TRILL Working Group Internet-Draft Intended status: Standards Track Expires: February 27, 2013 H. Zhai ZTE T. Senevirathne Cisco Systems R. Perlman Intel Labs D. Eastlake 3rd M. Zhang Huawei F. Hu ZTE August 26, 2012

# RBridge: Pseudo-Nickname draft-hu-trill-pseudonode-nickname-03

#### Abstract

At the edg of TRILL campus, some RBridges provide end-station services to their attached end stations; these RBridges are called edge RBridges. To avoid potential frame duplication or loops in TRILL campus, only one edge RBridge is permitted to provide such services in a VLAN at all times to its attached end stations even they are also attached to other edge RBridges. However, in some application scenarios, for example in Link Aggregation Group (LAG), more than one edge RBridge is required to provide such services to an end station even in a VLAN to improve resiliency and maximize the available network bandwidth, which causes the flip-flopping of the egress RBridge nickname for such an end station in remote RBridges' forwarding tables. In this document, the concept of Virtual RBridge, along with its pseudo-nickname, is introduced to address the above problem.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <u>http://datatracker.ietf.org/drafts/current/</u>.

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."

Zhai, et al.

Expires February 27, 2013

[Page 1]

This Internet-Draft will expire on February 27, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents (<u>http://trustee.ietf.org/license-info</u>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

# Table of Contents

$\underline{1}$ . Introduction	· <u>4</u>
<u>1.1</u> . Terminology and Acronyms	. <u>5</u>
2. Problem Statement	
2.1. Appointed Forwarders on Shared Links	. <u>5</u>
2.2. Multi-homing and Link Aggregation to TRILL Network	. <u>6</u>
3. Concept of Virtual RBridge and Pseudo-nickname	· <u>7</u>
3.1. VLAN-x Appointed Forwarder for member interfaces in RBv	. 8
<u>3.2</u> . Announcing Pseudo-Nickname of RBv	. <u>8</u>
$\underline{4}$ . Distribution Trees for Member RBridges in RBv $\ldots$ $\ldots$	. <u>8</u>
5. Frame Processing	. <u>10</u>
<u>5.1</u> . Data Frames	
<u>5.1.1</u> . Native Frames Ingressing	. <u>10</u>
<u>5.1.2</u> . TRILL Data Frames Egressing	. <u>10</u>
<u>5.1.2.1</u> . Unicast TRILL Data Frames	. <u>10</u>
<u>5.1.2.2</u> . Multi-Destination TRILL Data Frames	. <u>11</u>
$\underline{6}$ . Member Link Failure in RBv	
<u>7</u> . OAM Frames	. <u>12</u>
8. Configuration Consistency	. <u>13</u>
9. IANA Considerations	. <u>13</u>
<u>10</u> . Security Considerations	. <u>13</u>
<u>11</u> . Acknowledgements	. <u>13</u>
<u>12</u> . Normative References	. <u>13</u>
Appendix A. Reasons for MAC Sharing among Member RBriges	
Authors' Addresses	. <u>15</u>

## **1**. Introduction

The IETF TRILL protocol [RFC6325] provides optimal pair-wise data frame forwarding without configuration, safe forwarding even during periods of temporary loops, and support for multi-pathing of both unicast and multicast traffic. TRILL accomplishes this by using [IS-IS] [RFC1195] link state routing and encapsulating traffic using a header that includes a hop count. The design supports VLANs and optimization of the distribution of multi-destination frames based on VLANs and IP derived multicast groups. Devices that implement TRILL are called RBridges.

In TRILL protocol, RBridges are identified by nicknames (16-bits). Different RBridge has different nickname(s). At the edge of TRILL network, some RBridges connect to legacy networks on one side and to TRILL network on the other side. These RBridges are called edge RBridges. For the connectivity between the two types of network, edge RBridges provide end-station services to end stations located in legacy networks. When receiving a native frame from such a local end station S, the service edge RBridge RB1 encapsulates the frame in a TRILL header, addressing the packet to RBridge RBx to which the destination end station D is attached. The TRILL header contains an "ingress RBridge nickname" field (RB1's nickname), an "egress RBridge nickname" field (RBx's nickname), and a hop count. On receiving such a frame, RBx removes the TRILL header and forwards it onto D in its native form. Meanwhile, based on the de-capsulation of such a frame, RBx learns the (ingress RBridge nickname, source MAC address, VLAN ID) triplet. Edge RBridges maintain such triplets in their forwarding table for the potential following transmission of native frames.

Due to failures, reconfiguration and other network dynamics, service edge RBridge for S may change over from RB1 to another edge RBridge. In this event, remote traffic addressed to S will be still forwarded to RB1 by remote RBridge RBx before perceiving this change, and then the traffic gets dropped at RB1, causing traffic disruption. Furthermore, to improve resiliency and maximize the available network bandwidth, an end station typically is multi-homed to several edge RBridges and treats all the uplink links as a Link Aggregation Group (LAG) bundle. In this scenario, all those edge RBridges work in an active-active load sharing model to provide end-station services for an end station even in same VLAN. When remote RBridge RB2 receives different frames, which are originated by such an end station S and ingressed into TRILL campus by different such edge RBridge, flipflopping of ingress RBridge nickname for MAC of S will be observed by RBx during de-capsulating such frames. This flip-flopping will cause disorder of different frames in traffic, worsening the traffic disruption.

In this document, concept of Virtual RBridge group, together with its Pseudo-nickname, is introduced to address the above issues. For a member RBridge in such a group, it uses the pseudo-nickname of this group, instead of its own device nickname, as ingress RBridge nickname when encapsulating a frame to its TRILL form with a TRILL header. So, in a RBridge Group, even if there are more than one RBridge providing end-station services for a end station or the service RBridge changes over from one member RBridge to another in same set of VLANs, the ingress RBridge nickname for the MAC of this end station will still remain unchanged in remote RBridges' forwarding tables.

This document is organized as following: <u>Section 2</u> is problem statement, which describes why virtual RBridge and its pseudonickname are required. <u>Section 3</u> gives the concept of virtual RBridge. <u>Section 4</u> describes the consideration for pseudo-nickname used in ingressing multi-destination frames. <u>Section 5</u> covers processing of transit frame traffic when considering pseudo-nickname.

Familiarity with [<u>RFC6325</u>] is assumed in this document.

#### **<u>1.1</u>**. Terminology and Acronyms

This document uses the acronyms defined in [<u>RFC6325</u>] and the following additional acronym:

AF - Appointed Forwarder

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 [<u>RFC2119</u>].

When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in [RFC2119].

#### 2. Problem Statement

#### 2.1. Appointed Forwarders on Shared Links

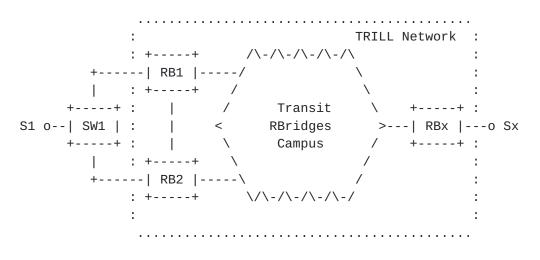
If there are multiple RBridges on a shared link, together with end stations, only one RBridge is permitted to provide end-station services in a VLAN at all times for the end stations to avoid possible frame duplication or loops in TRILL campus. The service RBridge is called VLAN-x Appointed Forwarder (AF) on a shared link.

However, AF for any set of VLANs on a shared link may change over

from one RBridge to another, due to failure, configurations and other network dynamics, etc. If such change occurs, local end stations may not perceive it, so the end station cannot timely notify remote RBridges to update the correspondence between ingress RBridge nickname and the MAC of this end station in their forwarding tables. As a result, remote RBridges may continue to forward traffic to the previous AF and the traffic may be dropped at the previous egress RBridge, causing traffic disruption.

### **<u>2.2</u>**. Multi-homing and Link Aggregation to TRILL Network

In order to improve the reliability of connection to TRILL network, multi-homing technique may be employed by a legacy device, such as a switch or end host. For example, in Figure 1, switch SW1 multi-homes to TRILL network by connecting to both RBridge RB1 and RB2 with respective links. Then the end stations (e.g., S1), attached to SW1, still get end-station services from TRILL network even if one connection of SW1 to TRILL network, e.g., SW1-RB1, fails. That is to say, the service RBridge for S1 changes over from RB1 to RB2 after the connection of SW1-RB1 fails, which causes traffic disruption temporarily (e.g., traffic from Sx to S1), similar to AF changing in <u>Section 2.1</u>.



# Figure 1 Multi-homing to TRILL Network

Furthermore, SW1 may treat the two links as a LAG (Link Aggregation Group) bundle, so that the two links form active-active load sharing model instead of previous active-stanby model. That is to say, in Figure 1, two service RBridges (e.g., RB1 and RB2) must provide end-station services simultaneously to S1 in that VLAN. And this will result in flip-flopping of the ingress RBridge for the MAC of S1 in remote RBridges' (e.g., RBx) forwarding tables. AS a result, this

flip-flopping will cause much more disorder packets and worsen the

Zhai, et al. Expires February 27, 2013 [Page 6]

traffic disruption.

Besides switches, end stations can also directly multi-home to TRILL network and treat the multi-homing links as a LAG bundle. The issue of traffic disruption also occurs in this scenario if such an end station balances different traffic load in a same VLAN among the member links.

In the following sections, concept of RBridge Group, together with its nickname, is introduced to fix these issues.

### 3. Concept of Virtual RBridge and Pseudo-nickname

A Virtual RBridge (RBv) represents a group of different ports on different edge RBridges, on which these RBridges provide end-station service to a set of their attached end stations. After joining RBv, such an RBridge port is called a member port of RBv, and such an RBridge becomes a member RBridge of RBv. In an RBridge RBv is identified by its virtual nickname in TRILL campus, and virtual nickname is also referred to as pseudo-nickname in this specification.

After joining an RBv, a member RBridge will announce its connection to RBv by including the information of that RBv, e.g., the pseudonickname of RBv, in its self-originated LSP. From such LSPs, remote RBridges that are not a members of RBv can deduce that, one or more shortest paths are available from itself to RBv.

When receiving a native frame on such a port, the member RBridge uses the RBv's nickname, instead of its own nickname, as ingress nickname in TRILL header if necessary to encapsulate the frame into TRILL data form. By de-capsulating such a TRILL-encapsulated data frame, a remote RBridge learns that S is reachable through RBv.

NOTE: An RBridge port can join at most one RBv at any time, but different ports on the same RBridge can join the same RBv or different RBvs. After joining an RBv, such a port becomes a member port of the RBv, and the RBridge becomes a member RBridge of the RBv. Furthermore, for a member RBridge, it MUST move out of RBv and clear the RBv's information from its self-originated LSPs when it loses the last member port from this group, due to port down, configuration, etc.

Use of the Appointed Forwarder framework specified in [<u>RFC6325</u>], this specification allows to utilize a single framework for both shared LAN and point-point edge connectivity. Additionally this allows to

- o Detect and protect against mis-configuration at the edge, e.g., on the device SW1 the two interfaces are not configured as LAG or
- Accept TRILL and native frames on the RBridge interface connecting S1 in above Figure 1.
- o Avoid loops in the event S1 and S2 were connected by a native Ethernet Link.

#### 3.1. VLAN-x Appointed Forwarder for member interfaces in RBv

If member RBridges in RBv cannot see each others' Hellos on their member ports (e.g., in the LAG scenario), then each RBridge becomes Designated RBridge (DRB) for that port and appoints itself as AF for all VLANs as it does not see any TRILL hellos on that port. However, it MAY acts as appointed forwarders only for parts of VLANs on that port, if it knows explicitly the sets of service VLANs on that port via other means. For example, administrator can statically configured the sets of service VLANs on that port, or a lower protocol (e.g., LAG protocol) informs TRILL the sets of service VLANs on that port, etc.

However, if they can see each others' Hellos on the member ports in RBv (e.g., in the shared link scenario), the TRILL Hello protocol in [<u>RFC6325</u>] is used for DRB election and for VLAN-x AFs appointment on those ports. Then the DRB appoints different member ports as AFs for different sets of VLANs.

Among the member RBridges of RBv, only the VLAN-x forwarder is responsible to ingress native traffic (both unicast and non-unicast traffic) in this VLAN into TRILL campus, but non-forwarder member RBridge is also permitted to egress unicast TRILL data traffic out of TRILL campus. For the multi-destination TRILL data frames, only the VLAN-x forwarder can egress their out of TRILL campus.

#### 3.2. Announcing Pseudo-Nickname of RBv

Each member RBridge advertises the RBv's pseudo-nickname using the nickname sub-TLV [rfc6326bis], along with its regular nickname or nicknames, in its LSPs. When a member RBridge leaves from RBv due to losing its last member ports in RBv, it MUST clear RBv's pseudo-nickname from its update LSPs.

#### **4**. Distribution Trees for Member RBridges in RBv

In TRILL, RBridges use distribution trees to forward multidestination frames. When a native frame either to destinations whose

location is unknown or to multicast/broadcast groups is necessary to be ingressed into TRILL campus, the ingress RBridge encapsulates it into multi-destination TRILL data frame and forwards it along a chosen distribution tree. In the TRILL header of the TRILL frame, the ingress nickname identifies the ingress RBridge and the egress nickname represents the root of the chosen tree. After receiving a multi-destination TRILL data frame, the RBridge performs Reverse Path Forwarding (RPF) check, along with other checks, on the multidestination frame to further control potentially looping traffic.

RPF specifies that a multi-destination TRILL data frame ingressed by an RBridge and forwarded along a distribution tree can only be received by an RBridge on an expected port. If not on that port, that frame MUST be dropped by that RBridge.

However, member RBridges employ RBv's pseudo-nickname other than their own nicknames as ingress nickname when they ingress native frames received on member ports, regardless unicast or non-unicast frames, into TRILL campus. Therefore, when these frames reach a remote RBridge, they will be treated, by that RBridge, as frames ingressed by the same RBridge, i.e., RBv. If they are multidestination frames and the same distribution tree is chosen by different member RBridges to forward these frames, they will travel along the tree and reach a remote RBridge on different ports. Then the RPF check is violated, and some of the frames reaching the RBridge on unexpected ports are dropped by the RBridge.

To fix the above issue, a scalable and resilient approach is proposed in [CMT], where different member RBridges are assigned different distribution trees for forwarding the multi-destination TRILL data frames that using RBv's pseudo-nickname as ingress nickname in their TRILL header. And a new TLV, named Affinity sub-TLV, is also introduced for a member RBridge to announce its assigned distribution tree for RBv in its self-originated LSPs. After receiving such LSPs, remote RBridges can calculate their RPF check information for RBv on those specified trees.

In this specification, the approach proposed in [CMT] is employed for RBv to assign different distribution trees to different member RBridges and the Affinity sub-TLV for member RBridges to announce their assigned trees in LSPs.

When a member RBridge joins into or leaves from a virtual RBridge group RBv due to its last member ports up/down or its configuration changing, etc., the distribution trees assigned to different member RBridges may change. That change and its influence on frame processing are beyond the scope of this document.

### 5. Frame Processing

Although, there are five types of Layer 2 frames in [<u>RFC6325</u>], e.g., native frame, TRILL data frame, TRILL control frames, etc., pseudonickname of RBv is only used for native frame and TRILL data frame in this specification.

#### 5.1. Data Frames

### 5.1.1. Native Frames Ingressing

When RB1 receives a native frame on one of its valid member ports of RBv, it uses the pseudo-nickname of RBv, instead of its own nickname, as ingress nickname, if it is the appointed forwarder for the VLAN of the frame on the port. If the frame is not received on a member port, RB1 MUST NOT use RBv's pseudo-nickname as ingress nickname when doing TRILL-encapsulation on the frame. Otherwise, the reverse traffic may be forwarded to another member RBridge that does not connect to the link containing the destination, which may cause the traffic disruption.

If the above native frame is ingressed by RB1 as a multi-destination TRILL data frame, e.g., its destination is unknown to RB1 or it is non-unicast frame, RB1 can only choose one of its assigned distribution trees for RBv to distribute the TRILL-encapsulated frame [CMT]. If not so, the multi-destination TRILL data frame will fail RPF check on another RBridge and be dropped.

Furthermore, for such a frame, its source MAC address information ( {
VLAN, Outer.MacSA, port } ) is learned by default if its source
address is unicast. Then the learned information is shared with
other member RBridges of RBv (See <u>Appendix A</u> for more details for the
information sharing).

# 5.1.2. TRILL Data Frames Egressing

This section describes egress processing of the received TRILL data frames on a member RBridge(RBn, say) in virtual RBridge group of RBv. <u>Section 5.1.2.1</u> describes unicast TRILL data frames egress processing. <u>Section 5.1.2.2</u> covers multi-destination TRILL data frames egress processing.

# 5.1.2.1. Unicast TRILL Data Frames

When receiving a unicast TRILL data frame, RBn checks the egress nickname in the TRILL header of the frame. If the egress nickname is one of RBn's own nicknames, the frame is processed as <u>Section 4.6.2.4</u> in [RFC6325].

If the egress nickname is RBv's pseudo-nickname and RBn is a valid member RBridge of RBv, the Inner.MacSA and Inner.VLAN ID are, by default, learned associated with the ingress nickname, unless that nickname is unknown or reserved or the Inner.MacSA is not unicast. If the learned {Inner.MacSA, Inner.VLAN ID, ingress nickname} triplet is a new one or updates the locally stored one, this triplet is shared with other member RBridges of RBv (See <u>Appendix A</u> for more details for the triplet sharing).

Then the frame being forwarded is de-capsulated to native form. The Inner.MacDA and Inner.VLAN ID are looked up in RBn's local forwarding address cache, and one of the three following cases occurs:

- o If the destination end station identified by the Inner.MacDA and Inner.VLAN ID is on a local link to RBv, this frame is sent onto the link containing the destination.
- o else if RBn can reach the destination through another RBridge RBk, it re-encapsulate the native frame into a unicast TRILL data frame and sends it to RBk. RBn uses RBk's own nickname, instead of RBv's pseudo-nickname as egress nickname for the re-encapsulation, and remains the ingress nickname unchanged.
- o Else, RBn does not know how to reach the destination; it sends the native frame out of all its member ports of RBv on which it is appointed forwarders for the Inner.VLAN.

### 5.1.2.2. Multi-Destination TRILL Data Frames

If the RBn is an appointed forwarder for the Inner.VLAN ID of the frame, the Inner.MacSA and Inner.VLAN ID are, by default, learned as associated with the ingress nickname unless that nickname is unknown or reserved or the Inner.MacSA is not unicast. If the learned {Inner.MacSA, Inner.VLAN ID, ingress nickname} triplet is a new one or updates the locally stored one, this triplet is shared among the members RBridges in virtual RBridge group RBv (See <u>Appendix A</u> for more details for the triplet sharing).

Then a copy of the frame is de-capsulated into its native form. Before the native frame is sent out of ports on which RBn is appointed forwarder for the frame's VLAN, the following extra check is performed:

Assigned Distribution Trees Check (ADTC): If the flag for this check (ADTC\_flag) is not zero on such a port, the distribution tree T along which the TRILL data frame arrives at RBn is checked. Only if T is one of RBn's assigned distribution trees in RBv, the native frame can be send out of this port. If not, the frame

cannot be sent out of this port.

The value of ADTC\_flag on a RBridge's end-station-servicing port depends on whether the port is a member port of RBv and RBn can not receive Hellos from other member RBridges on that port or not.

If a port is a member port of RBv and RBn is the appointed forwarder for all VLANs on that port, the ADTC\_flag MUST be set 1. For all other cases ADTC\_flag MUST be set to zero.

#### 6. Member Link Failure in RBv

In Figure 1, if the link SW1-RB1 fails, RB1 loses its only local link to S1. When that failure is detected, the MAC entries through the failed link to S1 are removed from RB1's forwarding table immediately, and other MAC entries to S1 shared by other member RBridges of RBv (See <u>Appendix A</u> for more details), e.g., RB2, are installed into RB1's forwarding table when RB1 still has at least one valid member port in RBv. Then when the TRILL-encapsulated traffic to S1 is delivered to RB1, it can be re-encapsulated by RB1 and forwarded, based on the available MAC entries, to another member RBridge which has direct link to S1 and egresses the traffic to S1.

On the other hand, if RB1 has lost all its member ports of RBv, it MUST updates its self-orignated LSPs to announce its giving up of membership of RBv and no longer utilizes pseudo-nickname of RBv to ingress/egress traffic into/out of TRILL campus until one of its member ports of RBv becomes valid.

NOTE: Although on an edge RBridge different ports that connect to different LAGs or LANs can join the same RBv, for simplicity, it is RECOMMENDED that on an edge RBridge different ports connecting to different LAGs or LANs join different RBvs in practical deployment, each RBv per LAG or per LAN. Then for such an edge RBridge, when all its member ports connecting to a LAG or LAN failed, it can move out of this RBv and no longer uses the RBv's pseudo-nickname to ingress/ egress data traffic into/out of TRILL campus.

#### 7. OAM Frames

Special attention must be paid when generate the OAM frames. When an OAM frame is generated with ingress nickname of RBv, originator RBridge's nickname MUST be included in the OAM message to ensure response is returned to the originating member of RBv group.

### **<u>8</u>**. Configuration Consistency

It is important that VLAN membership of member ports of end switch SW1 is consistent across all of the member ports it is attaching to member RBridges of RBv in the point-piont scenario. Any inconsistencies in VLAN membership may result in packet loss or having to through an extra hop RBridge before the packet reaches its destination end station.

As an example consider, in Figure 1, on RB1 link SW1-RB1 has VLAN1 and VLAN2 configured. Consider only VLAN1 is configured on RB2 on SW1-RB2 link. Both RB1 and RB2 use the same ingress nickname RBv for all frames originating from S1. Hence, RBx will learn MAC address from S1 on VLAN2 as originating from RBv. As a result, on the return path RBx may deliver VLAN2 traffic to RB2. RB2, does not have VLAN2 configured on SW1-RB2 link and hence may drop the frame or has to forward the frame to RB1 to egress it to S1 if RB2 knows RB1 can reach S1 in VLAN2.

# 9. IANA Considerations

TBD.

### **<u>10</u>**. Security Considerations

TBD.

### 11. Acknowledgements

We would like to thank Mingjiang Chen for his contributions to this document. Additionally, we would like to thank Erik Nordmark, Les Ginsberg, Ayan Banerjee, Dinesh Dutt, Anoop Ghanwani, Janardhanan Pathang, and Jon Hudson for their good questions and comments.

### **<u>12</u>**. Normative References

- [CMT] Senevirathne, T., Pathangi, J., and J. Hudson, "Coordinated Multicast Trees (CMT)for TRILL", <u>draft-ietf-trill-cmt-00.txt</u> Work in Progress, April 2012.
- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", <u>RFC 1195</u>, December 1990.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate

Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, March 1997.

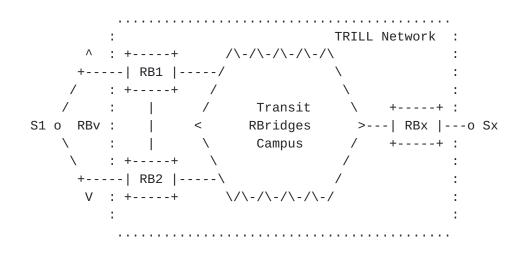
[RFC6325] Perlman, R., Eastlake, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", <u>RFC 6325</u>, July 2011.

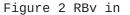
#### [rfc6326bis]

Eastlake 3rd, D., Banerjee, A., Ghanwani, A., and R. Perlman, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", <u>draft-eastlake-isis-rfc6326bis-07.txt</u> Work in Progress, March 2012.

# Appendix A. Reasons for MAC Sharing among Member RBriges

With the introduction of virtual RBridge, MAC flip-flopping problem in LAN or LAG is resolved. However, in order to forward traffic effectively, member RBridges should share some of their learned MAC addresses with each other. For example, see Figure 2 shown below.





#### LAG scenario

In VLAN-x, native frames from S1 to Sx will enter TRILL campus through one member RBridge of the RBv, such as RB1 in Figure 2, so RB1 learns the location of S1 in VLAN-x; but with regard to reverse traffic, RBx may deliver it to RB2 if it thinks the shortest path to RBv is through RB2. Then, if RB2 knows the location of S1 and the link RB2-S1 is good, it egresses the traffic directly to S1. However, if the link fails and RB2 has not learn the location of S1 in VLAN-x, RB2 cannot transmit the traffic properly to S1.

Internet-Draft

Pseudo-Nickname

Thus, the MAC addresses of attached end stations on one member RBridge SHOULD be shared with the rest member RBridges in an RBv. With these informations shared, when RB2 receives reverse frames, it can determine how to forward them to S1, for example, forward them to RB1 if the link RB2-S1 fails.

On the other hand, RBx always delivers the reverse traffic to RB2 if it thinks the shortest path to RBv is through RB2. Then RB2 egresses the traffic and learns the location of Sx in the case its link to S1 is good. RB1 will not know where Sx is if neighbor it has other ways to get the location of Sx nor RB2 shares this information with it. As a result, it has to always treat the traffic from S1 to Sx as unknown destination traffic and multicast it in TRILL. Always multicasting such traffic adds additional forwarding burden on TRILL network.

Therefore, in addition to local attached end station MAC addresses, the learned remote MAC addresses should also be shared among all other member RBridges in an RBv. With such information sharing, RB1 can treat the traffic to Sx as known destination traffic and unicast it to RBx.

Although we can extend ESADI (End Stations Address Distribution Information) protocol or LAG protocol, etc., for such MAC sharing, ways for the sharing are beyond the scope of this document.

Authors' Addresses

Hongjun Zhai ZTE 68 Zijinghua Road, Yuhuatai District Nanjing, Jiangsu 210012 China

Phone: +86 25 52877345 Email: zhai.hongjun@zte.com.cn

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

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

Internet-Draft

Radia Perlman Intel Labs 2200 Mission College Blvd Santa Clara, CA 95054-1549 USA Phone: +1-408-765-8080 Email: Radia@alum.mit.edu Donald Eastlake 3rd Huawei 155 Beaver Street Milford, MA 01757 USA Phone: +1-508-333-2270 Email: d3e3e3@gmail.com Mingui Zhang Huawei Huawei Building, No.156 Beiging Rd. Beijing, Beijing 100095 China Email: zhangmingui@huawei.com Fangwei Hu ZTE 889 Bibo Road, Pudong District Shanghai, Shanghai 201203 China

Phone: +86 21 68896273 Email: hu.fangwei@zte.com.cn