

TRILL Working Group
Internet Draft
Intended status: Standards Track
Expires: April 2012

Y. Li
W. Hao
Huawei Technologies
D. Bond
IBM
V. Manral
Hewlett Packard Co.
October 31, 2011

OAM tool for RBridges: Multi-destination Ping
draft-yizhou-trill-multi-destination-ping-01.txt

Status of this Memo

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

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

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

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

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

This Internet-Draft will expire on April 31, 2009.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

Internet-Draft

OAM along Multi-destination Path

October 2011

carefully, as they describe your rights and restrictions with respect to this document.

Abstract

Unicast and multi-destination data frame may follow the different path in TRILL network. We need the ping and traceroute like applications for the connectivity testing and fault isolation on the multi-destination path in addition to the unicast path. This document specifies the format and handling of the new TRILL OAM protocol messages and TLVs which can be used for the multi-destination OAM.

Table of Contents

1.	Introduction	3
2.	Conventions used in this document.....	3
3.	Motivations	3
4.	RBridge Channel Message Format.....	4
5.	OAM Protocol Frame Format for Echo in the Long Format	4
6.	TLV Encodings	6
	6.1. Target RBridge	6
	6.2. Jitter	7
7.	Processing Echo Messages for Multi-destination Path.....	8
	7.1. Sending an echo request.....	8
	7.2. Receiving an echo request.....	9
	7.2.1. If H Flag Is Not Set.....	9
	7.2.2. If H Flag Is Set.....	10
	7.3. Sending an echo reply.....	11
	7.4. Receiving an echo reply.....	12
8.	Security Considerations.....	12
9.	IANA Considerations	12
10.	References	12
	10.1. Normative References.....	12
	10.2. Informative References.....	13
11.	Acknowledgments	13

1. Introduction

When RBridges are deployed in a real network, a number of applications are necessary for error detection/reporting and diagnostic purpose. TRILL RBridge channel [[RBridgeChannel](#)] was designed for carrying the OAM relevant messages. [[RBridgeOAM](#)] has defined the ping and traceroute applications for unicast path and also the error reporting mechanisms.

Multi-destination data path in TRILL network has different characteristics from the unicast path. One or more distribution trees are formed for multi-destination traffic. RBridges advertise their interests in receiving the traffic of the specific VLANs. The distribution tree may or may not be pruned based on VLAN ID. Troubleshooting on the multi-destination path is a desirable feature of TRILL OAM. This document specifies the messages and mechanisms used by multi-destination OAM.

2. Conventions used in this document

The same terminology and acronyms are used in this document as in [RF6325].

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC-2119](#) [[RFC2119](#)].

3. Motivations

In an RBridge campus, unicast and multi-destination traffic may follow different paths between the same ingress and egress RBridges. [[RBridgeOAM](#)] specifies some OAM along unicast path. For diagnostic purposes it is also desirable to check the connectivity between two or more RBridges along a particular distribution tree.

There are various things we want to test for multi-destination path.

- Along a distribution tree, who are the leaf nodes of an inner VLAN?

The leaf nodes here refer to the R Bridges announcing the given inner VLAN as their interested VLAN in INT-VLAN sub-TLV. It is useful when we want to check if the configuration/provisioning are consistent with the design.

- Along a distribution tree, check the connectivity from the ingress RBridge to one or more leaf nodes of an inner vlan. It can be used as the first step in diagnosis when we suspect multi-destination data path to certain RBridge fails. Transit nodes do not decapsulate the

multi-destination data frame; therefore we do not think it is much of interest to check the connectivity to any non-leaf R Bridges.

- Along a distribution tree, trace the multi-destination data path hop-by-hop to a target RBridge. It is useful when we want to find out where exactly is the failed hop.

This document specifies new messages and TLVs used by multi-destination OAM applications like multi-destination ping and traceroute. Processing of these messages is also discussed in the draft.

[4.](#) RBridge Channel Message Format

The RBridge Channel Header fields is as follows,

- o CHV (Channel Header Version): zero.
- o Channel Protocol: 0x006 (Echo in the Long Format) (TBD)
- o Flags: The SL and NA bits SHOULD be zero, the MH bit SHOULD be one
- o ERR: zero.

[5.](#) OAM Protocol Frame Format for Echo in the Long Format

The frame format is shown as follows. In the rest of this document, echo request and echo reply are brief ways to refer to the Echo Request in Long Format and Echo Reply in Long Format messages.

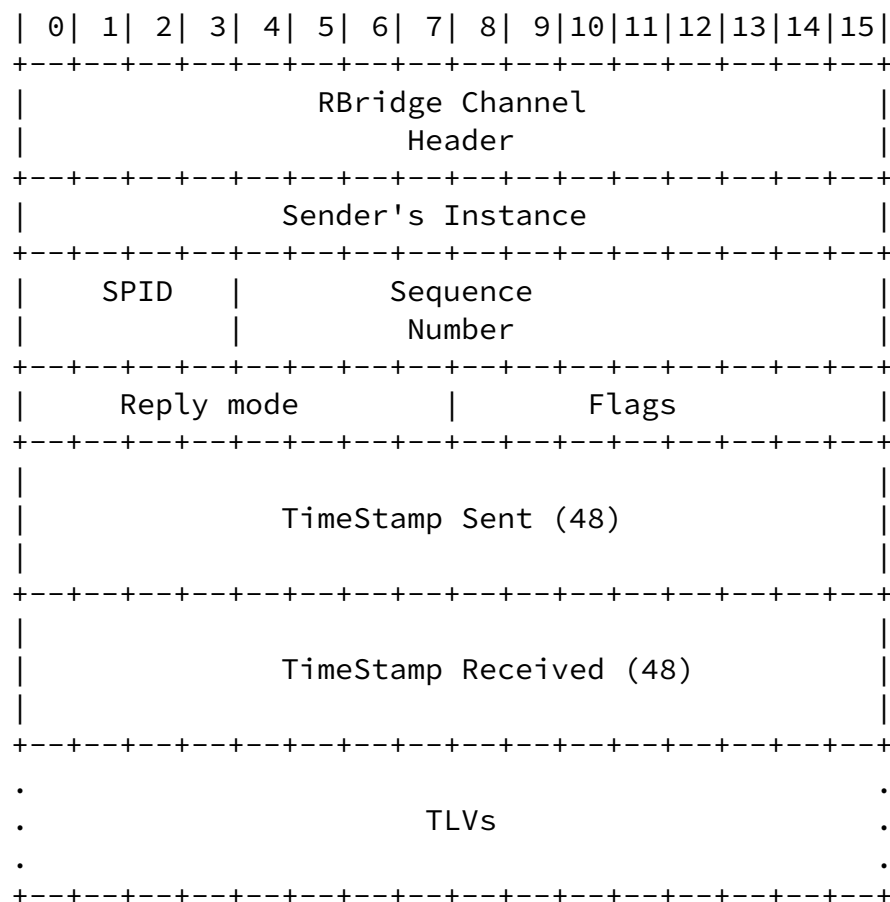


Figure 1 Echo Request with Long Format

o Sender's Instance: An instance ID used by sender to associate the echo operation with different application instances, e.g. different Telnet sessions. Echo reply should return the value

unchanged.

- o SPID:

- 1 - Echo Request in the Long Format
- 2 - Echo Reply in the Long Format

- o Sequence Number: An arbitrary 28-bit unsigned integer used to aid in matching reply messages to echo requests. It MAY be zero.

- o Reply Mode: Default is 2. It can take one of the following values.

- 1 - Do not reply. It can be used for one-way connectivity check. The receiving RBridge may perform monitoring and statistics collection on delay and/or jitter using one-way echo operation.

2 - Reply with Echo Reply in the Long Format and send back unicast in TRILL OAM channel. This value would be used by echo request in most cases.

- o Flags: A bit vector with the following format. Currently only the H (Respond Only When Hop Count is Zero) flag is defined. In practice, we set H flag to be 0 for ping type applications and 1 for traceroute type applications of multi-destination OAM. With H flag set, it will help to prevent the duplicate echo replies from the same RBridge triggered by echo request with different hop count value in the same traceroute operation. H flag is only significant in echo request and MUST NOT be set in echo reply. The detailed processing based on the value of H flag is explained in [section 7.2](#).

```
| 0| 1| 2| 3| 4| 5| 6| 7|
+---+---+---+---+---+---+---+
|      MBZ      | H|
+---+---+---+---+---+---+---
```

- o TimeStamp Sent: time-of-day (3 octets for seconds and 3 octets for microseconds) in NTP format that the echo request was sent according to the sender's clock.

- o TimeStamp Received: time-of-day (3 octets for seconds and 3 octets for microseconds) in NTP format that the corresponding echo request was received according to the receiver's clock. This value is significant only in echo reply and MUST be set to all zeros in echo request and ignored on receipt of an echo request.

- o TLVs: A set of type, length, value encoded fields as specified in next section.

6. TLV Encodings

6.1. Target RBridge

```
| 0| 1| 2| 3| 4| 5| 6| 7| 8| 9|10|11|12|13|14|15|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Type = 0x05      | Length = 2 + 2*n      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Number of Target RBridges      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
.      Target RBridge Nickname 1      .
.      ...                             .
.      Target RBridge Nickname n      .
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

- o Number of Target RBridges: The number of nicknames specified in the following fields, the maximum number is 127.

- o Target RBridge Nickname: The nickname of a Target RBridge.

This TLV MAY appear in an echo request. It SHOULD be copied back in the corresponding echo reply messages.

Target RBridge TLV is used by multi-destination OAM. For ping along the multi-destination path, the Target RBridge TLV with multiple nicknames MAY be included in echo request. It implies RBridges with any of the nicknames in the TLV should reply. While for traceroute like application, only a single nickname can be included in this TLV. If there was more than one nickname in the TLV, only the first nickname MUST be used as target nickname for tracing purpose.

When Target RBridge TLV is not included in an echo request, it implies the unspecified target. If an echo request with unspecified target was sent by ping like applications, then all leaf nodes in distribution tree pruned by the given inner VLAN SHOULD send back echo reply. If echo request with unspecified target was sent by traceroute like application, RBridges receiving the incoming frame with hop count value 1 would process the echo request and send back 'Hop Count is Zero' error notification.

6.2. Jitter

```
| 0| 1| 2| 3| 4| 5| 6| 7| 8| 9|10|11|12|13|14|15|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Type = 0x07      |      Length = 0x02      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                        Jitter time                |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---
```

o Jitter time: Set to the upper bound of the jitter period in milliseconds. A responding node SHOULD wait a random amount of time between zero milliseconds and the value specified.

This TLV MAY appear in an Echo Request in the Long format. It SHOULD NOT be present in echo reply messages.

7. Processing Echo Messages for Multi-destination Path

7.1. Sending an echo request

The inner frame header and TRILL header fields are as follows:

- o Inner.MacSA: MAC address of RBridge originating the echo request
- o Inner.MacDA: Defaults to All-Egress-RBridges. It can be set to L2 multicast address derived from IP multicast group.
- o Inner.VLAN ID: Defaults to 1. It can be any enabled VLAN ID on the ingress RBridge.
- o Ingress RBridge Nickname: the nickname of RBridge originating the echo request
- o Egress RBridge Nickname: the nickname of a distribution tree root
- o M bit: 1
- o Hop Count: defaults to maximum value 0x3F.

- For ping like applications, it can be any value which is believed to be no less than the number of hops from ingress RBridge to the most distant target RBridge in the tree.

- For traceroute like applications, hop count value starts from 1

and is increased by one for each sending of echo request.

H(Respond Only When Hop Count is Zero) flag in echo request is set to 1 for traceroute like applications and 0 for ping like applications.

The originating RBridge chooses the values of Sender's Instance and Sequence Number for the echo request. Sequence number should be increased by 1 for each new subsequent echo request of the same Sender's Instance. The Timestamp Sent is set to the time-of-day in NTP format [[NTP](#)] according to the sender's clock. The Timestamp Received is set to zero.

The originating RBridge MAY use Target RBridge TLV to specify the target. For ping like applications, multiple nicknames MAY be present in one such TLV if sender wants to ping multiple targets at one time. For traceroute like applications, the TLV should at most contain one nickname as the tracing target. If there is more than one nickname, only the first one takes effect.

Echo request without Target RBridge TLV means the originating RBridge potentially wants to target every RBridge in the distribution tree. We also call it echo request with the unspecified target. For ping like applications, echo request with the unspecified target implies the sender wants to know who are the leaf nodes of the inner VLAN in the distribution tree. For traceroute like applications, it implies the sender wants to know the whole distribution tree structure hop-by-hop.

The Originating RBridge MAY include the Jitter TLV (see [section 6.2](#)) in the echo request in order to randomize the delay of the replying echo message from multiple RBridges.

[7.2](#). Receiving an echo request

RBridge receiving an echo request with M bit set with EtherType of RBridge channel [[RBridgeChannel](#)] SHOULD replicate it to the control plane for processing and also forward it as normal multi-destination data frame. When a RBridge receives an incoming frame with hop count is 1 in TRILL header, it will not forward the frame further. If reply

mode is 1, no echo reply is generated. For the sub sections below, we assume the reply mode is set to 2.

7.2.1. If H (Respond Only When Hop Count is Zero) Flag Is Not Set

- If Target RBridge TLV is not present in the echo request:

All leaf nodes of the distribution tree in the inner VLAN MUST process the incoming echo request and send back echo reply.

- If Target RBridge TLV is present in the echo request:

An RBridge owning any one of the specified target nicknames in the incoming echo request MUST send back echo reply when it is a leaf node of the distribution tree in the inner VLAN.

If echo reply has already been generated for the incoming echo request, RBridge will not generate 'Hop Count is Zero' error notification even when the hop count value in the incoming echo request is one.

Echo request with 'H' flag unset is for ping like application. It should be noted that if an RBridge receives an echo request with its own nickname listed as one of the targets, it does not send back the echo reply if the RBridge did not advertise its interest of inner

VLAN. That is to say, the connectivity check using ping in multi-destination path is constrained by inner VLAN. Normally VLAN 1 is the default VLAN and enabled on every RBridge. Therefore it is recommended to put inner VLAN to be 1 when we want to check the connectivity without the constraint of a particular customer VLAN. We may use the echo replies from that to plot the whole distribution tree.

7.2.2. If H (Respond Only When Hop Count is Zero) Flag Is Set

When hop count of the incoming echo request is not one, RBridge would never generate any echo reply or 'hop count is zero' error notification.

If the hop count is one in the incoming echo request:

- If Target RBridge TLV is not present in the echo request:

RBridge receiving the incoming frame with hop count equal to 1 MUST send back error notification of 'Hop Count is Zero'. RBridges MUST not generate any echo reply in this case. If hop count in incoming echo request is more than 1, control plane will not do anything. RBridge forwards the frame as normal multi-destination TRILL frame in data plane.

- If Target RBridge TLV is present in the echo request:

RBridges owning the only target nickname listed in TLV MUST send back echo reply if it is a leaf node of the inner VLAN in the distribution tree. If it is not a leaf node of the inner vlan, no echo reply will be generated by the owner RBridge; however, 'Hop Count is Zero' error notification will be sent back instead.

If an RBridge not owning the only target nickname listed in TLV receives the incoming frame with hop count equal to 1, it SHOULD check its LSDB. If it sits in-between of the ingress RBridge and the target RBridge along the specified distribution tree, RBridge MUST send back the error notification of 'Hop Count is Zero'; otherwise the RBridge should not generate such error notification. The purpose of suppressing the error notification here is to make sure the ingress only receives the error notification along the real data path and to reduce the processing burden at ingress.

If RBridge not owning the first target nickname listed in TLV receives the incoming frame with hop count greater than 1, the frame is forwarded as usual.

Echo request with 'H' flag set is for traceroute like applications. For traceroute with unspecified target, the ingress RBridge will be able to construct the whole distribution tree (when tree is not pruned) or the distribution tree of inner vlan (when tree is pruned by inner VLAN) according to the returned error notifications. For traceroute with a specified target in an inner VLAN, the ingress RBridge will receive the error notifications from the RBridges along the path to the target in the tree. If the target announced its interest of the inner VLAN, it will finally send back echo reply to the ingress. If the target did not announce its interest of the inner VLAN, either the target will not receive the echo request (e.g. it is located in the tree path being pruned) or the target will send back

error notification of 'Hop Count is Zero' instead of echo reply.

[7.3.](#) Sending an echo reply

The inner frame header and TRILL header fields are as follows,

- o Inner.MacSA: The MAC address of the RBridge generating the echo reply
- o Inner.MacDA: All-Egress-RBridges
- o Inner.VLAN ID: same as Inner.VLAN ID in the received echo request to which the echo reply responds
- o Ingress RB Nickname: the nickname of the RBridge generating the echo reply.
- o Egress RBridge Nickname: the ingress RBridge nickname in the corresponding received echo request
- o M bit: 0
- o Hop Count: defaults to the maximum value 0x3F. It can be any value that is believed to be larger than the number of hops from ingress to egress RBridge.

The values of Sender's Instance, Sequence Number and Timestamp sent in an echo reply MUST be same as those in its corresponding echo request. H flag MUST be zero in echo reply. The value of Timestamp Received is set to the time-of-day in NTP format [[NTP](#)] according to the receiver's clock.

If Target RBridge TLV was present in the echo request, the corresponding echo reply SHOULD copy it.

Next Hop Nickname and Incoming Port ID TLV [[RBridgeOAM](#)] MAY be included in echo reply.

When an echo reply is going to be sent to the originator RBridge, 'Hop Count is Zero' error notification MUST not be sent in response to the same echo request.

[7.4.](#) Receiving an echo reply

An RBridge SHOULD use the Sender's Instance and Sequence Number to match up the received echo reply with the echo request it sent. If there is no match found, the RBridge should discard the echo reply.

If Jitter TLV was present in the echo request, the round trip time should not be calculated based on the difference between the arriving time of echo reply and the value of "TimeStamp sent" in the replying frame. However the single trip time is always correct to be calculated on Timestamp Received minus Timestamp Sent when the clocks of sender and receiver are synchronized.

When an RBridge receives either an echo reply or 'hop count is zero' error notification from the target RBridge for traceroute like application, it SHOULD stop sending echo request with increased hop count value.

[8.](#) Security Considerations

The security vulnerabilities raised in [\[RBridgeOAM\]](#) also apply to the multi-destination RBridge ping in this document. The same mechanisms can be used to prevent or alleviate the security issues.

[9.](#) IANA Considerations

New error notification sub-code needs to be allocated by IANA as specified in [Section 7](#).

[10.](#) References

[10.1.](#) Normative References

[RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", [RFC 6325](#), July 2011.

[RBridgeChannel] Eastlake, D., Manral, V., Yizhou, L., Aldrin, S., and D. Ward, "RBridges: TRILL RBridge Channel Support", [draft-ietf-trill-rbridge-channel-02](#) (work in progress), July 2011.

[RBridgeOAM] D. Bond, and V. Manral, "RBridges: Operations,

Administration, and Maintenance (OAM) Support", [draft-ietf-trill-rbridge-oam-01](#) (work in progress), October 2011.

[NTP] Mills, D., "Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI", [RFC 2030](#), October 1996.

[10.2](#). Informative References

[RFC6165] Banerjee, A. and D. Ward, "Extensions to IS-IS for Layer-2 Systems", [RFC 6165](#), April 2011.

[RFC6326] Eastlake, D., Banerjee, A., Dutt, D., Perlman, R., and A. Ghanwani, "TRILL Use of IS-IS", [RFC 6326](#), July 2011.

[11](#). Acknowledgments

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

Authors' Addresses

Yizhou Li
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Phone: +86-25-56624558
Email: liyizhou@huawei.com

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Phone: +86-25-56623144
Email: haoweiguo@huawei.com

International Business Machines
2051 Mission College Blvd.
Santa Clara, CA 95054
US

Phone: +1-603-339-7575
EMail: mokon@mokon.net
URI: <http://mokon.net>

Vishwas Manral
Hewlett Packard Co.
19111 Pruneridge Ave,
Cupertino, CA 95014 USA

Phone: +1-408-447-1497
EMail: vishwas.manral@hp.com