

TRILL Working Group  
INTERNET-DRAFT  
Intended status: Proposed Standard  
Updates: [6325](#), [6327](#)

Donald Eastlake  
Mingui Zhang  
Huawei  
Puneet Agarwal  
Broadcom  
Radia Perlman  
Intel Labs  
Dinesh Dutt  
Cumulus Networks  
December 31, 2012

Expires: June 30, 2013

TRILL: Fine-Grained Labeling  
<[draft-ietf-trill-fine-labeling-04.txt](#)>

## Abstract

The IETF has standardized TRILL (TRAnsparent Interconnection of Lots of Links), a protocol for least cost transparent frame routing in multi-hop networks with arbitrary topologies and link technologies, using link-state routing and a hop count. The TRILL base protocol standard supports labeling of TRILL data with up to 4K IDs. However, there are applications that require more fine-grained labeling of data. This document updates [RFC 6325](#) and [RFC 6327](#) by specifying optional extensions to the TRILL base protocol to safely accomplish this.

## Status of This Memo

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

Distribution of this document is unlimited. Comments should be sent to the TRILL working group mailing list.

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

INTERNET-DRAFT

TRILL: Fine-Grained Labeling

## Table of Contents

<a href="#">1. Introduction.....</a>	<a href="#">3</a>
<a href="#">1.1 Terminology.....</a>	<a href="#">3</a>
<a href="#">1.2 Contributors.....</a>	<a href="#">4</a>
<a href="#">2. Fine-Grained Labeling.....</a>	<a href="#">5</a>
<a href="#">2.1 Goals.....</a>	<a href="#">5</a>
<a href="#">2.2 Base Protocol TRILL Data Labeling.....</a>	<a href="#">6</a>
<a href="#">2.3 Fine-Grained Labeling (FGL).....</a>	<a href="#">7</a>
<a href="#">2.4 VL, Limited FGL, and Full FGL TRILL Switches.....</a>	<a href="#">8</a>
<a href="#">3. VL versus FGL Label Differences.....</a>	<a href="#">10</a>
<a href="#">4. FGL TRILL Interaction with VL TRILL.....</a>	<a href="#">11</a>
<a href="#">4.1 FGL and VL Mixed Campus Topology.....</a>	<a href="#">11</a>
<a href="#">4.2 FGL and VL Mixed Campus Characteristics.....</a>	<a href="#">12</a>
<a href="#">4.3 FGL and VL Mixed Links.....</a>	<a href="#">12</a>
<a href="#">5. FGL Details.....</a>	<a href="#">14</a>
<a href="#">5.1 Ingress Processing.....</a>	<a href="#">14</a>
<a href="#">5.2 Transit Processing.....</a>	<a href="#">15</a>
<a href="#">5.2.1 Unicast Transit Processing.....</a>	<a href="#">15</a>
<a href="#">5.2.2 Multi-Destination Transit Processing.....</a>	<a href="#">15</a>
<a href="#">5.3 Egress Processing.....</a>	<a href="#">16</a>
<a href="#">5.4 Appointed Forwarders and the DRB.....</a>	<a href="#">16</a>
<a href="#">5.5 Address Learning.....</a>	<a href="#">17</a>
<a href="#">5.6 ESADI Extensions.....</a>	<a href="#">17</a>
<a href="#">6. IS-IS Extensions.....</a>	<a href="#">18</a>
<a href="#">7. Comparison to Goals.....</a>	<a href="#">19</a>
<a href="#">8. Allocation Considerations.....</a>	<a href="#">20</a>
<a href="#">8.1 IEEE Allocation Considerations.....</a>	<a href="#">20</a>
<a href="#">8.2 IANA Considerations.....</a>	<a href="#">20</a>
<a href="#">9. Security Considerations.....</a>	<a href="#">21</a>
<a href="#">Acknowledgements.....</a>	<a href="#">22</a>
<a href="#">Normative References.....</a>	<a href="#">23</a>

Informative References.....	<a href="#">23</a>
<a href="#">Appendix A</a> : Serial Unicast.....	<a href="#">24</a>
<a href="#">Appendix Z</a> : Change History.....	<a href="#">25</a>

## [1](#). Introduction

The IETF has standardized the TRILL (TRansparent Interconnection of Lots of Links) protocol [[RFC6325](#)] that provides a solution for least cost transparent routing in multi-hop networks with arbitrary topologies and link technologies, using [[IS-IS](#)] [[RFC6165](#)] [[RFC6326bis](#)] link-state routing and a hop count. TRILL switches are sometimes called RBridges (Routing Bridges).

The TRILL base protocol standard supports labeling of TRILL data with up to 4K IDs. However, there are applications that require more fine-grained labeling of data for configurable isolation based on different tenants, service instances, or the like. This document updates [[RFC6325](#)] and [[RFC6327](#)] by specifying optional extensions to the TRILL base protocol to safely accomplish this. TRILL switches that support fine-grained labeling as specified herein have capabilities that are supersets of those specified in [[RFC6325](#)].

Familiarity with [[RFC6325](#)] and [[RFC6326bis](#)] is assumed in this document.

### [1.1](#) Terminology

The terminology and acronyms of [[RFC6325](#)] are used in this document with the additions listed below.

DEI - Drop Eligibility Indicator [[802.1Q](#)].

Edge TRILL switch - A TRILL switch announcing VLAN or Fine Grained

Label interest in its LSP.

FGL - Fine-Grained Labeling or Fine-Grained Labeled or Fine-Grained Label.

FGL TRILL switch - A TRILL switch that support both FGL and VL.  
See also "Limited FGL TRILL" and "Full FGL TRILL".

Full FGL TRILL - Capabilities that support FGL and VL ingress, transit, and egress.

Limited FGL TRILL - Capabilities that support VL ingress, transit, and egress but only FGL transit, not FGL ingress and egress.

TRILL Switch - Alternative name for an RBridge.

VL - VLAN Labeling or VLAN Labeled or VLAN Label.

VL TRILL switch - A TRILL switch that supports VL but does not support any FGL capabilities.

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](#)].

## [1.2](#) Contributors

Thanks for the contributions of the following:

Tissa Senevirathne

## [2.](#) Fine-Grained Labeling

The essence of Fine-Grained Labeling (FGL) is that (a) when TRILL Data frames are ingressed or created they may incorporate a data label from a set consisting of significantly more than 4K labels, (b) TRILL switch (RBridge) ports can be labeled with a set of such data labels, and (c) an FGL TRILL Data frame cannot be egressed through a TRILL switch port unless its fine-grained label (FGL) matches one of the data labels of the port.

[Section 2.1](#) lists FGL goals. [Section 2.2](#) briefly outlines the more coarse TRILL base protocol standard [[RFC6325](#)] data labeling. [Section 2.3](#) outlines FGL for TRILL Data frames. And [Section 2.4](#) compares VL, Limited FGL, and Full FGL TRILL switches.

## [2.1](#) Goals

There are several goals that would be desirable for FGL TRILL. They are briefly described in the list below in approximate order by priority with the most important first.

### 1. Fine-Grained

Some networks have a large number of entities that need configurable isolation, whether those entities are independent customers, applications, or branches of a single endeavor or some combination of these or other entities. The labeling supported by [\[RFC6325\]](#) provides for only  $(2^{12} - 2)$  valid identifiers or labels. A substantially larger number is required.

### 2. Silicon

Fine-grained labeling (FGL) should, to the extent practical, use existing features, processing, and fields that are already supported in many fast path silicon implementations that support the TRILL base protocol. To the extent that such silicon does not support Full FGL TRILL, it would be desirable to support Limited FGL TRILL (see [Section 2.4](#)).

### 3. Base RBridge Compatibility

To support some incremental conversion scenarios, it is desirable that not all RBridges in a campus using FGL be required to be FGL aware. That is, it is desirable if RBridges not implementing the FGL features and performing at least the transit forwarding function can usefully process TRILL Data frames that incorporate FGL.

### 4. Alternate Priority

It would be desirable for an ingress TRILL Switch to be able to assign a different priority to an FGL TRILL Data frame for its ingress-to-egress propagation from the priority of the original native frame. The original priority should be restored on egress. This enables traffic from attached non-TRILL networks to be

handled with different priority while transiting a TRILL network if desired.

## 2.2 Base Protocol TRILL Data Labeling

This section provides a brief review of the [RFC6325] TRILL Data frame internal VL Labeling and changes the description of the TRILL Header by moving its end point. This descriptive change does not involve any change in the bits on the wire or in the behavior of existing [RFC6325] VL RBridges.

VL TRILL Data frames have the structure shown below:

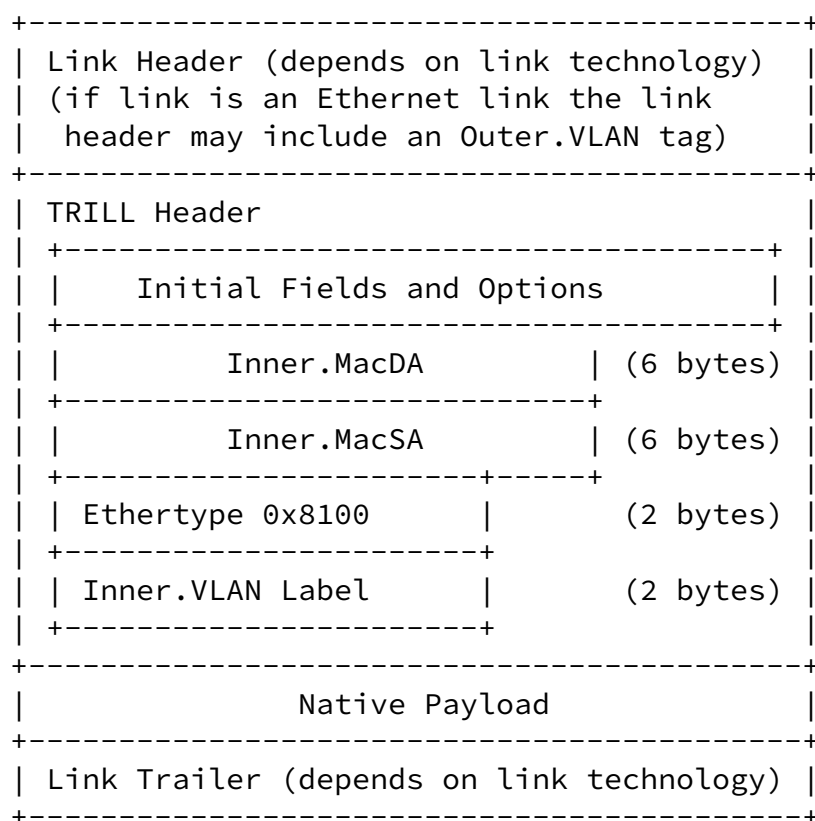


Figure 1. TRILL Data with VL

In the base protocol as specified in [RFC6325] the 0x8100 value is always present and is followed by the Inner.VLAN field which includes the 12-bit VL.

### [2.3](#) Fine-Grained Labeling (FGL)

FGL expands the variety of data labels available under the TRILL protocol to include a fine-grained label (FGL) with a 12-bit high order part and a 12-bit low order part. In this document, FGLs are usually denoted as "(X.Y)" where X is the high order part and Y is the low order part of the FGL.

FGL TRILL Data frames have the structure shown below.

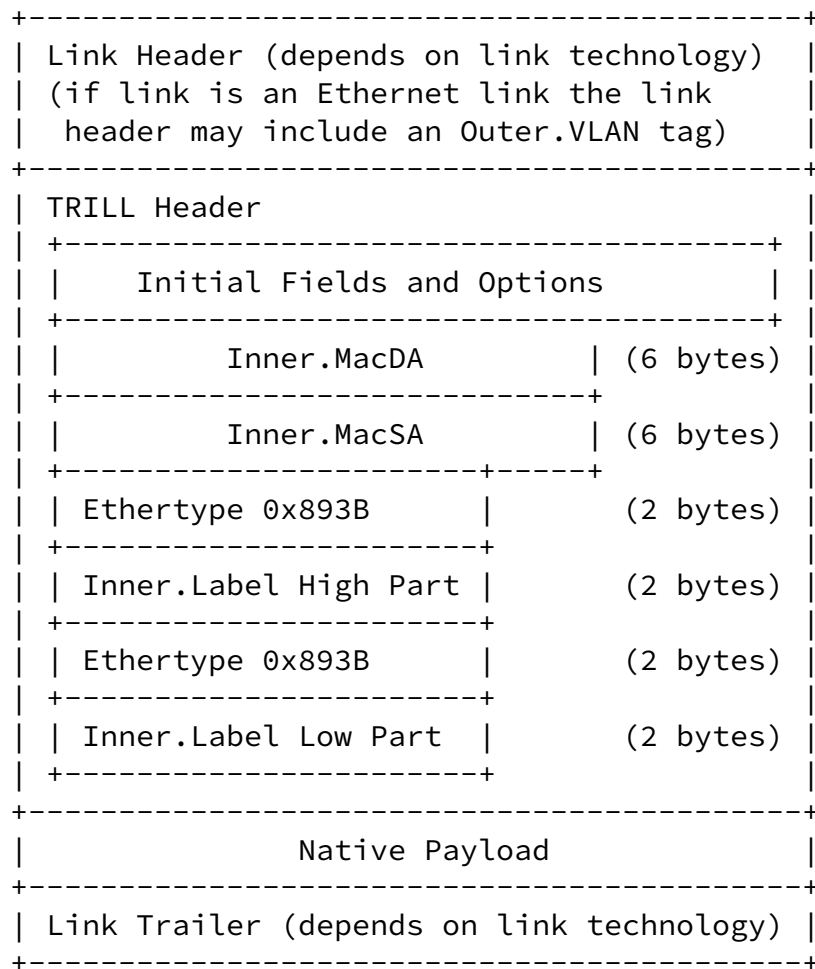


Figure 2. TRILL Data with FGL

For FGL frames, the inner MAC address fields are followed by the FGL information using 0x893B.

The two bytes following each 0x893B have, in their low order 12 bits, fine-grained label information. The upper 4 bits of those two bytes are used for a 3-bit priority field and one drop eligibility indicator (DEI) bit.

The priority field of the Inner.Label High Part is the priority used for frame transport across the TRILL campus from ingress to egress. The label bits in the Inner.Label High Part are the high order part



INTERNET-DRAFT

TRILL: Fine-Grained Labeling

of the FGL and those bits in the Inner.Label Low Part are the low order part of the FGL.

The appropriate FGL value for an ingressed or locally originated native frame is determined by the ingress TRILL switch port as specified in [Section 5.1](#).

#### [2.4](#) VL, Limited FGL, and Full FGL TRILL Switches

For a number of reasons, as listed below, it is desirable for FGL TRILL switches to be able to handle both VL and FGL TRILL Data frames.

- o Continued support of VL frames makes interoperation between VL and FGL RBridges easier as discussed in [Sections 4.1](#) and [4.2](#).
- o Due to the way TRILL works, it may be desirable to have a maintenance VLAN or FGL [[OAMframework](#)] in which all TRILL switches in the campus indicate interest. It will be simpler to use the same type of label for all TRILL switches for this purpose. That implies using VL if there might be any VL TRILL switches in the campus.
- o If a campus is being upgraded from VL to FGL TRILL switches, it avoids a requirement to immediately reconfigure all ports with FGL configuration.

The table below summarizes the differences in the capabilities of VL TRILL switches, Limited FGL TRILL switches, and Full FGL TRILL switches. As discussed in [Section 4](#), VL TRILL switches conformant to [[RFC6325](#)] should discard an FGL TRILL Data frame as malformed. FGL features are optional and VL TRILL switches are still a conformant implementation of the TRILL protocol.

TRILL Data Frame Type		VL TRILL Switch	FGL TRILL Switch	
			Limited	Full
VL		ingress	ingress	ingress
		transit	transit	transit
		egress	egress	egress

				ingress	
FGL		discard		transit	
				egress	

Figure 3. TRILL Switch Capabilities Comparison

The primary reason to specify and permit Limited FGL TRILL switches in a campus is that they might be upgraded VL TRILL switches not capable of efficiently performing FGL ingress or egress but capable of safely transiting FGL frames.

### [3.](#) VL versus FGL Label Differences

There are differences between the semantics across a TRILL campus for TRILL Data frames that are VL data labeled and those that are FGL data labeled.

With VL, data label IDs have the same meaning throughout the campus and are from the same label space as the C-VLAN IDs used on Ethernet links to end stations.

With FGL TRILL switches, many things remain the same because an FGL can appear only as the Inner.Label inside a TRILL Data frame. As such, only a TRILL-aware device will see a fine-grained label. The Outer.VLAN that may appear on native frames and that may appear on TRILL Data frames if those TRILL Data frames are on an Ethernet link can only be a C-VLAN tag. Thus ports of FGL TRILL switches, up through the usual VLAN and priority processing, act as they do for VL TRILL switches: TRILL switch ports provide a C-VLAN ID for an incoming frame and accept a C-VLAN ID for a frame being queued for output. Appointed Forwarders [[RFC6439](#)] on a link are still appointed for a C-VLAN. The Designated VLAN for an Ethernet link is still a C-VLAN.

The larger FGL data label space is a different space from the VL data label space. For ports configured for FGL, the C-VLAN on an ingressed native frame is mapped to the FGL data label space with a potentially

different mapping for each port. A similar FGL to C-VLAN mapping occurs per port on egress. Thus, for ports configured for FGL, the native frame C-VLAN on one link corresponding to an FGL can be different from the native frame C-VLAN corresponding to that same FGL on a different link elsewhere in the campus or even a different link attached to the same TRILL switch. The FGL label space is flat and does not hierarchically encode any particular number of native frame C-VLAN bits or the like. FGLs appear only inside TRILL Data frames after the inner MAC addresses.

As shown in [Section 2.4](#), FGL TRILL switches have capabilities that are a superset of those for VL TRILL switches. FGL TRILL switch ports can be configured for FGL or VL with VL being the default. As with a base protocol [\[RFC6325\]](#) TRILL switch, an unconfigured FGL TRILL switch port reports an untagged frame it receives as being in VLAN 1.

## [4.](#) FGL TRILL Interaction with VL TRILL

This section discusses various ramifications of attempting to mix FGL and VL TRILL switches in a campus. [Section 4.1](#) discusses what behaviors are needed to render such mixed campuses safe while [Section 4.2](#) discusses some resulting network characteristics. [Section 4.3](#) gives further details of link local mixed campus behavior.

### [4.1](#) FGL and VL Mixed Campus Topology

It is not possible for VL TRILL switches to safely handle FGL frames even if the VL TRILL switch is only acting in the transit capacity. This is because VL frames are required to have 0x8100 at the beginning of the data label where FGL frames have 0x893B. VL-only TRILL switches conformant to [\[RFC6325\]](#) should discard frames with this new value after the inner MAC addresses. If they do not discard

such frames, they will be confused and could egress them into the wrong VLAN (see [Section 9](#) below) or re-order them due to miscomputing flows for ECMP. Such difficulties are avoided by stopping TRILL data communication between VL and FGL TRILL switches as specified below.

FGL TRILL switches will report their FGL capability in LSPs. TRILL IS-IS communication is not affected by the blocking of TRILL data connectivity between VL and RFL Trill switches. So the link state data base will include the entire TRILL campus regardless of the presence of a mixture of VL and FGL TRILL switches. Thus FGL TRILL switches (and any management system with access to the link state database) will be able to detect the existence of TRILL switches in the campus that do not support FGL.

If both VL and FGL TRILL switches are present on a link then, although all other aspects of the adjacency machinery work as normal [[RFC6327](#)], any FGL TRILL switches on the link will not create a pseudo node for the link if they are DRB and will excommunicate the link by not announcing any adjacencies on the link. As a result, although adjacencies between two or more VL TRILL switch ports on such a link could become part of the campus topology and pass VL TRILL Data frames, no adjacency from an FGL TRILL switch port to a VL TRILL switch port or to a pseudo node will be reported on such a mixed FGL/VL link. Since an adjacency must be reported up by both ends before it becomes part of the campus topology, even though a VL TRILL switch might report an adjacency to an FGL TRILL switch, no TRILL Data can flow between an FGL and a VL TRILL switch port.

#### [4.2](#) FGL and VL Mixed Campus Characteristics

The provisions described in [Section 4.1](#) to block TRILL data communication between VL and FGL TRILL switches, although implemented only by FGL TRILL switches, have relatively symmetric effects.

A campus of mostly FGL TRILL switches with a few isolated VL TRILL switches scattered throughout will work well in terms of connectivity for end stations attached to those FGL switches except that they will be unable to communicate with any end stations for which a VL switch

is appointed forwarder. The VL TRILL switches will be isolated and will only be able to route TRILL Data to the extent they happen to be contiguously connected to other VL TRILL switches. Distribution trees computed by the FGL switches will not include any VL switches (see Section 2.1 of [[ClearCorrect](#)]). A campus of mostly VL TRILL switches with a few isolated FGL TRILL switches scattered throughout will also work reasonably well as described immediately above with all occurrences of "FGL" and "VL" swapped. However, a campus so badly misconfigured that it consists of an intermingled general mixture of VL and FGL TRILL switches is likely to offer abysmal data service.

There are possible future extensions to TRILL that would eliminate the requirement to block TRILL Data traffic between FGL and VL TRILL switches improving backwards compatibility. For example, with support for multiple topologies [[MultiTopo](#)], it would be possible to put all TRILL switches into one topology on which VL frames were sent and all FGL TRILL switches into a second topology on which FGL frames were sent. Assuming that the FGL TRILL switches are fully connected, everything would work fine including data traffic between VL ports in the same VLAN regardless of whether those VL ports were on VL or FGL TRILL switches

#### [4.3](#) FGL and VL Mixed Links

The usual DRB election operates on a link with mixed FGL and VL ports. If an FGL TRILL switch port is DRB, it MUST handle all native traffic or appoint only other FGL TRILL switch ports as Appointed Forwarder for one or more VLANs, so that all end stations will get service to the FGL campus. If a VL TRILL switch port is DRB, it will not understand that FGL TRILL switch ports are different. To the extent that a VL DRB handles native frames or appoints other VL TRILL switch ports on a link to handle native frames for one or more VLANs, the end stations sending and receiving those native frames will be isolated from the FGL campus and will receive only service from the VL campus to the extent the VL campus has connectivity. When a VL DRB happens to appoint an FGL port as Appointed Forwarder for one or more VLANs, the end stations sending and receiving native frames in those VLANs will get service to the FGL campus. This corner case behavior

is considered acceptable because it is assumed that the campus is intended to be VL or FGL and TRILL switches of the other type are

infrequent misconfigurations.

For links configured as point-to-point, if the TRILL switches at each end are both VL or both FGL, a bi-directional adjacency can be formed and reported as usual. If one is VL and one is FGL but the point-to-point link is otherwise correctly configured, the VL TRILL switch will report an adjacency but the FGL one will not. As a result, the link will not become part of the topology and TRILL Data cannot flow over the link, excommunicating the switches from each other for data.

## 5. FGL Details

This section specifies ingress, transit, egress, and other processing details with regard to FGL TRILL switches. A transit or egress FGL TRILL switch determines that a TRILL Data frame is FGL by detecting that the Inner.MacSA is followed by 0x893B.

### 5.1 Ingress Processing

It MUST be possible to configure the ports of a Full FGL TRILL switch to ingress native frames as FGL. Any ports not so configured performs the previously specified [\[RFC6325\]](#) VL ingress processing on native frames resulting in a VL TRILL Data frame. (There is no change in Appointed Forwarder logic (see [Section 5.4](#)).)

Thus Full FGL TRILL switches MUST support configurable per port mapping from the C-VLAN of a native frame, as reported by the ingress port, to an FGL. FGL TRILL switches MAY support other methods to determine the FGL of an incoming native frame, such as based on the protocol of the native frame or local knowledge.

The FGL ingress process MUST copy the priority and DEI associated with an ingressed native frame to the upper 4 bits of the Inner.Label Low Order part. It SHOULD also associate a possibly different mapped priority and DEI with an ingressed frame but a TRILL switch might not be able to do so because of implementation limitations. The mapped priority is placed in the Inner.Label High Part. If such mapping is not supported then the original priority and DEI MUST be placed in the Inner.Label High Part.

An FGL ingress TRILL switch MAY serially TRILL unicast a multi-destination TRILL Data frame to the relevant egress TRILL switches by using a known unicast TRILL Header (M=0) and SHOULD unicast such a multi-destination TRILL Data frame if there is only one relevant egress FGL TRILL switch. For FGL TRILL switches, this permits serial unicast of multi-destination frames by the ingress as an alternative to the use of a distribution tree. The relevant egress TRILL switches are determined by starting with those announcing interest in the frame's (X.Y) label. That set SHOULD be further filtered based on multicast listener and multicast router LSP announcements if the native frame was a multicast frame.

Using a TRILL unicast header for a multi-destination frame when it has only one actual destination RBridge can improve traffic spreading



and decrease latency as discussed in [Appendix A](#). How to decide whether to use a distribution tree or serial unicast for a multi-destination TRILL Data frame that has more than one destination TRILL switch is beyond the scope of this document.

## [5.2](#) Transit Processing

Any FGL TRILL switch, Limited or Full, MUST be capable of TRILL Data frame transit processing. Such processing is fairly straightforward as described in [Section 5.2.1](#) for known unicast TRILL Data frames and in [Section 5.2.2](#) for multi-destination TRILL Data frames.

### [5.2.1](#) Unicast Transit Processing

There is very little change in TRILL Data frame unicast transit processing. A transit TRILL switch forwards any unicast TRILL Data frame to the next hop towards the egress TRILL switch as specified in the TRILL Header. All transit TRILL switches, whether VL or FGL, MUST take the priority and DEI used to forward a frame from the Inner.VLAN label or the FGL Inner.Label High Part. These bits are in the same place in the frame.

An FGL TRILL switch, including a Limited FGL TRILL switch that might have limited knowledge of FGL formats, MUST properly distinguish flows if it provides ECMP for FGL frame.

### [5.2.2](#) Multi-Destination Transit Processing

Multi-destination TRILL Data frames are forwarded on a distribution tree selected by the ingress TRILL switch except that an FGL ingress TRILL switch may TRILL unicast such a frame to all relevant egress TRILL switches as described in [Section 5.1](#). The distribution trees do not distinguish between FGL and VL multi-destination frames except, possibly, in pruning behavior. All distribution trees are calculated as provided for in the TRILL base protocol standard [[RFC6325](#)] as updated by [[ClearCorrect](#)]. There is no change in the Reverse Path Forwarding Check.

An FGL TRILL switch, say RB1, having an FGL multi-destination frame

for label (X.Y) to forward on a distribution tree, SHOULD prune that tree based on whether there are any edge TRILL switches on a tree branch that are advertising connectivity to label (X.Y). In addition, RB1 SHOULD prune multicast frames based on reported multicast listener and multicast router attachment in (X.Y).

Pruning is an optimization. If a transit TRILL switch does less pruning than it could, there may be greater link utilization than strictly necessary but the campus will still operate correctly. A transit TRILL switch MAY prune based on an arbitrary subset of the bits in the FGL label, for example only the High Part or only the Low Part of the label.

### [5.3](#) Egress Processing

Egress processing is generally the reverse of ingress processing described in [Section 5.1](#).

A Full FGL TRILL switch MUST be able to covert in a configurable fashion between the FGL in an FGL TRILL Data frame it is egressing and a C-VLAN ID for the resulting native frame on a per port basis. The priority and DEI of the egressed native frame are taken from the Inner.Label Low Order Part. A port MAY be configured to strip output VLAN tagging.

It is the responsibility of the network manager to properly configure the TRILL switches in the campus to obtain the desired mappings.

An FGL TRILL switch egresses FGL frames similarly to the egressing of VL frames, as follows:

1. A known unicast FGL TRILL Data frame is egressed to the FGL port or ports matching its FGL and Inner.MacDA. If there are no such ports, it is flooded out all FGL ports that have its FGL except any ports for which the TRILL switch has knowledge that the frame's Inner.MacDA cannot be present on the link out that port.
2. A multi-destination FGL frame is decapsulated and flooded out all ports with its FGL, subject to multicast pruning.

An FGL TRILL switch, including a Limited FGL TRILL switch that might have limited knowledge of FGL formats, MUST NOT egress an FGL frame

with label (X.Y) to any port not configured with that label even if the port is configured to egress VL frames in VLAN X.

FGL TRILL switches MUST accept multi-destination TRILL Data frames that are sent to them as TRILL unicast frames, that is, frames that may have a multicast or broadcast Inner.MacDA (or are being sent to an unknown unicast Inner.MacDA) and the TRILL Header M bit set to 0. They locally egress such frames, if appropriate, but MUST NOT forward them (other than egressing them as native frames on their local links).

#### [5.4](#) Appointed Forwarders and the DRB

There is no change in Adjacency [[RFC6327](#)] or Appointed Forwarder logic [[RFC6439](#)] on a link regardless of whether some or all the ports on the link are for FGL TRILL switches except for the refusal of FGL TRILL switches to report adjacencies on links with mixed VL and FGL TRILL switches, as described in [Section 4](#) above.

#### [5.5](#) Address Learning

An FGL TRILL switch learns addresses on ports configured for FGL based on the fine-grained label rather than the native frame's VLAN. Addresses learned from ingressed native frames on FGL ports are logically represented by { MAC address, FGL, port, confidence, timer } while remote addresses learned from egressing FGL frames are logically represented by { MAC address, FGL, remote TRILL switch nickname, confidence, timer }.

#### [5.6](#) ESADI Extensions

The TRILL ESADI (End Station Address Distribution Information) protocol is specified in [[RFC6325](#)] as optionally transmitting MAC address connection information through TRILL Data frames between participating TRILL switches over the virtual link provided by the TRILL multicast frame distribution mechanism. In [[RFC6325](#)], the VL to which an ESADI frame applies is indicated only by the Inner.VLAN label and no indication of that VL is allowed within the ESADI payload.

ESADI is extended to support FGL by providing for the indication of the FGL to which an ESADI frame applies only in the Inner.Label of that frame and no indication of that FGL is allowed within the ESADI payload.

## [6.](#) IS-IS Extensions

Extensions to the TRILL use of IS-IS are required to support FGL include the following:

1. An method for a TRILL switch to announce itself in its LSP as supporting FGL (see [Section 8.2](#)).
2. A sub-TLV analogous to Interested VLANs and Spanning Tree Roots sub-TLV of the Router Capabilities TLV but indicating FGLs rather than VLs. This is called the Interested Labels and Spanning Tree Roots sub-TLV in [rfc6326bis].
3. Sub-TLVs analogous to the GMAC-ADDR sub-TLV of the Group

Address TLV that specifies an FGL rather than a VL. These are called the GLMAC-ADDR, GLIP-ADDR, and GLIP6 ADDR sub-TLVs in [rfc6326bis].

## 7. Comparison to Goals

Comparing TRILL FGL, as specified in this document, with the goals given in [Section 2.1](#), we find as follows:

1. Fine-Grained: FGL provides  $2^{24}$  labels, vastly more than the 4K

VL labels provided in [\[RFC6325\]](#).

2. Silicon: Existing TRILL fast path silicon chips can perform base TRILL Header insertion and removal to support ingress and egress. In addition, it is believed that most such silicon can also perform the native frame to FGL mapping and the encoding of the FGL as specified herein, as well as the inverse decoding and mapping. Some existing silicon can perform only one of these operations on a frame in one pass through the fast path and so is likely to be suitable as a full speed Limited FGL TRILL switch or a reduced rate Full FGL TRILL switch requiring two passes for ingress/egress; however, other existing chips are believed to be able to perform both operations on the same frame in one pass through their fast path and are thus suitable for full rate Full FGL TRILL processing.
3. Base RBridge Compatibility: As described in [Section 3](#), FGL is not generally compatible with TRILL switches conformant to the base specification [\[RFC6325\]](#) although, as described in [Section 4.3](#), there are possible future TRILL extensions that would enable data intercommunication in some topologies. In particular, an FGL-ignorant TRILL switch cannot be even a Limited FGL TRILL switch in a restricted topology because there is a risk that it would receive multi-destination FGL frames and egress them to the wrong VLAN. However, a contiguous set of VL TRILL switches can exchange VL frames regardless of the presence of FGL TRILL switches in the campus.
4. Alternate Priority: The encoding specified in [Section 2.3](#) and the ingress/egress processing specified in [Section 5](#) provide for a new priority and DEI in the Inner.Label High Part and a place to preserve the original user priority and DEI in the Low Part, so it can be restored on egress.

## [8.](#) Allocation Considerations

Allocations by the IEEE Registration Authority and IANA are listed below.

### [8.1](#) IEEE Allocation Considerations

The IEEE Registration Authority has assigned Ethertype 0x893B for use as the TRILL FGL Ethertype.

### [8.2](#) IANA Considerations

IANA is requested to allocate capability bit TBD in the TRILL-VER sub-TLV capability bits [[RFC6326bis](#)] to indicate a TRILL switch is FGL-capable (either Limited FGL or Full FGL).

INTERNET-DRAFT

TRILL: Fine-Grained Labeling

## 9. Security Considerations

See [[RFC6325](#)] for general TRILL Security Considerations.

As with any communications system, end-to-end encryption and authentication should be considered for sensitive data. In this case that would be encryption and authentication extending from a source end station to a destination end station.

Confusion between a frame with VL X and FGL (X.Y) or confusion due to a malformed frame is a potential problem if an FGL TRILL switch did not properly check for the occurrence of 0x8100 or 0x893B immediately after the Inner.MacSA (see Sections [2.2](#) and [2.3](#)) and handled the frame appropriately.

[RFC6325] requires that the Ethertype immediately after the Inner.MacSA be 0x8100. A VL TRILL switch that did not discard a frame with some other value there could cause problems. If it received a TRILL Data frame with FGL (X.Y) or with junk after the Inner.MacSA that included X where a VLAN ID would appear, then:

1. It could egress the frame to an end station in VLAN-X. If the frame was a well formed FGL frame, the payload of such an egressed native frame would appear to begin with Ethertype 0x893B that would likely be discarded by an end station. In any case, such an egress would almost certainly be a violation of security policy requiring the configurable separation of differently labeled data.
2. If the frame was multi-destination and the TRILL switch pruned the distribution tree, it would incorrectly prune it on the basis of VLAN-X. For an FGL frame, this would probably lead to the multi-destination data frame not being delivered to all of its intended recipients.

Possible problems with an FGL TRILL switch that received a TRILL Data frame with junk after the Inner.MacSA that included X where a VLAN ID would appear and did not check the Ethertype immediately after the Inner.MacSA would be that it could improperly egress the frame in VLAN-X, violating security policy. If the frame was multi-destination and was improperly forwarded, it should be discarded by properly



implemented TRILL switches downstream in the distribution tree and never egressed but the propagation of the frame would still waste bandwidth.

To avoid these problems all TRILL switches MUST check the Ethertype immediately after the Inner.MacSA and, if it is a value they do not know how to handle, either discard the frame or make no decisions based on any data after that Ethertype. In addition, care must be taken to avoid FGL frames being sent to or through VL TRILL switches

that will discard them if the VL TRILL switch is properly implemented or mishandle them if it is not properly implemented. This is accomplished, as described in [Section 4](#), by FGL TRILL switches not announcing adjacency to VL TRILL switches. As a result, no TRILL data frames can be exchanged between VL and FGL TRILL switches and they will be isolated from each other for data purposes.

#### Acknowledgements

The comments and suggestions of the following are gratefully acknowledged:

Anoop Ghanwani, Sujay Gupta, Weiguo Hao, Jon Hudson, Phanidhar Koganti, Yizhou Li, Vishwas Manral, Rajeev Manur, Thomas Narten, Gayle Nobel, Erik Nordmark, Olen Stokes, Ilya Varlashkin, and Xuxiaohu.

The document was prepared in raw nroff. All macros used were defined within the source file.

INTERNET-DRAFT

TRILL: Fine-Grained Labeling

#### Normative References

- [IS-IS] - ISO/IEC 10589:2002, Second Edition, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", 2002.
- [802.1Q] - IEEE 802.1, "IEEE Standard for Local and metropolitan area networks - Virtual Bridged Local Area Networks", IEEE Std 802.1Q-2011, May 2011.
- [RFC2119] - Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC6325] - Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", [RFC 6325](#), July 2011.
- [RFC6327] - Eastlake 3rd, D., Perlman, R., Ghanwani, A., Dutt, D., and V. Manral, "Routing Bridges (RBridges): Adjacency", [RFC 6327](#), July 2011.
- [RFC6326bis] - Eastlake, D., Banerjee, A., Dutt, D., Perlman, R., and A. Ghanwani, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", [draft-ietf-isis-rfc6326bis](#), Work in

Progress.

[ClearCorrect] - D. Eastlake, M. Zhang, A. Ghanwani, A. Banerjee, V. Manral, [draft-ietf-trill-clear-correct-06.txt](#), in RFC Editor's queue.

## Informative References

[MultiTopo]

- [draft-tissa-trill-mt-encode](#), Work in Progress.
- [draft-eastlake-trill-rbridge-multi-topo](#), Work in Progress.

[OAMframework] - [draft-ietf-trill-oam-framework](#), Work in Progress.

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

[RFC6439] - Perlman, R., Eastlake, D., Li, Y., Banerjee, A., and F. Hu, "Routing Bridges (RBridges): Appointed Forwarders", [RFC 6439](#), November 2011.

## Appendix A: Serial Unicast

This appendix discusses advantages and disadvantages of using serial unicast instead of a distribution tree for multi-destination TRILL Data frames. See Sections [5.1](#) and [5.3](#).

Consider a large TRILL campus with hundreds of TRILL switches in which, say, 300 end stations are in some particular FGL data label.

At one extreme, if all 300 end stations were on links attached to a single TRILL switch, then no other TRILL switch would be advertising interest in that FGL and likely a multi-destination (say broadcast) frame from one such end station would, even if put on a distribution tree, because of pruning, not get sent to any another TRILL switch.

At the other extreme, assume the 300 end stations are attached, one each, to 300 different TRILL switches; in that case you are almost

certainly better off using a distribution tree because if you tried to serially unicast you would probably have to output multiple copies through the same port and would cause much higher link utilization.

Now assume these 300 end stations are connected to exactly two TRILL switches, say 200 to one and 100 to the other. Using unicast TRILL Data frames between these two TRILL switches is best because the frames will follow least cost paths, possibly with such traffic spread over a number of equal cost least cost paths. On the other hand, if a distribution trees were used, each frame would be constrained to the tree used for that frame and would likely follow a more circuitous route, depending on where the tree's root is, and only a single path would be available per tree. Thus this document says that unicast "SHOULD" be used if there are exactly two TRILL switches involved.

It is a more complex decision whether to use a distribution tree or serial unicast if the end stations are connected to a number of TRILL switches greater than two. Which would be better would depend on many factors including network topology and application data patterns. How to make this decision in such more complex cases is beyond the scope of this document.

## Appendix Z: Change History

RFC Editor Note: Please delete this appendix before publication.

From -00 to -01:

Update author info and make editorial changes.

From -01 to -02

1. Change the value after the inner MAC addresses for FGL frames from 0x8100 to 0x893B
2. As a consequence of item 1 above, for safety prohibit use for TRILL Data of links between FGL and VL RBridges, isolating any VL RBridges. Make appropriate changes throughout document, including Security Considerations section, based on this change.
3. Reference and contributor updates.
4. Minor editorial changes.

From -02 to -03

1. Addition of the terms "Limited FGL" and "Full FGL".
2. Addition of [Appendix A](#).
3. Clarifications:
  - 3.a That FGL TRILL switches also support VL ports and frames (Add [Section 2.4](#), etc.).
  - 3.b That the FGL extensions to TRILL are optional. A VL TRILL switch is still a conformant implementation.
  - 3.c The utility of the alternate priority goal.
4. Expand Security Considerations discussion of misparsed frames.
5. Substantial editorial changes.

From -03 to -04

1. Typo and grammar fixes.
2. Update acknowledgements, date, and version as usual.

Donald Eastlake 3rd  
Huawei Technologies  
155 Beaver Street  
Milford, MA 01757 USA

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

Mingui Zhang  
Huawei Technologies Co., Ltd  
Huawei Building, No.156 Beiqing Rd.  
Z-park, Shi-Chuang-Ke-Ji-Shi-Fan-Yuan, Hai-Dian District,  
Beijing 100095 P.R. China

Email: zhangmingui@huawei.com

Puneet Agarwal  
Broadcom Corporation  
3151 Zanker Road  
San Jose, CA 95134 USA

Phone: +1-949-926-5000  
Email: pagarwal@broadcom.com

Radia Perlman  
Intel Labs  
2200 Mission College Blvd.  
Santa Clara, CA 95054 USA

Phone: +1-408-765-8080  
Email: Radia@alum.mit.edu

Dinesh G. Dutt  
Cumulus Networks  
1089 West Evelyn Avenue  
Sunnyvale, CA 94086 USA

Email: ddutt.ietf@hobbesdutt.com

INTERNET-DRAFT

TRILL: Fine-Grained Labeling

## Copyright, Disclaimer, and Additional IPR Provisions

Copyright (c) 2012 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. The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions. For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of [RFC 5378](#). No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under [RFC 5378](#), shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

