiating RBridge. This may be done, for e.g., by staggering the replies. 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 [TRILL-OAM-REQ], multicast connectivity verification must provide the following granularity of operation:

A. Un-pruned Tree

- Connectivity verification for un-pruned multidestination distribution tree. The operator in this case supplies the tree identifier (root RBridge nickname) and campus wide diagnostic VLAN.

B. Pruned Tree

- Connectivity verification for a VLAN or fine-grain label in a given multidestination 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 multidestination 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.

5. Performance Management

Performance Management functions can be performed both proactively and on-demand. Proactive management involves a scheduling function, where the performance management probes can be triggered on a recurring basis. Since the basic performance management 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.

Packet loss probes must be initiated from a MEP and must target a MEP. This function must be supported on sections, as defined in [TRILL-OAM-REQ]. 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.

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.

6. Security Considerations

TRILL OAM must provide mechanisms for:

- Preventing denial of service attacks caused by exploitation of the OAM message channel.

- Optionally authenticate communicating endpoints (MEPs and MIPs)

- 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].

7. IANA Considerations

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

8. Acknowledgements

We invite feedback and contributors.

9. References

9.1 Normative References

[TRILL-OAM-REQ] Senevirathne, "Requirements for Operations, Administration and Maintenance (OAM) in TRILL", draft-tissa-trill-oam-req, work in progress.

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

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

[RFC6136] Sajassi, A., Ed., and D. Mohan, Ed., "Layer 2 Virtual Private Network (L2VPN) Operations, Administration, and Maintenance (OAM) Requirements and Framework", RFC 6136, March 2011.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 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.

[RFC6371] Busi & Allan, "Operations, Administration, and Maintenance Framework for MPLS-Based Transport Networks", RFC 6371, September 2011.

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

9.2 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, work in progress, June 2012.

[TRILL-OAM] T. Senevirathne, et al., "Use of 802.1ag for TRILL OAM Messages", draft-tissa-trill-8021ag, work in progress, June 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