

TRILL Working Group
Internet Draft
Intended status: Standard Track
Updates: [6325](#)

Tissa Senevirathne
Consultant
Janardhanan Pathangi
DELL
Jon Hudson
Brocade

October 6, 2015

Expires: April 2016

Coordinated Multicast Trees (CMT) for TRILL
draft-ietf-trill-cmt-09.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/1id-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 March 6, 2009.

Copyright Notice

Copyright (c) 2015 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

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.

Abstract

TRILL facilitates loop free connectivity to non-TRILL networks via choice of an Appointed Forwarder for a set of VLANs. Appointed Forwarders provide load sharing based on VLAN with an active-standby model. High performance applications require an active-active load-sharing model. The Active-Active load-sharing model can be accomplished by representing any given non-TRILL network with a single virtual RBridge. Virtual representation of the non-TRILL network with a single RBridge poses serious challenges in multi-destination RPF (Reverse Path Forwarding) check calculations. This document specifies required enhancements to build Coordinated Multicast Trees (CMT) within the TRILL campus to solve related RPF issues. CMT provides flexibility to RBridges in selecting desired path of association to a given TRILL multi-destination distribution tree. This document updates [RFC 6325](#).

Table of Contents

1.	Introduction.....	3
1.1.	Scope and Applicability.....	5
1.2.	Contributors.....	5
2.	Conventions used in this document.....	5
2.1.	Acronyms and Phrases.....	5
3.	The AFFINITY sub-TLV.....	6
4.	Multicast Tree Construction and Use of Affinity Sub-TLV.....	6
4.1.	Update to RFC 6325.....	8
4.2.	Announcing virtual RBridge nickname.....	9
4.3.	Affinity Sub-TLV Capability.....	9
5.	Theory of operation.....	10
5.1.	Distribution Tree Assignment.....	10
5.2.	Affinity Sub-TLV advertisement.....	10
5.3.	Affinity sub-TLV conflict resolution.....	11
5.4.	Ingress Multi-Destination Forwarding.....	11
5.4.1.	Forwarding when $n < k$.....	12
5.5.	Egress Multi-Destination Forwarding.....	12
5.5.1.	Traffic Arriving on an assigned Tree to RBk-RBv.....	12
5.5.2.	Traffic Arriving on other Trees.....	12
5.6.	Failure scenarios.....	13
5.6.1.	Edge RBridge RBk failure.....	13

5.7. Backward compatibility.....	14
6. Security Considerations.....	14
7. IANA Considerations.....	14
8. References.....	15
8.1. Normative References.....	15
8.2. Informative References.....	16
9. Acknowledgments.....	16
Appendix A. Change History.....	17

1. Introduction

TRILL (Transparent Interconnection of Lots of Links) presented in [RFC6325] and other related documents, provides methods of utilizing all available paths for active forwarding, with minimum configuration. TRILL utilizes IS-IS (Intermediate System to Intermediate System [IS-IS]) as its control plane and uses a TRILL header with hop count.

[RFC6325], [RFC7177], and [RFC6439] provide methods for interoperability between TRILL and Ethernet end stations and bridged networks. [RFC6439], provides an active-standby solution, where only one of the R Bridges on a link with end stations is in the active forwarding state for end station traffic for any given VLAN. That R Bridge is referred to as the Appointed Forwarder (AF). All frames ingressed into a TRILL network via the Appointed Forwarder are encapsulated with the TRILL header with a nickname held by the ingress AF R Bridge. Due to failures, re-configurations and other network dynamics, the Appointed Forwarder for any set of VLANs may change. R Bridges maintain forwarding tables that contain destination MAC address and Data Label (VLAN or Fine Grained Label (FGL)) to egress R Bridge binding. In the event of an AF change, forwarding tables of remote R Bridges may continue to forward traffic to the previous AF and that traffic may get discarded at the egress, causing traffic disruption.

High performance applications require resiliency during failover. The active-active forwarding model minimizes impact during failures and maximizes the available network bandwidth. A typical deployment scenario, depicted in Figure 1, may have either End Stations and/or bridges attached to the R Bridges. These devices typically are multi-homed to several R Bridges and treat all of the uplinks independently using a Local Active-Active Link Protocol (LAALP [RFC7379]) such as a single Multi-Chassis Link Aggregation (MC-LAG) bundle or Distributed Resilient Network Interconnect [8021AX]. The Appointed Forwarder designation presented in [RFC6439] requires each of the edge R Bridges to exchange TRILL Hello packets. By design, an LAALP does not forward packets received on one of the member ports

of the MC-LAG to other member ports of the same MC-LAG. As a result the AF designation methods presented in [\[RFC6439\]](#) cannot be applied to deployment scenario depicted in Figure 1. [\[RFC7379\]](#)

An active-active load-sharing model can be implemented by representing the edge of the network connected to a specific edge group of R Bridges by a single virtual R Bridge. Each virtual R Bridge MUST have a nickname unique within its TRILL campus. In addition to an active-active forwarding model, there may be other applications that may requires similar representations.

Sections [4.5.1](#) and [4.5.2](#) of [\[RFC6325\]](#) as updated by [\[RFC7180bis\]](#) specify distribution tree calculation and RPF (Reverse Path Forwarding) check calculation algorithms for multi-destination forwarding. These algorithms strictly depend on link cost and parent R Bridge priority. As a result, based on the network topology, it may be possible that a given edge R Bridge, if it is forwarding on behalf of the virtual R Bridge, may not have a candidate multicast tree that the edge R Bridge can forward traffic on because there is no tree for which the virtual R Bridge is a leaf node from the edge R Bridge.

In this document we present a method that allows R Bridges to specify the path of association for real or virtual child nodes to distribution trees. Remote R Bridges calculate their forwarding tables and derive the RPF for distribution trees based on the distribution tree association advertisements. In the absence of distribution tree association advertisements, remote R Bridges derive the SPF (Shortest Path First) based on the algorithm specified in [section 4.5.1 of \[RFC6325\]](#) as updated by [\[RFC7180bis\]](#). This document updates [\[RFC6325\]](#) by changing, when CMT sub-TLVs are present, [\[RFC6325\]](#) mandatory provisions as to how distribution tree are constructed.

Other applications, beside the above mentioned active-active forwarding model, may utilize the distribution tree association framework presented in this document to associate to distribution trees through a preferred path.

This proposal requires the presence of multiple multi-destination trees within the TRILL campus and updating all the R Bridges in the network to support the new Affinity sub-TLV ([Section 3](#)). It is expected that both of these requirements will be met as they are control plane changes, and will be common deployment scenarios. In case either of the above two conditions are not met, R Bridges MUST support a fallback option for interoperability. Since the fallback is expected to be a temporary phenomenon till all R Bridges are

upgraded, this proposal gives guidelines for such fallbacks, and does not mandate or specify any specific set of fallback options.

1.1. Scope and Applicability

This document specifies an Affinity sub-TLV to solve RPF issues at the active-active edge. Specific methods in this document for making use of the Affinity sub-TLV are applicable where a virtual RBridge is used to represent multiple RBridges connected to an edge CE through an LAALP such as multi-chassis link aggregation or some similar arrangement where the RBridges cannot see each other's Hellos.

This document DOES NOT provide other required operational elements to implement an active-active edge solution, such as methods of multi-chassis link aggregation. Solution specific operational elements are outside the scope of this document and will be covered in other documents. (See, for example [[TRILLPN](#)].)

Examples provided in this document are for illustration purposes only.

1.2. Contributors

The work in this document is a result of many passionate discussions and contributions from following individuals. Their names are listed in alphabetical order:

Ayan Banerjee, Dinesh Dutt, Donald Eastlake, Mingui Zhang, Radia Perlman, Sam Aldrin, Shivakumar Sundaram, and Zhai Hongjun.

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 [[RFC2119](#)].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying [[RFC2119](#)] significance.

2.1. Acronyms and Phrases

The following acronyms and phrases are used in this document:

AF: Appointed Forwarder [[RFC6439](#)].

CE: Customer Ethernet device, that is, a device that performs forwarding based on 802.1Q bridging. This also can be end-station or a server.

Data Label: VLAN or FGL.

FGL: Fine Grained Label [[RFC7172](#)].

LAALP: Local Active-Active Link Protocol [[RFC7379](#)].

MC-LAG: Multi-Chassis Link Aggregation is a proprietary extension to [[8021AX](#)], that facilitates connecting group of links from an originating device (A) to a group of discrete devices (B). Device (A) treats, all of the links in a given Multi-Chassis Link Aggregation bundle as a single logical interface and treats all devices in Group (B) as a single logical device for all forwarding purposes. Device (A) does not forward packets receive on Multi-Chassis Link bundle out of the same Multi-Chassis link bundle. Figure 1 depicts a specific use case example.

RPF: Reverse Path Forwarding. See [section 4.5.2 of \[RFC6325\]](#).

3. The AFFINITY sub-TLV

Association of an RBridge to a multi-destination distribution tree through a specific path is accomplished by using a new IS-IS sub-TLV, the Affinity sub-TLV.

The AFFINITY sub-TLV appears in Router Capability TLVs or MT Capability TLVs that are within LSP PDUs, as described in [[RFC7176](#)] which specifies the code point and data structure for the Affinity sub-TLV.

4. Multicast Tree Construction and Use of Affinity Sub-TLV

Figure 1 and Figure 2 below show the reference topology and a logical topology using CMT to provide active-active service.

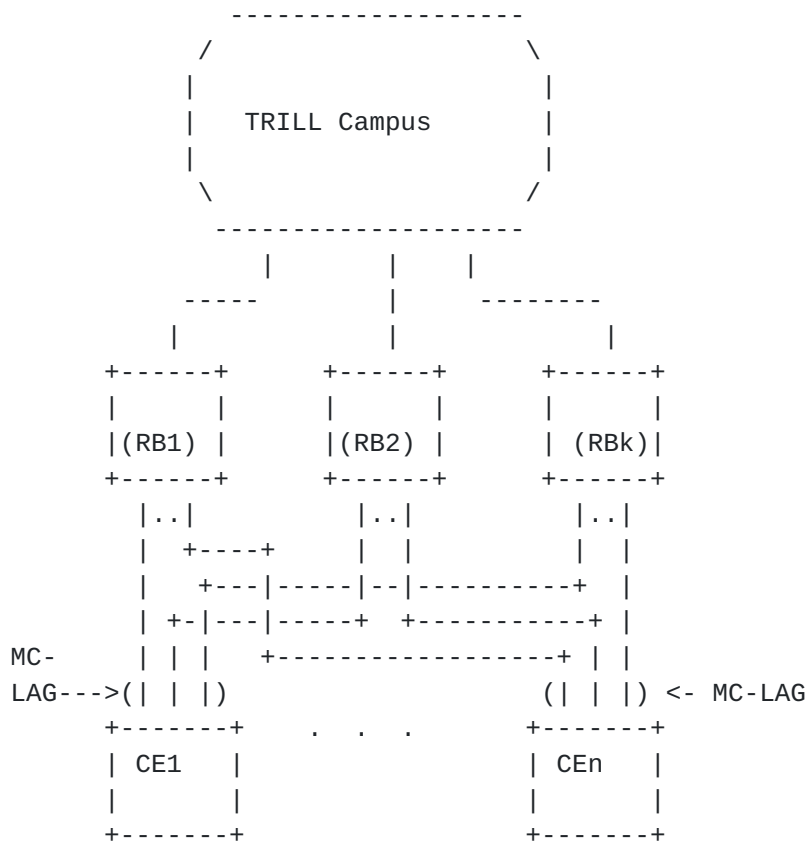


Figure 1 Reference Topology

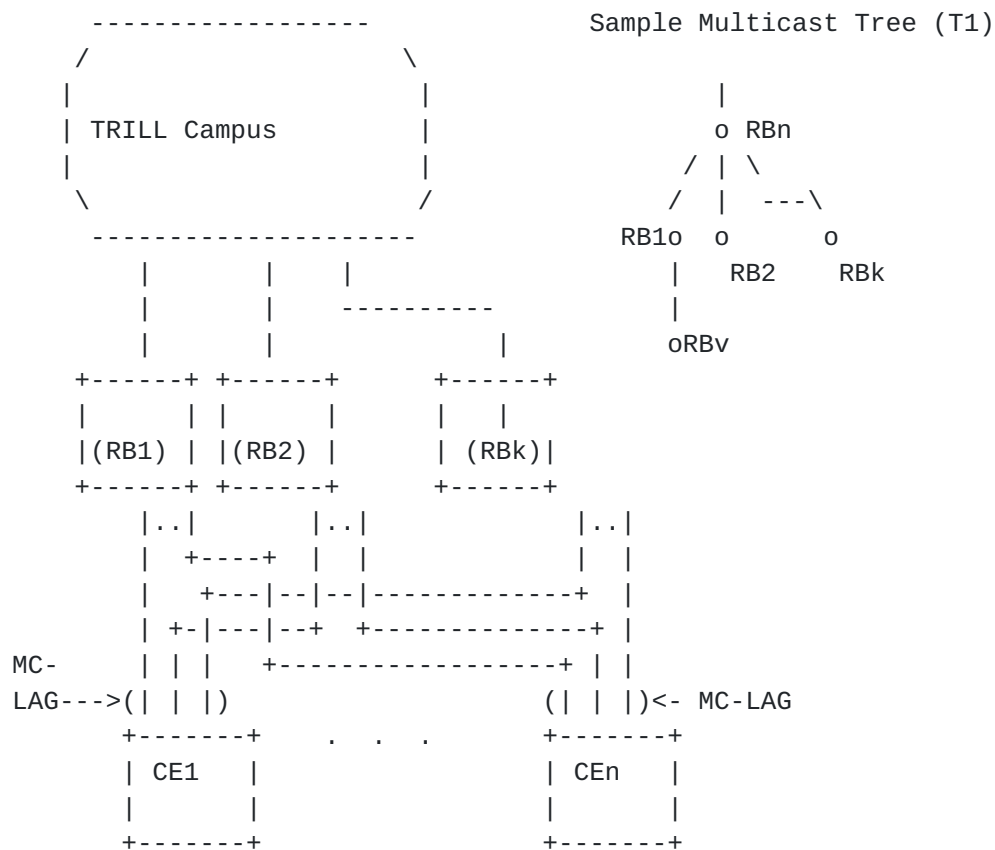


Figure 2 Example Logical Topology

4.1. Update to [RFC 6325](#)

[Section 4.5.1 of \[RFC6325\]](#), is updated to change the calculation of distribution trees as below:

Each RBridge that desires to be the parent RBridge for child RBridge RBy in a multi-destination distribution tree x announces the desired association using an Affinity sub-TLV. The child is specified by its nickname. If an RBridge RB1 advertises an AFFINITY sub-TLV designating one its own nicknames N1 as its 'child' in some distribution tree, the effect is that that nickname N1 is ignored when constructing other distribution trees. Thus the RPF check will enforce that only RB1 can use nickname N1 to do ingress/egress on

tree x. (This has no effect on least cost path calculations for unicast traffic.)

When such an Affinity sub-TLV is present, the association specified by the affinity sub-TLV MUST be used when constructing the multi-destination distribution tree except in case of conflicting Affinity sub-TLV, which are resolved as specified in [Section 5.3](#). In the absence of such an Affinity sub-TLV, or if there are any RBridges in the campus that do not support Affinity sub-TLV, distribution trees are calculated as specified in the [section 4.5.1 of \[RFC6325\]](#) as updated by [\[RFC7180bis\]](#). [Section 4.3](#). below specifies how to identify RBridges that support Affinity sub-TLV capability.

[4.2. Announcing virtual RBridge nickname](#)

Each edge RBridge RB1 to RBk advertises in its LSP virtual RBridge nickname RBv using the Nickname sub-TLV (6), [\[RFC7176\]](#), along with their regular nickname or nicknames.

It will be possible for any RBridge to determine that RBv is a virtual RBridge because each RBridge (RB1 to RBk) this appears to be advertising that it is holding RBv is also advertising an Affinity sub-TLV asking that RBv be its child in one or more trees.

Virtual RBridges are ignored when determining the distribution tree roots for the campus.

All RBridges outside the edge group assume that multi-destination packets with ingress nickname RBv might use any of the distribution trees that any member of the edge group is advertising that it might use.

[4.3. Affinity Sub-TLV Capability.](#)

RBridges that announce the TRILL version sub-TLV [\[RFC7176\]](#) and set the Affinity capability bit ([Section 7](#).) support the Affinity sub-TLV and calculation of multi-destination distribution trees and RPF checks as specified herein.

5. Theory of operation

5.1. Distribution Tree Assignment

Let's assume there are n distribution trees and k edge RBridges in the edge group of interest.

If $n \geq k$

Let's assume edge RBridges are sorted in numerically ascending order by IS-IS System ID such that $RB1 < RB2 < RBk$. Each RBridge in the numerically sorted list is assigned a monotonically increasing number j such that; $RB1=0$, $RB2=1$, $RBi=j$ and $RBi+1=j+1$. (See [Section 4.5 of \[RFC6325\]](#) as modified by [Section 3.4 of \[RFC7180bis\]](#) for how tree numbers are determined.)

Assign each tree to RBi such that tree number $((tree_number) \% k)+1$ is assigned to edge group RBridge i for tree_number from 1 to n . where n is the number of trees, k is the number of edge group RBridges considered for tree allocation, and '%' is the integer division remainder operation.

If $n < k$

Distribution trees are assigned to edge group RBridges $RB1$ to RBn , using the same algorithm as $n \geq k$ case. RBridges $RBn+1$ to RBk do not participate in active-active forwarding process on behalf of RBv .

5.2. Affinity Sub-TLV advertisement

Each RBridge in the $RB1$ through RBk domain advertises an Affinity TLV for RBv to be its child.

As an example, let's assume that $RB1$ has chosen Trees $t1$ and $tk+1$ on behalf of RBv .

$RB1$ advertises affinity TLV; $\{RBv, Num\ of\ Trees=2, t1, tk+1\}$.

Other RBridges in the $RB1$ through RBk edge group follow the same procedure.

5.3. Affinity sub-TLV conflict resolution

In TRILL, multi-destination distribution trees are built outward from the root by each RBridge so that they all derive the same set of distribution trees [[RFC6325](#)].

If an RBridge RB1 advertises an Affinity sub-TLV with an AFFINITY RECORD that asks for RBridge RRoot to be its child in a tree rooted at RRoot, that AFFINITY RECORD is in conflict with TRILL distribution tree root determination and MUST be ignored.

If an RBridge RB1 advertises an Affinity sub-TLV with an AFFINITY RECORD that asks for nickname RBn to be its child in any tree and RB1 is neither adjacent to RBn nor does nickname RBn identify RB1 itself, that AFFINITY RECORD is in conflict with the campus topology and MUST be ignored.

If different RBridges advertise Affinity sub-TLVs that try to associate the same virtual RBridge as their child in the same tree or trees, those Affinity sub-TLVs are in conflict with each other for those trees. The nicknames of the conflicting RBridges are compared to identify which RBridge holds the nickname that is the highest priority to be a tree root, with the System ID as the tiebreaker

The RBridge with the highest priority to be a tree root will retain the Affinity association. Other RBridges with lower priority to be a tree root MUST stop advertising their conflicting Affinity sub-TLV, re-calculate the multicast tree affinity allocation, and, if appropriate, advertise a new non-conflicting Affinity sub-TLV.

Similarly, remote RBridges MUST honor the Affinity sub-TLV from the RBridge with the highest priority to be a tree root (use system-ID as the tie-breaker in the event of conflicting priorities) and ignore the conflicting Affinity sub-TLV entries advertised by the RBridges with lower priorities to be tree roots.

5.4. Ingress Multi-Destination Forwarding

If there is at least one tree on which RBv has affinity via RBk, then RBk performs the following operations, for multi-destination frames received from a CE node:

1. Flood to locally attached CE nodes subjected to VLAN and multicast pruning.

2. Ingress in the TRILL header and assign ingress RBridge nickname as RBv (nickname of the virtual RBridge).
3. Forward to one of the distribution trees, tree x in which RBv is associated with RBk.

5.4.1. Forwarding when $n < k$

If there is no tree on which RBv can claim affinity via RBk (probably because the number of trees n built is less than number of RBridges k announcing the affinity sub-TLV), then RBk MUST fall back to one of the following

1. This RBridge should stop forwarding frames from the CE nodes, and should mark that port as disabled. This will prevent CE nodes from forwarding data on to this RBridge, and only use those RBridges which have been assigned a tree -
-OR-
2. This RBridge tunnels multi-destination frames received from attached native devices to an RBridge RBy that has an assigned tree. The tunnel destination should forward it to the TRILL network, and also to its local access links. (The mechanism of tunneling and handshake between the tunnel source and destination are out of scope of this specification and may be addressed in other documents such as [[ChannelTunnel](#)].)

Above fallback options may be specific to active-active forwarding scenario. However, as stated above, Affinity sub-TLV may be used in other applications. In such event the application SHOULD specify applicable fallback options.

5.5. Egress Multi-Destination Forwarding

5.5.1. Traffic Arriving on an assigned Tree to RBk-RBv

Multi-destination frames arriving at RBk on a Tree x, where RBk has announced the affinity of RBv via x, MUST be forwarded to CE members of RBv that are in the frame's VLAN. Forwarding to other end-nodes and RBridges that are not part of the network represented by the RBv virtual RBridge MUST follow the forwarding rules specified in [[RFC6325](#)].

5.5.2. Traffic Arriving on other Trees

Multi-destination frames arriving at RBk on a Tree y, where RBk has not announced the affinity of RBv via y, MUST NOT be forwarded to CE members of RBv. Forwarding to other end-nodes and RBridges that are

not part of the network represented by the RBv virtual RBridge MUST follow the forwarding rules specified in [[RFC6325](#)].

5.6. Failure scenarios

The below failure recovery algorithm is presented only as a guideline. Implementations MAY include other failure recover algorithms. Details of such algorithms are outside the scope of this document.

5.6.1. Edge RBridge RBk failure

Each of the member RBridges of a given virtual RBridge edge group is aware of its member RBridges through configuration, LSP advertisements, or some other method.

Member RBridges detect nodal failure of a member RBridge through IS-IS LSP advertisements or lack thereof.

Upon detecting a member failure, each of the member RBridges of the RBv edge group start recovery timer T_{rec} for failed RBridge RB_i . If the previously failed RBridge RB_i has not recovered after the expiry of timer T_{rec} , members RBridges perform the distribution tree assignment algorithm specified in [section 5.1](#). Each of the member RBridges re-advertises the Affinity sub-TLV with new tree assignment. This action causes the campus to update the tree calculation with the new assignment.

RB_i , upon start-up, advertises its presence through IS-IS LSPs and starts a timer T_i . Other member RBridges of the edge group, detecting the presence of RB_i , start a timer T_j .

Upon expiry of timer T_j , other member RBridges recalculate the multi-destination tree assignment and advertised the related trees using Affinity sub-TLV. Upon expiry of timer T_i , RB_i recalculate the multi-destination tree assignment and advertises the related trees using Affinity TLV.

If the new RBridge in the edge group calculates trees and starts to use one or more before the existing RBridges in the edge group recalculate, there could be duplication of packets (for example more than one edge group RBridge could decapsulate and forward a multi-destination frame on links into the active-active group) or loss of packets (for example due to the Reverse Path Forwarding

Check in the rest of the campus if two edge group RBridges are trying to forward on the same tree those from one will be discarded). Alternatively, if the new RBridge in the edge group calculates trees and starts to use one or more after the existing RBridges recalculate, there could be loss of data due to frames arriving at the new RBridge being black holed. Timers T_i and T_j should be initialized to values designed to minimize these problems keeping in mind that, in general, duplicating is a more serious problem than dropping. It is RECOMMENDED that T_j be less than T_i and a reasonable default is $1/2$ of T_i .

5.7. Backward compatibility

Implementations MUST support backward compatibility mode to interoperate with pre Affinity sub-TLV RBridges in the network. Such backward compatibility operation MAY include, however is not limited to, tunneling and/or active-standby modes of operations.

Example:

- Step 1. Stop using virtual RBridge nickname for traffic ingressing from CE nodes
- Step 2. Stop performing active-active forwarding. And fall back to active standby forwarding, based on locally defined policies. Definition of such policies is outside the scope of this document and may be addressed in other documents.

6. Security Considerations

In general, the RBridges in a campus are trusted routers and the authenticity of their link state information (LSPs) and link local PDUs (Hellos, etc.) can be enforced using regular IS-IS security mechanisms [[IS-IS](#)] [[RFC5310](#)]. This including authenticating the contents of the PDUs used to transport Affinity sub-TLVs.

The particular Security Considerations involved with different applications of the Affinity sub-TLV will be covered in the document(s) specifying those applications.

For general TRILL Security Considerations, see [[RFC6325](#)].

7. IANA Considerations

This document serves as the reference for "TRILL-VER Sub-TLV Capability Flags" registration "Affinity sub-TLV support." (bit 0) so that reference should be updated when this document is published as an RFC.

This document mentions "Sub-TLVs for TLV 144" registration "AFFINITY" (value 17), but should not be listed as a reference for that registration which should remain [[RFC7176](#)].

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.
- [RFC5310] Bhatia, M., et.al. "'IS-IS Generic Cryptographic Authentication'", [RFC 5310](#), February 2009.
- [RFC6325] Perlman, R., et.al. "'RBridge: Base Protocol Specification'", [RFC 6325](#), July 2011.
- [RFC7177] Eastlake 3rd, D. et.al., "'RBridge: Adjacency'", [RFC 7177](#), May 2014.
- [RFC6439] Eastlake 3rd, D. et.al., "'RBridge: Appointed Forwarder'", [RFC 6439](#), November 2011.
- [RFC7172] Eastlake 3rd, D., Zhang, M., Agarwal, P., Perlman, R., and D. Dutt, "Transparent Interconnection of Lots of Links (TRILL): Fine-Grained Labeling", [RFC 7172](#), May 2014, <<http://www.rfc-editor.org/info/rfc7172>>.
- [RFC7176] Eastlake 3rd, D. et.al., "'Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS'", [RFC 7176](#), May 2014.
- [RFC7180bis] Eastlake 3rd, D. et.al., "'TRILL: Clarifications, Corrections, and Updates'", [draft-ietf-trill-rfc7180bis](#), work in progress.
- [IS-IS] ISO/IEC, "'Intermediate System to Intermediate System Routing Information Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)'" ISO/IEC 10589:2002.

8.2. Informative References

- [RFC7379] Li, Y., Hao, W., Perlman, R., Hudson, J., and H. Zhai, "Problem Statement and Goals for Active-Active Connection at the Transparent Interconnection of Lots of Links (TRILL) Edge", [RFC 7379](#), October 2014.
- [TRILLPN] Zhai, H., et.al 'RBridge: Pseudonode Nickname', [draft-hu-trill-pseudonode-nickname](#), Work in progress, November 2011.
- [8021AX] IEEE, 'Link Aggregation', IEEE Std 802.1AX-2014, December 2014.
- [ChannelTunnel] D. Eastlake and Y. Li, "TRILL: RBridge Channel Tunnel Protocol", [draft-ietf-trill-channel-tunnel](#), work in progress.

9. Acknowledgments

Authors wish to extend their appreciations towards individuals who volunteered to review and comment on the work presented in this document and provided constructive and critical feedback. Specific acknowledgements are due for Anoop Ghanwani, Ronak Desai, Gayle Noble, and Varun Shah. Very special Thanks to Donald Eastlake for his careful review and constructive comments.

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

[Appendix A](#). Change History.

From -01 to -02:

Replaced all references to ''LAG'' with references to Multi-Chassis (MC-LAG) or the like.

Expanded, Security Considerations section.

Other editorial changes.

From -02 to -03

Minor editorial changes

From -03 to -04

Minor editorial changes and version update.

From -04 to -05

Editorial, reference, and other minor changes based on Document Shepherd review.

From -05 to -6

Minor editorial fixes.

From -06 to -07

Clarifications primarily based on AD review.

Authors' Addresses

Tissa Senevirathne
Consultant

Email: tsenevir@gmail.com

Janardhanan Pathangi
Dell/Force10 Networks
Olympia Technology Park,
Guindy Chennai 600 032

Phone: +91 44 4220 8400
Email: Pathangi_Janardhanan@Dell.com

Jon Hudson
Brocade
130 Holger Way
San Jose, CA 95134 USA

Phone: +1-408-333-4062
Email: jon.hudson@gmail.com