Internet Engineering Task Force (IETF)
Request for Comments: 7968
Category: Standards Track
ISSN: 2070-1721

Y. Li
D. Eastlake 3rd
W. Hao
H. Chen
Huawei Technologies
S. Chatterjee
Cisco

August 2016

Transparent Interconnection of Lots of Links (TRILL): Using Data Labels for Tree Selection for Multi-Destination Data

Abstract

TRILL (Transparent Interconnection of Lots of Links) uses distribution trees to deliver multi-destination frames. Multiple trees can be used by an ingress Routing Bridge (RBridge) for flows, regardless of the VLAN, Fine-Grained Label (FGL), and/or multicast group of the flow. Different ingress RBridges may choose different distribution trees for TRILL Data packets in the same VLAN, FGL, and/or multicast group. To avoid unnecessary link utilization, distribution trees should be pruned based on one or more of the following: VLAN, FGL, or multicast destination address. If any VLAN, FGL, or multicast group can be sent on any tree, for typical fast-path hardware, the amount of pruning information is multiplied by the number of trees, but there is limited hardware capacity for such pruning information.

This document specifies an optional facility to restrict the TRILL Data packets sent on particular distribution trees by VLAN, FGL, and/or multicast groups, thus reducing the total amount of pruning information so that it can more easily be accommodated by fast-path hardware.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7968.

Table of Contents

1. Introduction

1.1. Background Description

One or more distribution trees, identified by their root nicknames, are used to distribute multi-destination data in a (Transparent Interconnection of Lots of Links) (TRILL) campus [RFC6325]. The Routing Bridge (RBridge) having the highest tree root priority announces the total number of trees that should be computed for the campus. It may also specify the list of trees that RBridges need to compute using the Tree Identifiers (TREE-RT-IDs) sub-TLV [RFC7176]. Every RBridge can specify the trees it will use for multi-destination TRILL Data packets it originates in the Trees Used Identifiers (TREE-USE-IDs) sub-TLV [RFC7176], and the VLANs or Fine-Grained Labels (FGLs) [RFC7172] it is interested in are specified in Interested VLANs and/or Interested Labels sub-TLVs [RFC7176]. It is suggested that by default the ingress RBridge uses the distribution tree whose root is the closest [RFC6325]. The TREE-USE-IDs sub-TLV is used to build the RPF (Reverse Path Forwarding) check table that is used for RPF checking. Interested VLANs and Interested Labels sub-TLVs are used for distribution tree pruning, and the multi-destination forwarding table with pruning information is built based on that RPF check table. To reduce unnecessary link loads, each distribution tree should be pruned per VLAN/FGL, eliminating branches that have no potential receivers downstream as specified in [RFC6325]. Further pruning based on Layer 2 or Layer 3 multicast addresses is also possible.

Defaults are provided, but how many trees are calculated, where the tree roots are located, and which tree or trees are to be used by an ingress RBridge are implementation dependent. With the increasing demand to use TRILL in data center networks, there are some features we can explore for multi-destination frames in the data center use case. In order to achieve non-blocking data forwarding, a fat tree structure is often used. Figure 1 shows a typical data center network based on the fat tree structure. RB1 and RB2 are aggregation switches, and RB11 through RB14 are access switches. It is a common practice to configure the tree roots to be at the aggregation switches for efficient traffic transportation. All the ingress RBridges that are access switches will then be equally distant from all the tree roots.

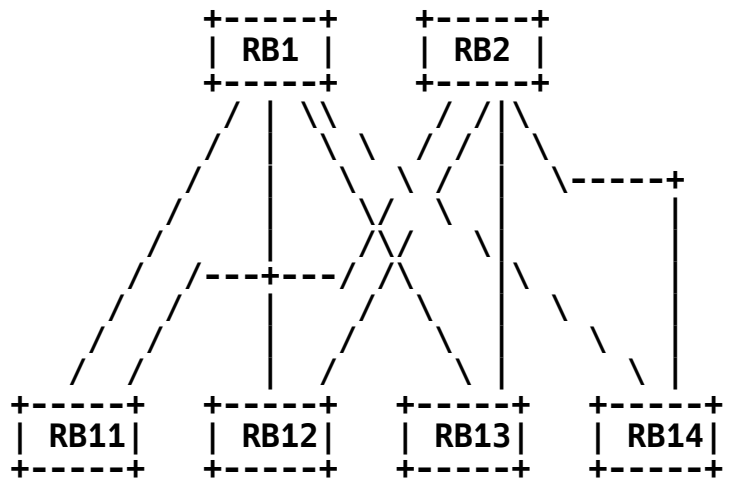


Figure 1: TRILL Network Based on Fat Tree Structure

1.2. Terminology Used in This Document

This document uses the terminology from [RFC6325] and [RFC7172], some of which is repeated below for convenience, along with some additional terms listed below:

Campus: The name for a network using the TRILL protocol in the same sense that a "bridged LAN" is the name for a network using bridging. In TRILL, the word "campus" has no academic implication.

Data Label: VLAN or FGL.

ECMP: Equal-Cost Multipath [RFC6325].

FGL: Fine-Grained Label [RFC7172].

Interested Labels sub-TLV: Short for "Interested Labels and Spanning Tree Roots sub-TLV" [RFC7176].

Interested VLANs sub-TLV: Short for "Interested VLANs and Spanning Tree Roots sub-TLV" [RFC7176].

IPTV: "Television" (video) over IP.

RBridge: An alternative name for a TRILL switch.

RPF: Reverse Path Forwarding.

TRILL: Transparent Interconnection of Lots of Links (or Tunneled Routing in the Link Layer).

TRILL switch: A device implementing the TRILL protocol. Sometimes called an RBridge.

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].

2. Motivations

In the structure of Figure 1, if we choose to put the tree roots at RB1 and RB2, the ingress RBridge (e.g., RB11) would find more than one equal-cost closest tree root (i.e., RB1 and RB2). An ingress RBridge has two options to select the tree root for multi-destination frames: choose one and only one as the distribution tree root, or use an ECMP-like algorithm to balance the traffic among the multiple trees whose roots are at the same distance from the RBridge.

- For the former (one distribution tree root), a single tree used by each ingress RBridge can have the problem of uneven or inefficient link usage. For example, if RB11 chooses the tree that is rooted at RB1 as the distribution tree, the link between RB11 and RB2 will not be used for multi-destination frames ingressed by RB11.

- For the latter (an ECMP-like algorithm), ECMP-based tree selection results in a linear increase in multicast forwarding table size with the number of trees, as explained in the next paragraph.

A multicast forwarding table at an RBridge is normally used to map the key of (distribution tree nickname + VLAN) to an index to a list of ports for multicast packet replication. The key used for mapping is simply the tree nickname when the RBridge does not prune the tree.

The key could be the distribution tree nickname augmented by the FGL and/or Layer 2 or 3 multicast address when the RBridge supports FGL and/or Layer 2 or 3 pruning information.

For any RBridge RBn, for each VLAN x, if RBn is in a distribution tree t used by traffic in VLAN x, there will be an entry of (t, x, port list) in the multicast forwarding table on RBn. Typically, each entry contains a distinct combination of (tree nickname, VLAN) as the lookup key. If there are n such trees and m such VLANs, the multicast forwarding table size on RBn is n*m entries. If an FGL is used [RFC7172] and/or finer pruning is used (for example, VLAN + multicast group address is used for pruning), the value of m increases. In the larger-scale data center, more trees would be necessary for purposes of better load-balancing; this results in an increased value for n. In either case, the number of table entries (i.e., n*m) will increase dramatically.

The left-hand table in Figure 2 shows an example of the multicast forwarding table on RB11 in the Figure 1 topology, with two distribution trees in a campus using typical fast-path hardware.

Before VLAN-Based Tree Selection

tree nickname	VLAN	port list
tree 1	1	
tree 1	2	
tree 1	...	
tree 1	...	
tree 1	...	
tree 1	4093	
tree 1	4094	
tree 2	1	
tree 2	2	
tree 2	...	
tree 2	...	
tree 2	...	
tree 2	...	
tree 2	4093	
tree 2	4094	

After VLAN-Based Tree Selection

tree nickname	VLAN	port list
tree 1	1	
tree 1	2	
tree 1	...	
tree 1	1999	
tree 1	2000	
tree 2	2001	
tree 2	2002	
tree 2	...	
tree 2	4093	
tree 2	4094	

Figure 2: Multicast Forwarding Table before and after Using VLAN-Based Tree Selection

The number of entries is approximately 2*4K in this case. If four distribution trees are used in a TRILL campus and RBn has 4K VLANs with downstream receivers, it consumes 16K table entries. The size of fast-path TRILL multicast forwarding tables is typically limited by hardware; therefore, the table entries are a precious resource.

In some implementations, the table is shared with Layer 3 IP multicast for a total of 16K or 8K table entries. Therefore, we want to reduce the table size consumed for TRILL distribution trees as much as possible and at the same time maintain load-balancing among the trees.

In cases where blocks of consecutive VLANs or FGLs can be assigned to a tree, the multicast forwarding table could be greatly compressed if entries could have a Data Label value and mask, with the fast-path hardware doing the longest prefix matching. But few, if any, fast-path implementations provide such logic.

A straightforward way to alleviate the problem of limited table entries is not to prune the distribution tree. However, this can only be used in restricted scenarios, for the following reasons:

- Not pruning wastes bandwidth for multi-destination packets. There is normally broadcast traffic, like ARP and unknown unicast, that can be pruned on a VLAN (or FGL) so that it is not sent down branches of a distribution tree where it is not needed. In addition, if there is a lot of Layer 3 multicast traffic, no pruning may result in a worst-case scenario where that user data is unnecessarily flooded all over the campus. The volume of flooded data could be very large if certain applications such as IPTV are supported. More precise pruning, such as pruning based on multicast groups, may be desirable in this case.

- Not pruning is only useful at pure transit nodes. Edge nodes always need to maintain the multicast forwarding table with the key of (tree nickname + VLAN (or FGL)), since the edge node needs to decide whether and how to replicate the frame to local access ports. It is likely that edge nodes are relatively low-end switches with a smaller shared table size, say 4K, available.

- Due to security concerns, VLAN-based (or FGL-based) traffic isolation is a basic requirement in some scenarios. No pruning may increase the risk of leakage of the traffic. Misbehaving RBridges may take advantage of this leakage of traffic.

In addition to the concern regarding multicast table size, some silicon does not currently support hashing-based tree nickname selection at the ingress RBridge but commonly uses VLAN-based tree selection. If the control plane of the ingress RBridge maps the incoming VLAN x to a tree nickname t, the data plane will always use tree t for VLAN x multi-destination frames. Such an ingress RBridge may choose multiple trees to be used for load-sharing; it can use one and only one tree for each VLAN. If we make sure that all ingress

RBridges campus-wide send VLAN x multi-destination packets only use tree t, then there would be no need to store the multicast table entry with the key of (tree-other-than-t, x) on any RBridge.

This document describes the TRILL control-plane support for distribution tree selection based on a VLAN, FGL, and/or multicast address to reduce the multicast forwarding table size. It is compatible with the silicon implementations mentioned in the previous paragraph.

3. Tree Selection Based on Data Labels

Data Label (VLAN-based or FGL-based) tree selection can be used as a distribution tree selection mechanism, especially when the multicast forwarding table size is a concern. This section specifies that mechanism and how to extend it so that tree selection can be based on multicast groups.

3.1. Overview of the Mechanism

The RBridge that has the highest priority to be a tree root announces the tree nicknames and the Data Labels allowed on each tree. Such announcements of correspondence of tree to Data Label can be based on static configuration or some predefined algorithm beyond the scope of this document. An ingress RBridge selects the tree-VLAN correspondence that it wishes to use from the list announced by the highest-priority tree root. It SHOULD NOT transmit VLAN x frames on tree y if the highest-priority tree root does not say that VLAN x is allowed on tree y.

If we make sure that a particular VLAN is allowed on one and only one tree, we can keep the number of multicast forwarding table entries on any RBridge fixed at 4K maximum (or up to 16M in the case of an FGL). Take Figure 1 as an example, where two trees are rooted at RB1 and RB2, respectively. The highest-priority tree root appoints tree 1 to carry VLAN 1-2000 and tree 2 to carry VLAN 2001-4094. With such an announcement by the highest-priority tree root, every RBridge that understands the announcement will not send VLAN 2001-4094 traffic on tree 1 and will not send VLAN 1-2000 traffic on tree 2. That way, no RBridge would need to store the entries for tree 1 / VLAN 2001-4094 or tree 2 / VLAN 1-2000. Figure 2 shows the multicast forwarding table on an RBridge before and after we use VLAN-based tree selection. The number of entries is reduced by a factor f, where f is the number of trees used in the campus. In this example, it is reduced from 2*4094 to 4094. This affects both transit nodes and edge nodes. The data-plane encoding does not change.

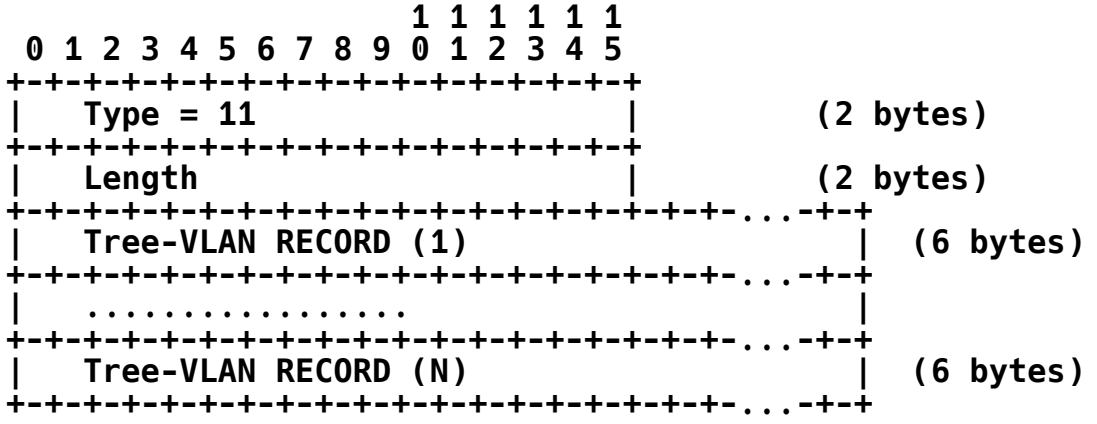
3.2. APPsub-TLVs Supporting Tree Selection

Six new APPsub-TLVs that can be carried in the TRILL GENINFO TLV [RFC7357] in Extended Level 1 Flooding Scope (E-L1FS) FS-Link State Protocol Data Units (FS-LSPs) [RFC7780] are defined below. The first four can be considered analogous to finer-granularity versions of the TREE-RT-IDs sub-TLV and the TREE-USE-IDs sub-TLV [RFC7176]. Two APPsub-TLVs supporting VLAN-based tree selection are specified in Sections 3.2.1 and 3.2.2. They are used by the highest-priority tree root to announce the allowed VLANs on each tree in the campus and by an ingress RBridge to announce the tree-VLAN correspondence that it selects from the list announced by the highest-priority tree root. Two APPsub-TLVs supporting FGL-based tree selection are specified in Sections 3.2.3 and 3.2.4 for the same purpose. Sections 3.2.5 and 3.2.6 define two APPsub-TLVs to support finer granularity in selecting trees based on multicast groups rather than Data Labels.

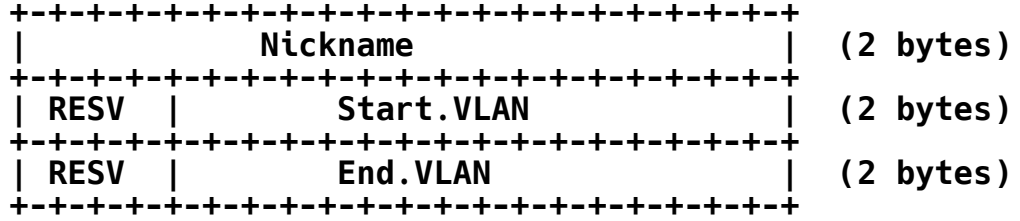
New APPsub-TLVs	Description
Tree and VLANs	announcement by the highest-priority tree root of the VLANs allowed per tree
Tree and VLANs Used	tree-VLAN correspondence that an ingress RBridge selects
Tree and FGLs	announcement by the highest-priority tree root of the FGLs allowed per tree
Tree and FGLs Used	tree-FGL correspondence that an ingress RBridge selects
Tree and Groups	announcement by the highest-priority tree root of the multicast groups allowed on each tree
Tree and Groups Used	tree and multicast group correspondence that an ingress RBridge selects

3.2.1. The Tree and VLANs APPsub-TLV

The RBridge that is the highest-priority tree root announces the VLANs allowed on each tree with the Tree and VLANs (TREE-VLANs) APPsub-TLV. Multiple instances of this APPsub-TLV may be carried. The same tree nicknames may occur in multiple Tree-VLAN RECORDs within the same APPsub-TLV or across multiple APPsub-TLVs. The APPsub-TLV format is as follows:



where each Tree-VLAN RECORD is of the form:



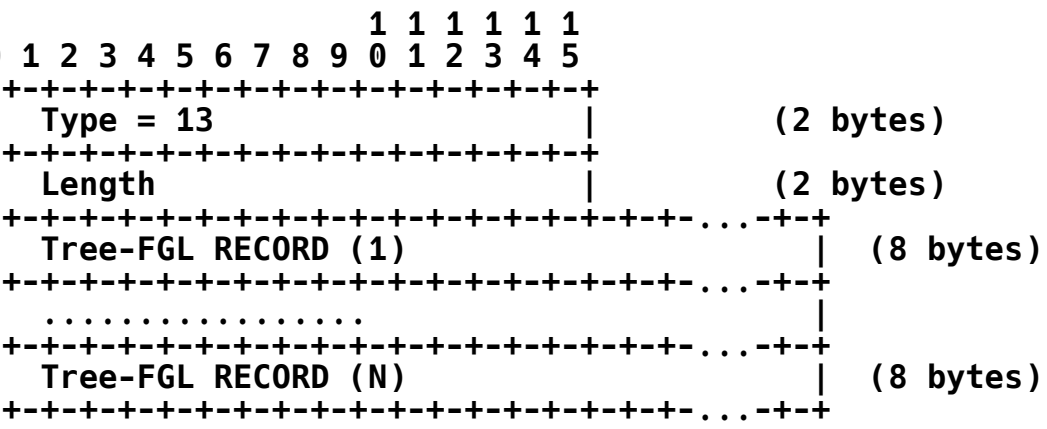
- Type: TRILL GENINFO APPsub-TLV type; set to 11 (TREE-VLANs).
- Length: 6*n bytes, where there are n Tree-VLAN RECORDs. Thus, the value of Length can be used to determine n. If Length is not a multiple of 6, the APPsub-TLV is corrupt and MUST be ignored.
- Nickname: The nickname identifying the distribution tree by its root.
- RESV: 4 bits that MUST be sent as zero and ignored on receipt.
- Start.VLAN, End.VLAN: These fields are the VLAN IDs of the allowed VLAN range on the tree, inclusive. To specify a single VLAN, the VLAN's ID appears as both the start and end VLAN. If End.VLAN is less than Start.VLAN, the Tree-VLAN RECORD MUST be ignored.

3.2.2. The Tree and VLANs Used APPsub-TLV

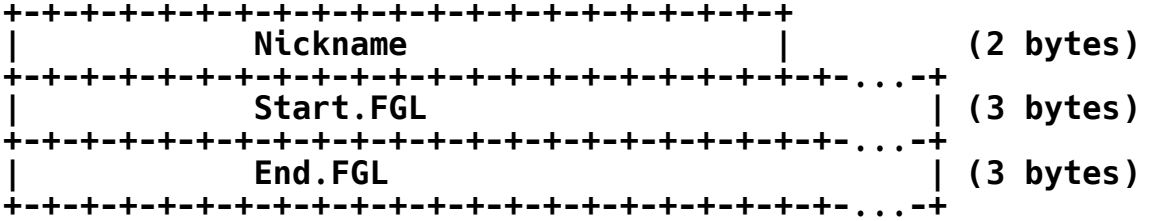
This APPsub-TLV has the same structure as the TREE-VLANs APPsub-TLV specified in Section 3.2.1. The differences are that its APPsub-TLV type is set to 12 (TREE-VLAN-USE) and the tree-VLAN correspondences in the Tree-VLAN RECORDs listed are those correspondences that the originating RBridge wants to use for multi-destination packets. This APPsub-TLV is used by an ingress RBridge to distribute the tree-VLAN correspondence that it selects from the list announced by the highest-priority tree root.

3.2.3. The Tree and FGLs APPsub-TLV

The RBridge that is the highest-priority tree root can use the Tree and FGLs (TREE-FGLs) APPsub-TLV to announce the FGLs allowed on each tree. Multiple instances of this APPsub-TLV may be carried. The same tree nicknames may occur in the multiple Tree-FGL RECORDs within the same APPsub-TLV or across multiple APPsub-TLVs. Its format is as follows:



where each Tree-FGL RECORD is of the form:



- Type: TRILL GENINFO APPsub-TLV type; set to 13 (TREE-FGLs).
- Length: 8*n bytes, where there are n Tree-FGL RECORDs. Thus, the value of Length can be used to determine n. If Length is not a multiple of 8, the APPsub-TLV is corrupt and MUST be ignored.
- Nickname: The nickname identifying the distribution tree by its root.
- RESV: 4 bits that MUST be sent as zero and ignored on receipt.
- Start.FGL, End.FGL: These fields are the FGL IDs of the allowed FGL range on the tree, inclusive. To specify a single FGL, the FGL's ID appears as both the start and end FGL. If End.FGL is less than Start.FGL, the Tree-FGL RECORD MUST be ignored.

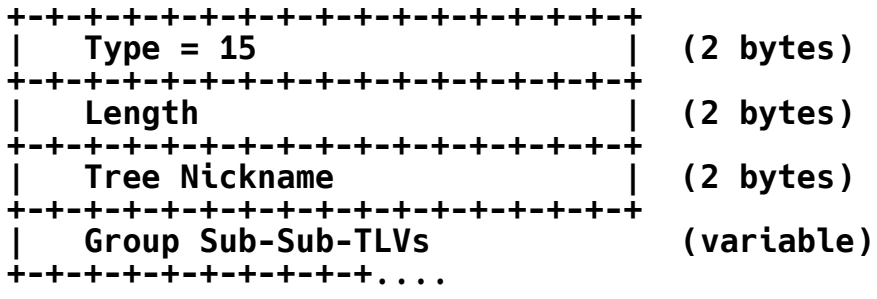
3.2.4. The Tree and FGLs Used APPsub-TLV

This APPsub-TLV has the same structure as the TREE-FGLs APPsub-TLV specified in Section 3.2.3. The differences are that its APPsub-TLV type is set to 14 (TREE-FGL-USE) and the Tree-FGL correspondences in the Tree-FGL RECORDs listed are those that the originating RBridge wants to use for multi-destination packets. This APPsub-TLV is used by an ingress RBridge to distribute the tree-FGL correspondence that it selects from the list announced by the highest-priority tree root.

3.2.5. The Tree and Groups APPsub-TLV

Tree selection based on Data Labels is easily extended to tree selection based on Data Label + Layer 2 or 3 multicast groups. We can appoint multicast group 1 in VLAN 10 to tree 1 and appoint group 2 in VLAN 10 to tree 2 for better load-sharing.

The RBridge that is the highest-priority tree root can announce the multicast groups allowed on each tree for each Data Label with the Tree and Groups (TREE-GROUPs) APPsub-TLV. Multiple instances of this APPsub-TLV may be carried. The APPsub-TLV format is as follows:



- Type: TRILL GENINFO APPsub-TLV type; set to 15 (TREE-GROUPs).
- Length: 2 + the length of the Group Sub-Sub TLVs that are included.
- Nickname: The nickname identifying the distribution tree by its root.
- Group Sub-Sub-TLVs: Zero or more of the TLV structures that are allowed as sub-TLVs of the Group Address (GADDR) TLV [RFC7176]. Each such TLV structure specifies a multicast group and either a VLAN or FGL. Although these TLV structures are considered sub-TLVs when they appear inside a GADDR TLV, they are technically sub-sub-TLVs when they appear inside a TREE-GROUPs APPsub-TLV that is in turn inside a TRILL GENINFO TLV [RFC7357].

3.2.6. The Tree and Groups Used APPsub-TLV

The Tree and Groups Used (TREE-GROUPs-USE) APPsub-TLV has the same structure as the TREE-GROUPs APPsub-TLV specified in Section 3.2.5. The differences are that its APPsub-TLV type is set to 16 (TREE-GROUPs-USE) and the Tree Nickname and Group sub-sub-TLVs listed in this APPsub-TLV are those that the originating RBridge wants to use for multi-destination packets. This APPsub-TLV is used by an ingress RBridge to distribute the tree-group correspondence that it selects from the list announced by the highest-priority tree root.

3.3. Detailed Processing

The highest-priority tree root RBridge MUST include all the necessary tree-related sub-TLVs defined in [RFC7176] as usual in its E-L1FS FS-LSP and MAY include the TREE-VLANs APPsub-TLV and/or the TREE-FGLs APPsub-TLV in its E-L1FS FS-LSP [RFC7780]. In this way, it MAY indicate that each VLAN and/or FGL is only allowed on one or some other number of trees less than the number of trees being calculated in the campus in order to save table space in the fast-path forwarding hardware.

An ingress RBridge that understands the TREE-VLANs APPsub-TLV SHOULD select the tree-VLAN correspondences that it wishes to use and put them in TREE-VLAN-USE APPsub-TLVs. If there are multiple tree nicknames announced in a TREE-VLANs APPsub-TLV for VLAN x, the ingress RBridge chooses one of them if it supports this feature. For example, the ingress RBridge may choose the closest (minimum-cost) root among them. How to make such a choice is out of scope for this document. It may be desirable to have some fixed algorithm to make sure that all ingress RBridges choose the same tree for VLAN x in this case. Any single Data Label that the ingress RBridge is

interested in should be related to only one tree ID in a TREE-VLAN-USE APPsub-TLV to minimize the multicast forwarding table size on other RBridges, but as long as the Data Label is related to less than all the trees being calculated, it will reduce the burden on the forwarding table size.

When an ingress RBridge encapsulates a multi-destination frame for Data Label x, it SHOULD use a tree nickname that it selected previously in a TREE-VLAN-USE or TREE-FGL-USE APPsub-TLV for Data Label x. However, that may not be possible because either (1) the RBridge may not have advertised such TREE-VLAN-USE or TREE-FGL-USE APPsub-TLVs, in which case it can use any tree that has been advertised as permitted for the Data Label by the highest-priority tree root RBridge, or (2) the tree or trees it advertised might be unavailable due to failures.

If RBridge RBn does not perform pruning, it builds the multicast forwarding table as specified in [RFC6325].

If RBn prunes the distribution tree based on VLANs, RBn uses the information received in TREE-VLAN-USE APPsub-TLVs to mark the set of VLANs reachable downstream for each adjacency and for each related tree. If RBn prunes the distribution tree based on FGLs, RBn uses the information received in TRILL-FGL-USE APPsub-TLVs to mark the set of FGLs reachable downstream for each adjacency and for each related tree.

Logically, an ingress RBridge that does not support VLAN-based or FGL-based tree selection is equivalent to the one that supports it but uses it in such a way as to gain no advantage; for example, it announces the use of all trees for all VLANs and FGLs.

3.4. Failure Handling

This section discusses failure scenarios for a distribution tree root for the case where that tree root is not the highest-priority root and the case where it is the highest-priority root. This section also discusses some other transient error conditions.

Failure of a tree root that is not the highest-priority tree root:
It is the responsibility of the highest-priority tree root to inform other RBridges of any change in the allowed tree-VLAN correspondence. When the highest-priority tree root learns that the root of tree t has failed, it should reassign the VLANs allowed on tree t to other trees or to a tree replacing the failed one.

Failure of the highest-priority tree root: It is suggested that the tree root of second-highest priority be pre-configured with the proper knowledge of the tree-VLAN correspondence allowed when the highest-priority tree root fails. The information announced by the RBridge that has the second-highest priority to be a tree root would be in the link state of all RBridges but would not take effect unless the RBridge noticed the failure of the highest-priority tree root. When the highest-priority tree root fails, the tree root that formerly had second-highest priority will become the highest-priority tree root of the campus. When an RBridge notices the failure of the original highest-priority tree root, it can immediately use the stored information announced by the tree root that originally had second-highest priority. It is suggested that the tree-VLAN correspondence information be pre-configured on the tree root of second-highest priority to be the same as that on the highest-priority tree root for the trees other than the highest-priority tree itself. This can minimize the change to multicast forwarding tables in the case of highest-priority tree root failure. For a large campus, it may make sense to pre-configure this information in a similar way on the third-priority, fourth-priority, or even lower-priority tree root RBridges.

In some transient conditions, or in the case of a misbehaving highest-priority tree root, an ingress RBridge may encounter the following scenarios:

- No tree has been announced for which VLAN x frames are allowed.

- An ingress RBridge is supposed to transmit VLAN x frames on tree t, but the root of tree t is no longer reachable.

For the second case, an ingress RBridge may choose another reachable tree root that allows VLAN x frames according to the highest-priority tree root announcement. If there is no such tree available, then it is the same as the first case above. The ingress RBridge should then be "downgraded" to a conventional RBridge with behavior as specified in [RFC6325]. A timer should be set to allow the temporary transient stage to complete before the change of the responsive tree or the downgrade takes effect. The value of the timer should be set to at least the LSP flooding time of the campus.

4. Backward Compatibility

RBridges MUST include the TREE-USE-IDs and INT-VLAN sub-TLVs in their LSPs when required by [RFC6325] whether or not they support the new TREE-VLAN-USE or TREE-FGL-USE APPsub-TLVs specified by this document.

RBridges that understand the new TREE-VLAN-USE APPsub-TLV sent from another RBridge RBn should use it to build the multicast forwarding table and ignore the TREE-USE-IDs and INT-VLAN sub-TLVs sent from the same RBridge. TREE-USE-IDs and INT-VLAN sub-TLVs are still useful for some purposes other than building the multicast forwarding table (e.g., building an RPF table, spanning tree root notification). If the RBridge does not receive TREE-VLAN-USE APPsub-TLVs from RBn, it uses the conventional way described in [RFC6325] to build the multicast forwarding table.

For example, there are two distribution trees, tree 1 and tree 2, in the campus. RB1 and RB2 are RBridges that use the new APPsub-TLVs described in this document. RB3 is an old RBridge that is compatible with [RFC6325]. Assume that RB2 is interested in VLANs 10 and 11 and RB3 is interested in VLANs 100 and 101. Hence, RB1 receives ((tree 1, VLAN 10), (tree 2, VLAN 11)) as a TREE-VLAN-USE APPsub-TLV and (tree 1, tree 2) as a TREE-USE-IDs sub-TLV from RB2 on port x. Also, RB1 receives (tree 1) as a TREE-USE-IDs sub-TLV and no TREE-VLAN-USE APPsub-TLV from RB3 on port y. RB2 and RB3 announce their interested VLANs in an INT-VLAN sub-TLV as usual. RB1 will then build the entry of (tree 1, VLAN 10, port x) and (tree 2, VLAN 11, port x) based on RB2's LSP and the mechanism specified in this document. RB1 also builds entries of (tree 1, VLAN 100, port y), (tree 1, VLAN 101, port y), (tree 2, VLAN 100, port y), and (tree 2, VLAN 101, port y) based on RB3's LSP in the conventional way.

The multicast forwarding table on RB1 with a merged entry would be like the following:

tree nickname	VLAN	port list
tree 1	10	x
tree 1	100	y
tree 1	101	y
tree 2	11	x
tree 2	100	y
tree 2	101	y

As expected, that table is not as small as the one where every RBridge supports the new TREE-VLAN-USE APPsub-TLVs. In a hybrid campus, the worst case would be where the number of entries is equal to the number of entries required by the current practice that does not support VLAN-based tree selection. Such an extreme case happens when the set of interested VLANs from the new RBridges is a subset of the set of interested VLANs from the old RBridges.

Tree selection based on the Data Label and multicast group is compatible with the current practice. Its effectiveness increases with more RBridges supporting this feature in the TRILL campus.

5. Security Considerations

This document does not change the general RBridge security considerations of the TRILL base protocol. The APPsub-TLVs specified can be secured using the IS-IS authentication feature [RFC5310]. See Section 6 of [RFC6325] for general TRILL security considerations.

6. IANA Considerations

IANA has assigned six new TRILL APPsub-TLV types from the range less than 255, as specified in Section 3, and updated the "TRILL APPsub-TLV Types under IS-IS TLV 251 Application Identifier 1" registry on <http://www.iana.org/assignments/trill-parameters/>, as shown below.

Type	Name of APPsub-TLV	Reference
11	Tree and VLANs	Section 3.2.1 of RFC 7968
12	Tree and VLANs Used	Section 3.2.2 of RFC 7968
13	Tree and FGLs	Section 3.2.3 of RFC 7968
14	Tree and FGLs Used	Section 3.2.4 of RFC 7968
15	Tree and Groups	Section 3.2.5 of RFC 7968
16	Tree and Groups Used	Section 3.2.6 of RFC 7968

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, DOI 10.17487/RFC6325, July 2011, <http://www.rfc-editor.org/info/rfc6325>.

[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, DOI 10.17487/RFC7172, May 2014, <http://www.rfc-editor.org/info/rfc7172>.

[RFC7176] Eastlake 3rd, D., Senevirathne, T., Ghanwani, A., Dutt, D., and A. Banerjee, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 7176, DOI 10.17487/RFC7176, May 2014, <http://www.rfc-editor.org/info/rfc7176>.

[RFC7357] Zhai, H., Hu, F., Perlman, R., Eastlake 3rd, D., and O. Stokes, "Transparent Interconnection of Lots of Links (TRILL): End Station Address Distribution Information (ESADI) Protocol", RFC 7357, DOI 10.17487/RFC7357, September 2014, <http://www.rfc-editor.org/info/rfc7357>.

[RFC7780] Eastlake 3rd, D., Zhang, M., Perlman, R., Banerjee, A., Ghanwani, A., and S. Gupta, "Transparent Interconnection of Lots of Links (TRILL): Clarifications, Corrections, and Updates", RFC 7780, DOI 10.17487/RFC7780, February 2016, <http://www.rfc-editor.org/info/rfc7780>.

7.2. Informative References

[RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <http://www.rfc-editor.org/info/rfc5310>.

Acknowledgments

The authors wish to thank David M. Bond, Liangliang Ma, Naveen Nimmu, Radia Perlman, Rakesh Kumar, Robert Sparks, Daniele Ceccarelli, and Sunny Rajagopalan for their valuable comments and contributions.

Authors' Addresses

Yizhou Li
Huawei Technologies
101 Software Avenue
Nanjing 210012
China

Phone: +86-25-56624629
Email: liyizhou@huawei.com

Donald Eastlake 3rd
Huawei Technologies
155 Beaver Street
Milford, MA 01757
United States of America

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Weiguo Hao
Huawei Technologies
101 Software Avenue
Nanjing 210012
China

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

Hao Chen
Huawei Technologies
101 Software Avenue
Nanjing 210012
China

Email: philips.chenhao@huawei.com

Somnath Chatterjee
Cisco Systems
SEZ Unit, Cessna Business Park
Outer Ring Road
Bangalore 560087
India

Email: somnath.chatterjee01@gmail.com