TRILL Working Group
Internet Draft
Intended status: Informational

Tissa Senevirathne
CISCO
David Bond
IBM
Sam Aldrin
Yizhou Li
Huawei
Rohit Watve
CISCO
Anoop Ghanwani
DELL
Jon Hudson
Brocade
Naveen Nimmu
Broadcom
Radia Perlman
Intel
Tal Mizrahi
Marvell

June 29, 2012
Expires: December 2012

Requirements for Operations, Administration and Maintenance (OAM) in TRILL
draft-tissa-trill-oam-req-02.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

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

Copyright Notice

Abstract

OAM (Operations, Administration and Maintenance) is a general term used to identify functions and toolsets to troubleshoot and monitor networks. This document presents, OAM Requirements applicable to TRILL. Also presented in this document is the proposed frame structure for TRILL OAM messages.

Table of Contents

1. Introduction

OAM (Operations, Administration and Maintenance) generally covers various production aspects of a network. In this document we use the term OAM as defined in [RFC6291].

Success of any mission critical network depends on the ability to proactively monitor networks for faults, performance, etc. as well as its ability to efficiently and quickly troubleshoot defects and failures. A well-defined OAM toolset is a vital requirement for wider adoption of TRILL as the next generation data forwarding technology in larger networks such as data centers.

In this document we define the Requirements for TRILL OAM. Also the proposed format for OAM frames is presented.

It is assumed that the readers are familiar with the OAM concepts and terminologies defined in other OAM standards such as [802.1ag], [RFC5860]. This document does not attempt to redefine the terms and concepts specified elsewhere.

1.1. Contributors

The following members were part of the design team that produced this document. Their names are listed below in alphabetical order. Anoop Ghanwani, David Bond, Donald Eastlake 3rd, Jon Hudson, Naveen Nimmu, Radia Perlman, Rohit Watve, Sam Aldrin, Shivakumar Sundaram, Tal Mizrahi, Thomas Narten, Tissa Senevirathne, Yizhou Li.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119]. Although this document is not a protocol specification, the use of this language clarifies the instructions to protocol designers producing solutions that satisfy the requirements set out in this document.

3. Terminology

Section: The term Section refers to a partial segment of a path between any two given RBridges. As an example, consider the case where RB1 is connected to RBx via RB2,RB3 and RB4. The segment between RB2 to RB4 is referred to as a Section of the path RB1 to RBx.

Flow: The term Flow indicates a set of packets that share the same path and per-hop behavior (such as priority). A flow is typically identified by a portion of the inner payload that affects the hop-by hop forwarding decisions. This may contain Layer 2 through Layer 4 information.

All Least Cost Paths: The term "all least cost paths" refers to all potentially available least cost paths from a given source to a specified destination RBridge as determined by the TRILL network topology learned from IS-IS. ECMP: Equal Cost Multi Path.

Connectivity: The term connectivity indicates reachability between an arbitrary RBridge RB1 and any other RBridge RB2. The specific path can be either explicit (e.g. specific flow) or unspecified. Unspecified means taking whatever the path connectivity verification message happen to take.

Continuity Verification: Continuity Verification refers to proactive verification of Connectivity between two RBridges at periodic intervals and generation of explicit notification when Connectivity failures occur.

Fault: The term Fault refers to an inability to perform a required action, e.g., an unsuccessful attempt to deliver a packet.

Defect: The term Defect refers to an interruption in the normal operation, such that over a period of time no packets are delivered successfully.

Failure: The term Failure refers to the termination of the required function over a longer period of time. Persistence of a defect for a period of time is interpreted as a failure.

4. OAM Requirements

4.1. Data Plane

OAM frames, utilized for connectivity verification, continuity checks, performance measurements, etc., MUST follow the same path as the specified flow.

OAM frames, utilized for connectivity verification, continuity checks, performance measurements, etc., when a specific flow is not specified MUST follow whatever the path TRILL choose based on current topology, per-hop forwarding behavior and default flow entropy.

RBridges MUST have the ability to identify OAM frames destined for them or which require processing by the OAM plane from normal data frames.

RBridges MUST have the ability to differentiate between OAM frames and data frames experiencing errors.

TRILL OAM frames MUST be confined to the TRILL domain and MUST NOT be forwarded out of the TRILL domain. E.g. they must not be sent as native frames on an end station service enabled port

OAM MUST have the capability to provide services for both IP and non-IP flows.

OAM MUST be able to function in IP and non-IP infrastructure.

4.2. Connectivity Verification

4.2.1. Unicast

OAM MUST have the ability to verify connectivity from an arbitrary RBridge RB1 to any other RBridge RB2.

OAM MUST have the ability to verify connectivity from an arbitrary RBridge RB1 to any other RBridge RB2 for a specific flow.

An RBridge SHOULD have the ability to verify the above connectivity tests on sections. As an example, assume RB1 is connected to RB5 via RB2, RB3 and RB4. An operator SHOULD be able to verify the RB1 to RB5 connectivity on the section from RB3 to RB5. The difference is that the ingress and egress TRILL nicknames in this case are RB1 and RB5 as opposed to RB3 and RB5, even though the message itself may originate at RB3.

OAM MUST have the ability to invoke the above functions on-demand.

4.2.2. Multicast

OAM MUST have the ability to verify connectivity, from an arbitrary RBridge RB1, to either to a specific set of RBridges or all member RBridges, for a specified multicast tree. This functionality is referred to as verification of the un-pruned multicast tree.

OAM MUST have the ability to verify connectivity, from an arbitrary RBridge RB1, to either to a specific set of RBridges or all member RBridges, for a specified multicast tree and for a specified flow. This functionality is referred to as verification of the pruned tree.

OAM MUST have the ability to invoke the above functions on-demand.

4.3. Continuity Check

[RFC5860] defines Continuity Check as the ability of end points to monitor liveliness of a path or a section of a path [RFC5860]. We use similar semantics in this document where end points are the ingress or egress RBridges.

OAM MUST provide functions that allow any arbitrary RBridge RB1 to perform a Continuity Check to any other RBridge.

OAM MUST provide functions that allow any arbitrary RBridge RB1 to perform a Continuity Check to any other RBridge for a specified flow.

OAM SHOULD provide functions that allow any arbitrary RBridge to perform a Continuity Check to any other RBridge over all available least cost paths.

OAM SHOULD provide the ability to perform a Continuity Check on sections of any path within the network.

OAM SHOULD provide the ability to perform a multicast Continuity Check for specified multi-destination tree(s) as well as specified multi-destination tree and flow combinations. The former is referred to as an un-pruned multi-destination tree Continuity Check and the latter is referred to as a pruned tree Continuity Check.

OAM implementations SHOULD support at least a minimum frequency of 1 second of Continuity check.

OAM implementations SHOULD support multiple concurrent Continuity sessions from and/or to the same RBridge.

4.4. Path Tracing

OAM MUST provide the ability to trace a path between any two RBridges per specified unicast flow.

OAM SHOULD provide the ability to trace all least cost paths between any two RBridges.

OAM SHOULD provide functionality to trace all branches of a specified multi-destination tree (un-pruned tree)

OAM SHOULD provide functionality to trace all branches of a specified multi-destination tree for a specified flow (pruned tree).

4.5. General Requirements

OAM MUST provide the ability to initiate and maintain multiple concurrent sessions for multiple OAM functions between any arbitrary RBridge RB1 to any other RBridge.

OAM MUST NOT require extensions to or modifications of the TRILL header.

OAM MUST provide a single OAM framework for all TRILL OAM functions

OAM, as practical and as possible, SHOULD provide a single framework between TRILL and other similar standards.

OAM MUST maintain related error and operational counters. SUCH counters MUST be accessible via network management applications (e.g. SNMP).

Operations of OAM MUST NOT result in errors on end devices.

OAM MAY be required to provide the ability to specify a desired response mode for a specific OAM message. The desired response mode can be either in-band, out-of band or none.

The OAM Framework MUST be extensible to future needs of TRILL and the needs of other standard organizations.

OAM MAY provide methods to verify control plane and forwarding plane alignments.

OAM SHOULD leverage existing OAM technologies, where practical.

4.6. Performance Monitoring

4.6.1. Packet Loss

In this document, term loss of a packet is used as defined in [RFC2680] (see Section 2.4 of RFC2680).

NOTE: Term simulated flow below indicates a flow that is generated by an RBRidge RB1 for OAM purposes. The fields of the simulated flow may or may not be identical to the actual data. However, simulated flow is required to follow the intended path.

OAM SHOULD provide the ability to measure packet loss statistics for a simulated flow from any arbitrary RBridge RB1 to any other RBridge.

OAM SHOULD provide the ability to measure packet loss statistics over a segment, for a simulated flow between any arbitrary RBridge RB1 to any other RBridge.

OAM SHOULD provide the ability to measure simulated packet loss statistics between any two RBridges over all least cost paths.

An RBridge SHOULD be able to perform the above packet loss measurement functions either proactively or on-demand.

4.6.2. Packet Delay

There are two types of packet delays -- one-way delay and two-way delay (Round Trip Delay).

One-way delay is defined in [RFC2679] as the time elapsed from the start of transmission of the first bit of a packet by an RBridge until the reception of the last bit of the packet by the destination RBridge.

Two-way delay is also referred to as Round Trip Delay is defined similar to [RFC2681]; i.e. the time elapsed from the start of transmission of the first bit of a packet by an RBridge until the reception of the last bit of the packet by the same RBridge.

OAM SHOULD provide functions to measure two-way delay between two RBridges for a specified flow.

OAM SHOULD provide functions to measure two-way delay between two RBridges for a specified flow over a specific section.

OAM MAY provide functions to measure one-way delay between two RBridges for a specified flow.

OAM MAY provide functions to measure one-way delay between two RBridges for a specified flow over a specific section.

4.7. ECMP Utilization

OAM MAY provide functionality to monitor the effectiveness of per-hop ECMP hashing. For example, individual RBridges could maintain counters that show how packets are being distributed across equal cost next hops for a specified destination RBridge or RBridges as a result of ECMP hashing.

4.8. Security and Operational considerations

Methods MUST be provided to protect against exploitation of OAM framework for security and denial of service attacks.

Methods SHOULD be provided to prevent OAM messages causing congestion in the networks. Periodically generated messages with

high frequencies may lead to congestion, hence methods such as shaping or rate limiting SHOULD be utilized.

4.9. Fault Indications

The term Fault refers to an inability to perform a required action, e.g., an unsuccessful attempt to deliver a packet [OAMOVER]. The unsuccessful attempt may be due to Hop Count expiry, invalid nickname, etc.

OAM MUST provide a Fault Indication framework to notify faults to the ingress RBRidge of the flow or other interested parties (such as syslog servers).

OAM MUST provide functions to selectively enable or disable different types of Fault Indications.

4.10. Defect Indications

[OAMOVER] defines "The term Defect refers to an interruption in the normal operation, such as a consecutive period of time where no packets are delivered successfully."

OAM SHOULD provide a framework for Defect Detection and Indication.

OAM implementations that provide Defect Indication MUST provide methods to selectively enable or disable Defect Detection per defect type.

OAM implementations that provide Defect Indication MUST provide methods to configure Defect Detection thresholds per different types of defects.

OAM implementations that provide Defect Indication facilities MUST provide methods to log defect indications to a locally defined archive such as log buffer or SNMP traps.

OAM implementations that provide Defect Indication facilities SHOULD provide a Remote Defect Indication framework that facilitates notifying the originator/owner of the flow experiencing the defect, which is the ingress RBridge.

Remote Defect Indication MAY be either in-band or out-of-band.

4.11. Live Traffic monitoring

OAM implementations MAY provide methods to utilize live traffic for troubleshooting and performance monitoring.

Implementations MAY leverage Data Driven CFM [802.1Q] or IPFIX [RFC5101] for the purpose of performance monitoring.

5. General Format of TRILL OAM Messages

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

As specified above, OAM Messages are required to follow the exact path as the data packets.

The proposed frame format for TRILL OAM messages is as follows.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                 |
.    Outer Header                 . (variable)
|                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                 |
+    TRILL Header                 + 8 bytes
|                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                 |
.   Flow Entropy                  . 128 bytes
.                                 .
|                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                 |
.   OAM Message Channel           . Variable
.                                 .
|                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1 Frame format of OAM Messages

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

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

Flow Entropy: This is a 128 byte Fixed size opaque field. The field MUST be padded with zeros when the flow entropy is less than 128 bytes. Flow entropy emulates the forwarding behavior of the desired data packets.

OAM Message Channel: This is a variable size section that carries OAM related information. Reusing existing OAM message definitions such as [RFC4379] and [802.1ag] will be explored.

5.1. Requirements of OAM Message channel

The OAM Message channel header MUST contain a version number

The OAM Message channel header MUST contain flags that facilitate hardware level processing. (e.g. indicate this is two-way delay measurement probe)

The OAM Message channel MUST contain time stamping in fixed locations to facilitate hardware level performance monitoring. (e.g. delay measurements).

The OAM Message channel MUST have the ability to be extensible to include future capabilities without requiring a change to the version of the message header. (e.g. include TLV structures)

The OAM Message header MUST contain a unique marker that allows for identifying the presence of the OAM channel. This marker MUST provide equal or better uniqueness compared to the IP checksum define in [RFC791]

The OAM Message SHOULD provide methods to include arbitrary data to test functions such as MTU testing.

6. Security Considerations

Security Requirements are specified in section 4.8.

7. IANA Considerations

<TBD>

8. References

8.1. Normative References

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

[OAMOVER] Mizrahi, T, et.al., "An Overview of Operations, Administration, and Maintenance (OAM) Mechanisms", draft-ietf-opsawg-oam-overview-06, Work in Progress, March 2012.

[RFC5860] Vigoureux, M., et.al., "Requirements for Operations, Administration and Maintenance (OAM) in MPLS Transport Networks", RFC5860, May 2010.

[RFC4377] Nadeau, T., et.al., "Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks", RFC 4377, February 2006.

8.2. Informative References

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

[RFC5101] Claise, B., "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", RFC5101, January 2008.

[RFC2680] Almes, G., et.al. "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.

[RFC2679] Almes, G., et.al. "A One-way Delay Metric for IPPM", RFC 2679, September 1999.

[RFC2681] Almes, G., et.al. "A Round-trip Delay Metric for IPPM", RFC 2681, September 1999.

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

[8021ag] IEEE, "Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management", 802.1ag, 2007.

[8021Q] IEEE, "Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks", IEEE Std 802.1Q-2011, August 31st , 2011.

[RFC791] "Internet Protocol", RFC 791, September 1981.

9. Acknowledgments

Special acknowledgments to IEEE 802.1 chair, Tony Jeffree for allowing us to solicit comments from IEEE 802.1 group. Also recognized are the comments received from IEEE group, Ayal Loir and others.

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

Authors' Addresses

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

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

David Bond
IBM
2051 Mission College Blvd
Santa Clara, CA 95054
USA

Phone: +1-603-339-7575
Email: mokon@mokon.net

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

Email: aldrin.ietf@gmail.com

Yizhou Li
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

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

Rohit Watve
CISCO Systems
375 East Tasman Drive
San Jose, CA 95134
USA.

Phone: +1-408-424-2091
Email: rwatve@cisco.com

Anoop Ghanwani
DELL
350 Holger Way
San Jose, CA 95134
USA.

Phone: +1-408-571-3500
Email: Anoop@duke.alumni.duke.edu

John Hudson
Brocade
120 Holger Way
San Jose, CA 95134
USA.

Email: jon.hudson@gmail.com

Naveen Nimmu
Broadcom
9th Floor, Building no 9, Raheja Mind space
Hi-Tec City, Madhapur,
Hyderabad - 500 081, INDIA

Phone: +1-408-218-8893
Email: naveen@broadcom.com

Radia Perlman
Intel Labs
2700 156th Ave NE, Suite 300,
Bellevue, WA 98007
USA.

Phone: +1-425-881-4824
Email: radia.perlman@intel.com

Tal Mizrahi
Marvell
6 Hamada St.
Yokneam, 20692 Israel

Email: talmi@marvell.com