Rbridges: Base Protocol Specification

Status of This Document

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Distribution of this document is unlimited. Comments should be sent to the TRILL working group mailing list <rbridge@postel.org>.

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.
Abstract

RBridges provide optimal pair-wise forwarding with zero configuration, safe forwarding even during periods of temporary loops, and support for multipathing of both unicast and multicast traffic. They achieve these goals using IS-IS routing and encapsulation of traffic with a header that includes a hop count.

RBridges are compatible with previous IEEE 802.1 customer bridges as well as IPv4 and IPv6 routers and end nodes. They are as invisible to current IP routers as bridges are and, like routers, they terminate the bridge spanning tree protocol.

The design supports VLANs and optimization of the distribution of multi-destination frames based on VLAN and IP derived multicast groups. It also allows forwarding tables to be sized according to the number of RBridges (rather than the number of end nodes), which allows internal forwarding tables to be substantially smaller than in conventional bridges.

Acknowledgements

Many people have contributed to this design, including, in alphabetic order, Alia Atlas, Ayan Banerjee, Suresh Boddapati, Caitlin Bestler, Stewart Bryant, James Carlson, Dino Farinacci, Don Fedyk, Bill Fenner, Eric Gray, Joel Halpern, Andrew Lange, David Melman, Erik Nordmark, Sanjay Sane, Pekka Savola, Matthew Thomas, Joe Touch, and Mark Townsley. We invite you to join the mailing list at http://www.postel.org/rbridge.
## Table of Contents Continued

4.2.4  TRILL ESADI IS-IS..................................40
4.2.4.1  TRILL ESADI Participation........................40
4.2.4.2  TRILL ESADI Information..........................41
4.3  Distribution Trees...................................41
4.3.1  Distribution Tree Calculation and Checks.............42
4.3.2  Pruning the Distribution Tree.........................43
4.3.3  Tree Distribution Optimization........................44
4.3.4  Forwarding Using a Distribution Tree..................45
4.4  Frame Processing Behavior...............................46
4.4.1  Receipt of a Native Frame.............................46
4.4.1.1  Native Unicast Case...............................46
4.4.1.2  Native Multicast and Broadcast Frames...............47
4.4.2  Receipt of a TRILL Frame................................48
4.4.2.1  TRILL IS-IS Frames................................48
4.4.2.2  TRILL ESADI Frames................................48
4.4.2.3  TRILL Data Frames..................................49
4.4.3  Receipt of a Control Frame.............................50
4.5  IGMP, MLD, and MRD Learning.............................50
4.6  End Station Address Details...............................51
4.6.1  Learning End Station Addresses.........................52
4.6.2  Forgetting End Station Addresses......................53
4.6.3  Shared VLAN Learning..................................54
4.7  RBridge Ports...........................................55
4.7.1  RBridge Port Configuration.............................55
4.7.2  RBridge Port Structure................................56
4.7.3  BPDU Handling.........................................58
4.7.3.1  Receipt of BPDUs....................................59
4.7.3.2  Root Bridge Changes.................................59
4.7.3.3  Transmission of BPDUs................................59
4.7.4  Dynamic VLAN Registration.............................60

5.  RBridge Addresses, Parameters, and Constants............61

6.  Security Considerations..................................63
6.1  VLAN Security Considerations............................63

7.  Assignment Considerations................................65
7.1  IANA Considerations......................................65
7.2  IEEE Registration Authority Considerations..............65

8.  Normative References.....................................66

9.  Informative References....................................67

Appendix A: Incremental Deployment Considerations...........69
A.1  Link Cost Determination..................................69
A.2  Appointed Forwarders and Bridged LANs...................69
A.3  Wiring Closet Topology...................................71
Table of Contents Continued

A.3.1 The RBridge Solution.................................72
A.3.2 The VLAN Solution..................................72
A.3.3 The Spanning Tree Solution............................72
A.3.4 Comparison of Solutions...............................73

Appendix B: Trunk and Access Port Configuration.........74

Appendix C: Multipathing..................................75

Appendix D: Determination of VLAN and Priority............77

Appendix Z: Revision History...............................78
Changes from -03 to -04..................................78
Changes from -04 to -05..................................79
Changes from -05 to -06..................................80
Changes from -06 to -07..................................80
Changes from -07 to -08..................................81
Changes from -08 to -09..................................83
Changes from -09 to -10..................................84

Disclaimer................................................85
Additional IPR Provisions..................................85

Authors’ Addresses........................................86
Table of Figures

Figure 2.1: Interconnected RBridges.......................14
Figure 2.2: An Ethernet Encapsulated TRILL Frame.........14
Figure 2.3: A PPP Encapsulated TRILL Frame...............15
Figure 2.4: RBridge Port Model............................19
Figure 3.1: TRILL Header................................21
Figure 3.2: Options Area Initial Flags Octet..............23
Figure 4.1: TRILL Data Encapsulation over Ethernet......28
Figure 4.2: VLAN Tag Information..........................29
Figure 4.3: Detailed RBridge Port Model....................57
Figure A.1: Link Cost of a Bridged Link..................69
Figure A.2: Wiring Closet Topology.........................71
Figure C.1: Multi-Destination Multipath....................75
Figure C.2: Known Unicast Multipath.......................76
1. Introduction

In traditional IPv4 and IPv6 networks, each subnet has a unique prefix. Therefore, a node in multiple subnets has multiple IP addresses, typically one per interface. This also means that when an interface moves from one subnet to another, it changes its IP address. Administration of IP networks is complicated because IP routers require significant configuration. Careful IP address management is required to avoid creating subnets that are sparsely populated, wasting addresses.

IEEE 802.1 bridges avoid these problems by transparently gluing many physical links into what appears to IP to be a single LAN [802.1D]. However, 802.1 bridge forwarding using the spanning tree protocol has some disadvantages:

- The spanning tree protocol blocks ports, limiting the number of forwarding links, and therefore creates bottlenecks by concentrating traffic onto selected links.
- Forwarding is not pair-wise shortest path, but is instead whatever path remains after the spanning tree eliminates redundant paths.
- The Ethernet header does not contain a hop count (or TTL) field. This is dangerous when there are temporary loops such as when spanning tree messages are lost or components such as repeaters are added.
- VLANs can partition when spanning tree reconfigures due to a node failure or topology change.

This document presents the design for RBridges (Routing Bridges [RBriges]) which implement the TRILL protocol and are poetically summarized below. Rbridges combine the advantages of bridges and routers and in most cases they can incrementally replace IEEE [802.1Q] or [802.1D] customer bridges. While they can be applied to a variety of link protocols, this specification focuses on IEEE [802.3] links.

For further discussion of the problem domain addressed by RBridges see [PAS].
1.1 Algorhyme V2, by Ray Perlner

I hope that we shall one day see
A graph more lovely than a tree.

A graph to boost efficiency
While still configuration-free.

A network where RBridges can
Route packets to their target LAN.

The paths they find, to our elation,
Are least cost paths to destination!

With packet hop counts we now see,
The network need not be loop-free!

RBridges work transparently.
Without a common spanning tree.

1.2 Normative Content and Precedence

The bulk of the normative material in this specification appears in Sections 2, 3, and 4 as follows:

- Section 2: general RBridge description
- Section 3: the TRILL header
- Section 4: other TRILL protocol details

In case of conflict, the order of precedence of these sections is as follows, with those appearing earlier in this list having precedence over those which appear later:

4 > 3 > 2

1.3 Terminology and Notation in this document

"TRILL" normally refers to the protocol specified herein while "RBridge" refers to the devices that implement that protocol. The second letter in Rbridge is case insensitive. Both Rbridge and RBridge are correct.

This document uses Hexadecimal Notation for MAC addresses. Each octet (that is, 8-bit byte) is represented by two hexadecimal digits giving the value of the octet as an unsigned integer and successive octets are separated by a hyphen. This document consistently uses
IETF bit ordering although the physical order of bit transmission within an octet on an IEEE [802.3] link is from the lowest order bit to the highest order bit, the reverse.

In this document, Layer 2 frames are divided into three categories, TRILL frames, control frames, and native frames, as follows:

- "control" frames are those with a multicast destination address in the range 01-80-C2-00-00-00 to 01-80-C2-00-00-0F or equal to 01-80-C2-00-00-21. There are two sub-categories of control frames as follows:
  - "high level control" frames are those with a destination address of 01-80-C2-00-00-00 (BPDU) or 01-80-C2-00-00-21 (VRP).
  - "low level control" frames are all other control frames.

- "TRILL" frames are those (1) with a multicast destination address allocated to the TRILL protocol (see Section 7.2) and (2) non-control frames with the TRILL Ethertype. There are three sub-categories of TRILL frames. RBridges do not generate and silently discard on receipt any TRILL frames which do not match one of these sub-categories as follows:
  - "TRILL IS-IS" frames have the All-IS-IS-RBridges multicast address as their destination address.
  - "TRILL ESADI" frames have the TRILL Ethertype and the All-ESADI-RBridges multicast address as the encapsulated destination address.
  - "TRILL data" frames have the TRILL Ethertype but are not TRILL ESADI frames.

- "native" frames are all frames other than TRILL and control frames.

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].

1.4 Acronyms

AllL1ISs - All Level 1 Intermediate Systems

AllL2ISs - All Level 2 Intermediate Systems

BPDU - Bridge PDU

CHbH - Critical Hop-by-Hop

CItE - Critical Ingress-to-Egress
CSNP - Complete Sequence Number PDU
DA - Destination Address
EAP - Extensible Authentication Protocol
ECMP - Equal Cost Multi-Path
EISS - Extended Internal Sublayer Service
ESADI - End Station Address Distribution Instance
DRB - Designated RBridge
FCS - Frame Check Sequence
GARP - Generic Attribute Registration Protocol
GVRP - GARP VLAN Registration Protocol
IEEE - Institute of Electrical and Electronics Engineers
IGMP - Internet Group Management Protocol
IP - Internet Protocol
IS-IS - Intermediate System to Intermediate System
ISS - Internal Sublayer Service
LAN - Local Area Network
LSP - Link State PDU
MAC - Media Access Control
MLD - Multicast Listener Discovery
MRD - Multicast Router Discovery
MVRP - Multiple VLAN Registration Protocol
NSAP - Network Service Access Point
PDU - Protocol Data Unit
PPP - Point-to-Point Protocol
RBridge - Routing Bridge
RPF - Reverse Path Forwarding
SA - Source Address
SPF - Shortest Path First
TLV - Type, Length, Value
TRILL - TRansparent Interconnection of Lots of Links
VLAN - Virtual Local Area Network
VRP - VLAN Registration Protocol
2. RBridges

This section provides a high level overview of RBridges, which implement the TRILL protocol. Sections 3 and 4 below provide the main specification.

RBridges run a link state protocol amongst themselves. This gives them enough information to compute pair-wise optimal paths for unicast, and calculate distribution trees for delivery of frames either to unknown MAC destinations or to multicast/broadcast groups.

To mitigate temporary loop issues, RBridges forward based on a header with a hop count. RBridges also specify the next hop RBridge as the frame destination when forwarding unicast frames across a shared-media link, which avoids spawning additional copies of frames during a temporary loop. A Reverse Path Forwarding Check and other checks are performed on multi-destination frames to further control potentially looping traffic (see Section 4.3.1).

The first RBridge that a unicast frame encounters in a campus, RB1, encapsulates the received frame with a TRILL header that specifies the last RBridge, RB2. RB1 is known as the "ingress RBridge" and RB2 is known as the "egress RBridge". To save room in the TRILL header, a dynamic nickname acquisition protocol is run among the RBridges to select a 2-octet nickname for each RBridge, unique within the campus, which is an abbreviation for the 6-octet IS-IS system ID of the RBridge. The 2-octet nicknames are used to specify the ingress and egress RBridges in the TRILL header.

Multipathing of multi-destination frames through alternative distribution tree roots and ECMP (Equal Cost MultiPath) of unicast frames are supported (see Appendix C).

RBridges run the IS-IS [ISO10589] election protocol to elect a "Designated RBridge" (DRB) on each bridged LAN ("link"). As with an IS-IS router, the DRB may give a pseudonode name to the link, issue an LSP (Link State PDU) on behalf of the pseudonode, and issues CSNPs (Complete Sequence Number PDUs) on the link. Additionally, the DRB specifies which VLAN will be the Designated VLAN used for communication between RBridges on that link.

The DRB either encapsulates/decapsulates all data traffic to/from the link, or, for load splitting, delegates this responsibility, for one or more VLANs, to other RBridges on the link. There must at all times be at most one RBridge on the link that encapsulates/decapsulates traffic for a particular VLAN. We will refer to the RBridge appointed to forward VLAN-x traffic on behalf of the link as the "appointed VLAN-x forwarder". (Section 2.3 discusses VLANs further.)
Rbridges SHOULD support SNMPv3 [RFC3411]. The Rbridge MIB will be specified in a separate document. If IP service is available to an RBridge, it SHOULD support SNMPv3 over IP; however, management can be used, within a campus, even by an RBridge that lacks an IP or other Layer 3 transport stack or which has zero configuration and thus no Layer 3 address, by transporting SNMP with Ethernet [RFC4789].

2.1 End Station Addresses

An RBridge, RBl, which is the VLAN-x forwarder on any of its links MUST learn the location of VLAN-x end nodes, both on the links for which it is VLAN-x forwarder, and on other links in the campus. RBl learns the port and Layer 2 (MAC) addresses of end nodes on links for which it is VLAN-x forwarder from the source address of frames received, as bridges do (for example, see section 8.7 of [802.1Q]), or through a Layer 2 explicit registration protocol such as IEEE 802.11 association and authentication. RBl learns the Layer 2 address of distant VLAN-x end nodes, and the corresponding RBridge to which they are attached, by looking at the ingress RBridge nickname in the TRILL header and the VLAN and source address of the inner frame of TRILL data frames that it decapsulates.

Additionally, an end station address distribution instance (ESADI) of IS-IS MAY be used by an RBridge that is the appointed VLAN-x forwarder on one or more links to announce some or all of the attached VLAN-x end nodes on those links. An ESADI could be used to announce end nodes that have been explicitly enrolled. Such information might be more authoritative than learning from data frames being decapsulated onto the link. Also, it can be more secure because not only might the enrollment be authenticated (for example by cryptographically based EAP methods via [802.1X]), but IS-IS also supports cryptographic authentication of its messages [RFC5304]. But even if an ESADI is used to announce attached end nodes, Rbridges MUST still learn from decapsulating data frames unless configured not to do so.

Advertising end nodes using an ESADI of IS-IS is optional, as is learning from these announcements.

(See Section 4.6 for further end station address details.)

2.2 RBridge Encapsulation Architecture

The Layer 2 technology used to connect Rbridges may be either IEEE [802.3] or some other technology such as PPP [RFC1661]. This is possible since the RBridge relay functionality is layered on top of
the Layer 2 technologies. However, this document specifies only an IEEE 802.3 encapsulation.

Figure 2.1 shows two R Bridges RB1 and RB2 interconnected through an Ethernet cloud. The Ethernet cloud may include point-to-point or shared media, hubs and IEEE 802.1D or 802.1Q bridges.

---

Figure 2.1: Interconnected R Bridges

Figure 2.2 shows the format of a TRILL data or ESADI frame traveling through the Ethernet cloud from RB1 to RB2.

---

Figure 2.2: An Ethernet Encapsulated TRILL Frame

In the case of media different from Ethernet, the outer Ethernet header is replaced by the header specific to that media. For example, Figure 2.3 shows a TRILL encapsulation over PPP.
The outer header is link-specific and, although this document specifies only Ethernet links, other links are allowed.

In both cases the Inner Ethernet Header and the Ethernet Payload come from the original frame and are encapsulated with a TRILL Header as they travel between RBridges. Use of a TRILL header offers the following benefits:

1. loop mitigation through use of a hop count field;
2. elimination of the need for original source and destination MAC address learning in transit RBridges;
3. direction of frames towards the egress RBridge (this enables forwarding tables of RBridges to be sized with the number of RBridges rather than the total number of end nodes); and,
4. provision of a separate VLAN tag for forwarding traffic between RBridges, independent of the VLAN of the native frame.

When forwarding unicast frames between RBridges across a shared-media, the outer header has the MAC destination address of the next hop Rbridge, to avoid frame duplication. Having the outer header specify the transmitting RBridge as source address ensures that any bridges inside the Ethernet cloud will not get confused, as they might be if multipathing is in use and they were to see the original source or ingress RBridge in the outer header.

From a forwarding standpoint, transit frames may be classified into two main categories: known-unicast and multi-destination.
2.2.1 Known-Unicast

These frames have a unicast inner MAC destination Address (Inner.MacDA) and are those for which the egress RBridge for that destination MAC address is known to the ingress RBridge.

Such frames are forwarded Rbridge hop by Rbridge hop to their egress Rbridge.

2.2.2 Multi-destination

These are frames that must be delivered to multiple destinations.

Multi-destination frames include the following:

1. unicast frames for which the destination is unknown: the Inner.MacDA is unicast, but the ingress RBridge does not know its location;

2. multicast frames for which the Layer 2 destination address is derived from an IP multicast address: the Inner.MacDA is multicast, from the set of Layer 2 multicast addresses derived from IPv4 [RFC1112] or IPv6 [RFC2464] multicast addresses; these frames are handled somewhat differently in different subcases:
   2.1 IGMP [RFC3376] and MLD [RFC2710] multicast group membership reports;
   2.2 IGMP [RFC3376] and MLD [RFC2710] queries and MRD [RFC4286] announcement messages;
   2.3 other IP derived Layer 2 multicast frames;

3. multicast frames for Layer 2 destination address is not derived from an IP multicast address: the Inner.MacDA is multicast, and not from the set of Layer 2 multicast addresses derived from IPv4 or IPv6 multicast addresses.

4. broadcast frames: the Inner.MacDA is broadcast (FF-FF-FF-FF-FF-FF).

RBridges build distribution trees (see Section 4.3) and use these trees for forwarding multi-destination frames. These distribution trees are pruned in different ways for different cases to avoid unnecessary propagation of the frame.
2.3 RBridges and VLANs

A VLAN is a way to partition end nodes in a campus into different Layer 2 communities [802.1Q]. Use of VLANs requires configuration. The default method of determining the VLAN of a frame sent by an end station is based on the port on which it is initially received. End stations can also explicitly insert this information in a frame.

IEEE 802.1Q bridges can be configured to support multiple VLANs over a single link by inserting/removing a VLAN tag in the frame. VLAN tags used by TRILL have the same format as VLAN tags defined in IEEE [802.1Q]. As shown in Figure 2.2 there are two places where such tags may be present in a TRILL-encapsulated frame sent over an IEEE 802.3 link: one in the outer header (Outer.VLAN) and one in the inner header (Inner.VLAN). Inner and outer VLANs are further discussed in Section 4.1.

RBridges enforce delivery of a native frame originating in a particular VLAN only to other links in the same VLAN; however, there are a few differences in the handling of VLANs between an RBridge campus and an 802.1 bridged LAN as described below.

(See Section 4.2.3.1 for further discussion of TRILL IS-IS operation on a link beyond that in the subsections below.)

2.3.1 Link VLAN Assumptions

In a network with a mix of bridges and RBridges, certain configurations of bridges or ports may prevent some nodes in some VLANs from communicating with each other.

TRILL requires that on each link in the campus there is at least one VLAN that gives full connectivity to all the RBridges on that link and that the RBridges are configured to know which VLAN that is if it is not the default VLAN 1.

Since there will be only one appointed forwarder for any VLAN, say VLAN A, on a link, if bridges are configured to cause VLAN A to be partitioned on a link, some end nodes on that link may be orphaned (unable to communicate with the rest of the campus).

It is possible for bridge and port configuration to cause VLAN mapping on a link (where a VLAN A frame turns into a VLAN B frame). TRILL detects this case through IS-IS Hellos, by inserting the initial VLAN tag into the Hello. TRILL includes mechanisms to detect VLAN mapping within a link and takes steps to ensure that there is at most a single appointed forwarder on the link, to avoid possible frame duplication or loops.
TRILL behaves as conservatively as possible, avoiding loops rather than avoiding partial connectivity. As a result, lack of connectivity may result from bridge or port misconfiguration.

2.4 R Bridges and IEEE 802.1 Bridges

As described below, RBridge ports are, for the most part, layered on top of IEEE [802.1Q] port facilities and RBridges can be incrementally deployed into an existing bridged LAN.

2.4.1 RBridge and 802.1 Layering

RBridges make use of 802.1Q port VLAN processing and lower level 802.1 protocols as well as the protocols for the link in use, such as port based access control [802.1X] or link aggregation (Clause 43 of [802.3]). The only exceptions are those protocols related to high level control frames including spanning tree. RBridge do not use spanning tree and do not block ports in the way that spanning tree blocks ports. (There may in the future be additional lower level 802.1 protocols that require different handling in an RBridge than in an 802.1 bridge.) Figure 2.4 shows a high level diagram of an RBridge port connected to an IEEE 802.3 link. Single lines represent the flow of control information, double lines the flow of frames and control information.
The upper interface to the lower level port/link control logic corresponds to the Internal Sublayer Service (ISS) in [802.1Q]. In R Bridges, high level control frames are processed above the ISS interface.

The upper interface to the port VLAN processing corresponds to the Extended Internal Sublayer Service (EISS) in [802.1Q]. In R Bridges, all native and TRILL frames are processed above the EISS interface and are subject to port VLAN and priority processing.
2.4.2 Incremental Deployment

Because RBridges are generally compatible with current IEEE [802.1Q] customer bridges, a bridged LAN can be upgraded by incrementally replacing such bridges with RBridges. Bridges that have not yet been replaced are transparent to RBridge traffic. The physical links directly interconnected by such bridges, together with the bridges themselves, constitute bridged LANs. These bridged LANs appear to RBridges to be multi-access links. If the bridges replaced by RBridges were unmanaged, zero configuration bridges, then their RBridge replacements will not require configuration.

The RBridge campus will work best if all IEEE 802.1 bridges are replaced with RBridges, assuming the RBridges have the same speed and capacity as the bridges. However, there may be intermediate states, where only some bridges have been replaced by RBridges, with inferior performance.

See Appendix A for further discussion of incremental deployment.
3. Details of the TRILL Header

The section specifies the TRILL header. Section 4 below provides other RBridge design details.

3.1 TRILL Header Format

The TRILL header is shown in Figure 3.1 and is independent of the data link layer used. When that layer is IEEE [802.3], it is prefixed with the 16-bit TRILL Ethertype [RFC5342], making it 64 bit aligned.

```
+------------------+-+------------------+-
| V | R | M | Op-Length | Hop Count |
+------------------+-+------------------+-
| Egress RBridge Nickname | Ingress RBridge Nickname |
+------------------+-+------------------+-
```

Figure 3.1: TRILL Header

The header contains the following fields which are described in the sections referenced:

- V (Version): 2-bit unsigned integer. See Section 3.2.
- R (Reserved): 2 bits. See Section 3.3.
- M (Multi-destination): 1 bit. See Section 3.4.
- Op-Length (Options Length): 5-bit unsigned integer. See Section 3.5.
- Egress RBridge Nickname: 16-bit identifier. See Section 3.7.1.
- Ingress RBridge Nickname: 16-bit identifier. See Section 3.7.2.

3.2 Version (V)

A single Ethertype is granted to a protocol and, under IEEE guidelines, it is the protocol’s responsibility to structure itself to support future revisions. Adhering to this guideline, there is a two bit Version field in the TRILL header. Version zero of TRILL is specified in this document. An RBridge that sees a message with a Version value it does not understand MUST silently discard the message because it may not be able to parse it.
3.3 Reserved (R)

The two R bits are reserved for future use in extensions to this version zero of the TRILL protocol. They MUST be zero on transmission and are ignored on receipt.

3.4 Multi-destination (M)

The Multi-destination bit (see Section 2.2.2) indicates that the frame is to be delivered to a class of destination end stations via a distribution tree and that the egress RBridge nickname field specifies the root RBridge for this tree. In particular:

- M = 0 (FALSE) - The egress RBridge nickname contains the nickname of the egress RBridge for a known unicast TRILL data frame;
- M = 1 (TRUE) - The egress RBridge nickname field contains the nickname of the RBridge that is the root of a distribution tree. This tree is selected by the ingress RBridge for a TRILL data frame or by the source RBridge for a TRILL ESADI (end station address distribution instance) frame.

3.5 TRILL Header Options

The TRILL Protocol includes an option capability in the TRILL Header. The Op-Length header field gives the length of the options in units of 4 octets which allows up to 124 octets of options area. If Op-Length is zero there are no options present. If options are present, they follow immediately after the Ingress RBridge Nickname field.

All Rbridges MUST be able to skip the number of 4-octet chunks indicated by the Op-Length field in order to find the inner frame, since Rbridges must be able to find the destination MAC address and VLAN tag in the inner frame. (Transit Rbridges need such information to filter VLANs, IP multicast, and the like. Egress Rbridges need to find the inner header to correctly decapsulate and handle the inner frame.)

To ensure backward compatible safe operation, when Op-Length is non-zero indicating that options are present, the top two bits of the first octet of the options area are specified as follows:
If the CHbH (Critical Hop by Hop) bit is one, one or more critical hop-by-hop options are present so transit RBridges that support no options MUST drop the frame. If the CHbH bit is zero, the frame is safe, from the point of view of options processing, for a transit RBridge to forward the frame even if it doesn’t understand the options. A transit RBridge that supports no options and forwards a frame MUST transparently forward the options area.

If the CItE (Critical Ingress to Egress) bit is a one, one or more critical ingress-to-egress options are present. If it is zero, no such options are present. If either CHbH or CItE is non-zero, egress RBridges that support no options MUST drop the frame. If both CHbH and CItE are zero, the frame is safe, from the point of view of options, for any egress RBridge to process even if it doesn’t understand options.

Options will be further specified in other documents and are expected to include provisions for hop-by-hop and ingress-to-egress options as well as critical and non-critical options.

Note: Most RBridges implementations are expected to be optimized for the simplest and most common cases of frame forwarding and processing. The inclusion of any options may, and the inclusion of complex or lengthy options almost certainly will, cause frame processing using a "slow path" with markedly inferior performance to "fast path" processing. Limited slow path throughput may cause such frames to be lost.

3.6 Hop Count

The Hop Count field is a 6-bit unsigned integer. Each RBridge that is about to forward a frame to another RBridge MUST check this field and discard the frame if this field is zero. If this field is greater than or equal to 1, it MUST be decremented in the forwarded frame.

For known unicast frames, the ingress RBridge MUST set the Hop Count to at least the number of RBridge hops it expects to the egress RBridge and SHOULD set it in excess of that number to allow for alternate routing later in the path.

For multi-destination frames, the Hop Count MUST be set by the ingress RBridge (or source RBridge for an end station address...
distribution TRILL IS-IS frame) to at least the expected number of hops to the most distant RBridge. To accomplish this, RBridge RBn calculates, for each branch from RBn of the distribution tree rooted at RBi, the maximum number of hops in that branch. When forwarding a multi-destination frame onto a branch, transit RBridge RBm MAY decrease the hop count by more than 1 unless decreasing the hop count by more than 1 would result in a Hop Count insufficient to reach all destinations in that branch of the tree rooted at RBi. Using a Hop Count close to the minimum on multi-destination frames reduces potential problems with temporary loops when forwarding.

Although the RBridge MAY decrease the hop count by more than 1, under the circumstances described above, the RBridge forwarding a frame MUST decrease the hop count by at least 1, and discards the frame if it cannot do so because the hop count is 0.

3.7 RBridge Nicknames

Nicknames are 16-bit dynamically assigned abbreviations for each RBridge’s 48-bit IS-IS System ID to achieve a more compact encoding. This assignment allows specifying up to 2**16 RBridges; however, the value 0x0000 is reserved to indicate that a nickname is not specified, the values 0xFFC0 through 0xFFFF are reserved for future specification, and the value 0xFFFF is permanently reserved. RBridges piggyback a nickname acquisition protocol on the link state protocol (see Section 3.7.3) to acquire a nickname unique within the campus.

3.7.1 Egress RBridge Nickname

There are two cases for the contents of the egress RBridge nickname field, depending on the M-bit (see Section 3.4). It is filled in by the ingress RBridge for TRILL data frames and by the source RBridge for TRILL ESADI frames.

- For known unicast TRILL data frames, M == 0 and the egress RBridge nickname field specifies the egress RBridge i.e. it specifies the RBridge that needs to remove the TRILL encapsulation and forward the native frame. Once the egress nickname field is set, it MUST NOT be changed by any subsequent transit RBridge.

- For multi-destination TRILL data frames and for TRILL ESADI (end station address distribution instance) frames, M == 1. The egress RBridge nickname field contains the nickname of the root RBridge of the distribution tree selected to be used to forward the frame. This root MUST NOT be changed by transit RBridges.
3.7.2 Ingress RBridge Nickname

The ingress RBridge nickname is set to the nickname of the ingress RBridge for all TRILL data frames and to the nickname of the source RBridge for all TRILL ESADI frames.

Once the ingress nickname field is set, it MUST NOT be changed by any subsequent transit RBridge.

3.7.3 RBridge Nickname Selection

The nickname selection protocol is piggybacked on the core TRILL IS-IS instance as follows:

- The nickname being used by an RBridge is carried in an IS-IS TLV (type-length-value data element) along with a priority of use value. Each RBridge chooses its own nickname.

- The nickname value MAY be configured. An RBridge that has been configured with a nickname value will have priority for that nickname value over all Rbridges with non-configured nicknames.

- The nickname values 0x0000 and 0xFFC0 through 0xFFFF are reserved and MUST NOT be selected by or configured for an RBridge. The value 0x0000 is used to indicate that a nickname is not known.

- The priority of use field reported with a nickname is an unsigned 8-bit value, where the most significant bit (0x80) indicates that the nickname value was configured. The bottom 7 bits have the default value 0x40, but MAY be configured to be some other value. Additionally, an RBridge MAY increase its priority after holding the nickname for some amount of time. However, the most significant bit of the priority MUST NOT be set unless the nickname value was configured.

- Once an RBridge has successfully acquired a nickname it SHOULD attempt to reuse it in the case of a reboot.

- Each RBridge is responsible for ensuring that its nickname is unique. If RB1 chooses nickname x, and RB1 discovers, through receipt of RB2’s LSP, that RB2 has also chosen x, then the RBridge with the numerically higher priority keeps the nickname, or if there is a tie in priority, the RBridge with the numerically higher IS-IS System ID keeps the nickname, and the other RBridge MUST select a new nickname. This can require an RBridge with a configured nickname to select a different nickname.

- To minimize the probability of nickname collisions, when an
RBridge selects a new nickname, it does so by randomly hashing some of its parameters, e.g., interface MAC addresses, time and date, and other entropy sources such as those given in [RFC4086]. There is no reason for all Rbridges to use the same algorithm for selecting nicknames.

If two RBridge campuses merge, then transient nickname collisions are possible. As soon as each RBridge receives the LSPs from the other RBrignes, the RBrignes that need to change nicknames select new nicknames that do not, to the best of their knowledge, collide with any existing nicknames. Some RBrignes may need to change their nickname more than once before the situation is resolved.

To minimize the probability of a new RBridge usurping a nickname already in use, an RBridge whose nickname is not configured SHOULD wait to acquire the link state database from a neighbor before it announces its own nickname.

An RBridge that will not act as an ingress, egress, or tree root need not have a nickname.
4. Other RBridge Design Details

Section 3 above specifies the TRILL header, while this Section specifies other RBridge design details.

4.1 Ethernet Data Encapsulation

TRILL data and ESADI frames in transit on Ethernet links are encapsulated with an outer Ethernet header (see Figure 2.2). This outer header looks, to a bridge on the path between two RBridges, like the header of a regular Ethernet frame and therefore bridges forward the frame as they normally would. To enable RBridges to distinguish such TRILL frames, a new TRILL Ethertype (see Section 7.2) is used in the outer header.

Figure 4.1 details a TRILL data frame with an outer VLAN tag traveling on an Ethernet link as shown at the top of the Figure, that is, between transit RBridges RB3 and RB4. The native frame originated at end station ESa, was encapsulated by ingress RBridge RB1 and will ultimately be decapsulated by egress RBridge RB2 and delivered to destination end station ESb. The encapsulation shown has the advantage, in the absence of TRILL options, of aligning the original Ethernet frame at a 64-bit boundary.

When a TRILL data frame is carried over an Ethernet cloud it has three pairs of addresses:

- Outer Ethernet Header: Outer Destination MAC Address (Outer.MacDA) and Outer Source MAC Address (Outer.MacSA): These addresses are used to specify the next hop RBridge and the transmitting RBridge, respectively.

- TRILL Header: Egress Nickname and Ingress Nickname. These specify the nickname values of the egress and ingress RBridges, respectively, unless the frame is multi-destination in which case the Egress Nickname specified the root of the distribution tree on which the frame is being sent.

- Inner Ethernet Header: Inner Destination MAC Address (Inner.MacDA) and Inner Source MAC Address (Inner.MacSA): These addresses are as transmitted by the original end station, specifying, respectively, the destination and source of the inner frame.

A TRILL data frame also potentially has two VLAN tags that can carry two different VLAN Identifiers and specify priority.
Flow:

| ESa | RB1 | RB3 | RB4 | RB2 | ESb |

| ingress | transit | transit | egress | transit |

Outer Ethernet Header:

<table>
<thead>
<tr>
<th>Outer Destination MAC Address (RB4)</th>
<th>Outer Destination MAC Address</th>
<th>Outer Source MAC Address</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ethertype = C-Tag [802.1Q]</td>
<td>Outer.VLAN Tag Information</td>
</tr>
</tbody>
</table>

TRILL Header:

<table>
<thead>
<tr>
<th>Ethertype = TRILL</th>
<th>V</th>
<th>R</th>
<th>M</th>
<th>Op-Length</th>
<th>Hop Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egress (RB2) Nickname</td>
<td>Ingress (RB1) Nickname</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inner Ethernet Header:

<table>
<thead>
<tr>
<th>Inner Destination MAC Address (ESb)</th>
<th>Inner Destination MAC Address</th>
<th>Inner Source MAC Address</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ethertype = C-Tag [802.1Q]</td>
<td>Inner.VLAN Tag Information</td>
</tr>
</tbody>
</table>

Payload:

<table>
<thead>
<tr>
<th>Ethertype of Original Payload</th>
<th>Original Ethernet Payload</th>
</tr>
</thead>
</table>

Frame Check Sequence:

| New FCS (Frame Check Sequence) |

Figure 4.1: TRILL Data Encapsulation over Ethernet
4.1.1 VLAN Tag Information

A "VLAN Tag", also known as a "C-tag" (formerly known as a Q-tag) for customer tag, includes a VLAN ID and a priority field as shown in Figure 4.2. The "VLAN ID" may be zero, indicating the no VLAN is specified, just a priority, although such frames are called "priority tagged" rather than a "VLAN tagged" [802.1Q].

(802.1Q S-tags or service tags are beyond the scope of this document.)

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Priority | C |                  VLAN ID                      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

Figure 4.2: VLAN Tag Information

As recommended in [802.1Q], RBridges SHOULD be implemented so as to allow use of the full range of VLAN IDs from 0x001 through 0xFFE. VLAN ID zero is the null VLAN identifier and indicates that no VLAN is specified while VLAN ID 0xFFF is reserved. RBridges MAY support a smaller number of simultaneously active VLAN IDs than the total number of different VLAN IDs they allow.

The VLAN ID 0xFFF is reserved and MUST NOT be used. RBridges MUST discard any frame they receive with an Outer.VLAN ID of 0xFFF. RBridges MUST discard any frame for which they examine the Inner.VLAN ID and find it to be 0xFFF; such examination is required at all egress RBridges which decapsulate a frame.

The "C" bit shown in Figure 4.2 is not used in TRILL. It MUST be set to zero and is ignored by receivers.

As specified in [802.1Q], the priority field contains an unsigned value from 0 through 7 where 1 indicates the lowest priority, 7 the highest priority, and the default priority zero is considered to be higher than priority 1 but lower than priority 2. The [802.1ad] amendment to [802.1Q] permits mapping some adjacent pairs of priority levels into a single priority level with and without drop eligibility. RBridges MAY also implement such configuration options. RBridges are not required to implement any particular number of distinct priority levels but may treat one or more adjacent priority levels in the same fashion.

Frames with the same source address, destination address, VLAN, and priority that are received on the same port as each other and are transmitted on the same port MUST be transmitted in the order received. (Such frames might not be sent out the same port if multipath is implemented. See Appendix C.) Differing priorities can cause frame re-ordering.
The C-Tag Ethertype [RFC5342] is 0x8100.

4.1.2 Inner VLAN Tag

The "Inner VLAN Tag Information" (Inner.VLAN) field contains the VLAN tag information associated with the native frame or the VLAN tag information associated with a TRILL ESADI (end station address distribution instance) frame when that frame was created. When a TRILL frame passes through a transit RBridge, the Inner.VLAN MUST NOT be changed except when VLAN mapping is being intentionally performed within that RBridge.

When a native frame arrives at an RBridge, the associated VLAN ID and priority are determined as specified in [802.1Q] (see Appendix D and [802.1Q] Section 6.7). If the RBridge is an appointed forwarder for that VLAN and the delivery of the frame requires transmission to one or more other links, this ingress RBridge forms a TRILL data frame with the associated VLAN ID and priority placed in the Inner.VLAN Information. Thus, in TRILL data frames, the Inner.VLAN tag always specifies a VLAN ID.

The VLAN ID is required at the ingress RBridge as one element in determining the appropriate egress RBridge for a known unicast frame and is needed at the ingress and every transit RBridge for multi-destination frames to correctly prune the distribution tree.

4.1.3 Outer VLAN Tag

TRILL frames sent by an RBridge, except for some TRILL IS-IS Hellos, use an Outer.VLAN ID specified by the Designated RBridge (DRB) for the link onto which they are being sent, referred to as the Designated VLAN. For TRILL data and ESADI frames, the priority in the Outer.VLAN tag SHOULD be set to the priority in the Inner.VLAN tag.

TRILL frames forwarded by a transit RBridge use the priority present in the Inner.VLAN of the frame as received. TRILL data frames are sent with the priority associated with the corresponding native frame when (see Appendix D). TRILL IS-IS frames SHOULD be sent with priority 7.

Whether an Outer.VLAN tag actually appears on the wire when a TRILL frame is sent depends on the configuration of the RBridge port through which it is sent in the same way as the appearance of a VLAN tag on a frame sent by an [802.1Q] frame depends on the configuration of the bridge port (see Section 4.7.2).
4.1.4 Frame Check Sequence (FCS)

Each Ethernet frame has a single Frame Check Sequence (FCS) that is computed to cover the entire frame, for detecting frame corruption due to bit errors on the link. Any received frame for which the FCS check fails MUST be discarded. The FCS normally changes on encapsulation, decapsulation, and every TRILL hop due to changes in the outer destination and source addresses, the decrementing of the hop count, etc.

Although the FCS is normally calculated just before transmission, it is desirable, when practical, for an FCS to accompany a frame within an RBridge after receipt. That FCS could then be dynamically updated to account for changes to the frame during Rbridge processing and used for transmission or checked against the FCS calculated for frame transmission. This optional, more continuous use of an FCS would be helpful in detecting a class of internal RBridge failures such as memory errors.

4.2 Link State Protocol (IS-IS)

TRILL uses an extension of IS-IS [ISO10589] as the routing protocol since it has the following advantages:

- it runs directly over Layer 2, so therefore may be run with zero configuration (no IP addresses need to be assigned);
- it is easy to extend by defining new TLV (type-length-value) data elements and sub-elements for carrying TRILL information;

4.2.1 IS-IS RBridge Identity

Each RBridge has a unique unsigned 48-bit (6-octet) IS-IS System ID. This ID may be derived from any of the RBridge’s unique MAC addresses.

A pseudonode is assigned a 7-octet ID by the DRB that created it, usually by taking a 6-octet ID owned by the DRB, and appending another octet. The only constraint is that the 7-octet ID be unique within the campus, and that the 7th octet be nonzero. An RBridge has a 7-octet ID consisting of its 6-octet system ID concatenated with a zero octet.

In this document we use the term "IS-IS ID" to refer to the 7-octet quantity that can either be the ID of an RBridge or a pseudonode.
4.2.2 IS-IS Instances

TRILL implements separate IS-IS instances from any used by Layer 3, that is, different from the one used by routers. Layer 3 IS-IS frames must be distinguished from TRILL IS-IS frames even when those Layer 3 IS-IS frames are transiting an RBridge campus.

Layer 3 IS-IS native frames have a special multicast destination address, either AllL1ISs or AllL2ISs. When they are TRILL encapsulated, these multicast addresses appear as the Inner.MacDA and the Outer.MacDA will be either unicast or the All-RBridges multicast address.

Within TRILL, there is a mandatory core IS-IS instance across all Rbridges in the campus as described in Section 4.2.3. This core instance uses TRILL IS-IS frames which are distinguished by having a multicast destination address of All-IS-IS-RBridges. TRILL IS-IS frames have the IS-IS NSAP protocol type and do not have a TRILL Header.

In addition, there can be optional end station address distribution instances (ESADIs) between the Rbridges on each supported VLAN as described in Section 4.2.4. They are similar to TRILL data frames where the encapsulated frame is an IS-IS protocol frame but are distinguished by the presence of an Inner.MacDA of All-ESADI-RBridges.

4.2.3 Core TRILL IS-IS

All Rbridges must participate in the core TRILL IS-IS instance. Core instance frames are never forwarded by an RBridge but are locally processed on receipt. (Such processing may cause the RBridge to send additional TRILL IS-IS frames.)

4.2.3.1 Core IS-IS Link Protocol

Rbridges send TRILL IS-IS Hello frames on a link in order to discover RBridge neighbors. As with Layer 3 IS-IS, one RBridge is elected DRB (Designated RBridge), based on configured priority (most significant field), and system ID. The DRB, as described in Section 4.2.3.2, designates the VLAN to be used on the link for inter-RBridge communication and appoints itself or other Rbridges on the link as appointed forwarder (see Section 4.2.3.3) for VLANs on the link.
4.2.3.1.1 Core Hello VLAN Tagging

By default, RBridges tag TRILL IS-IS Hello frames with VLAN 1. Because a link may be a bridged LAN with different connectivity for different VLANs, and since an RBridge may be configured so that it cannot use VLAN 1 on a port, Hellos may need to be sent out a port with additional and/or other VLANs for connectivity and safety.

An RBridge RBn maintains for each port the same VLAN information as a customer IEEE [802.1Q] bridge, including the set of VLANs enabled for output through that port (see Section 4.7.2). In addition, RBn maintains the following TRILL specific VLAN parameters per port:

a) Desired Designated VLAN: the VLAN that RBn, if it is DRB, will specify in its Hellos to be used by all RBridges on the link to communicate all TRILL frames, except some Hellos as described below. This MUST be a VLAN enabled on RBn’s port. It defaults to the numerically lowest enabled VLAN ID, which is VLAN 1 for a zero configuration RBridge.

b) Designated VLAN: the VLAN being used on the link for all TRILL frames except some Hellos. This is RBn’s Desired Designated VLAN if RBn believes it is the DRB or the Designated VLAN in the DRB’s Hellos if RBn is not the DRB.

c) Announcing VLANs set. This defaults to the enabled VLANs set on the port but may be configured to be a subset of the enabled VLANs.

d) Forwarding VLANs set: the set of VLANs for which an RBridge port is appointed VLAN forwarder on the port. This MUST only contain enabled VLANs for the port, possibly all enabled VLANs.

On each of its ports an RBridge sends Hellos Outer.VLAN tagged with each VLAN in a set of VLANs. For each port, this set depends on the RBridge’s DRB status and the above VLAN parameters. All RBridges send Hellos Outer.VLAN tagged with the Designated VLAN, unless that VLAN is not enabled. In addition, the DRB sends Hellos Outer.VLAN tagged with each enabled VLAN in its Announcing VLANs set. All non-DRB RBridges send Hellos Outer.VLAN tagged with all enabled VLANs that are in the intersection of their Forwarding VLANs set and their Announcing VLANs set. More symbolically, Hellos are sent as follows:

If it is DRB

intersection ( Enabled VLANs,
    union ( Designated VLAN, Announcing VLANs ) )

If it is not DRB

intersection ( Enabled VLANs,
    union ( Designated VLAN,
Configuring the Announcing VLANs set to be null minimizes the number of Hellos. In that case, Hellos are only tagged with the Designated VLAN.

The number of Hellos is maximized, within this specification, by configuring the Announcing VLANs set to be the set of all enabled VLAN IDs, which is the default. In that case, the DRB will send Hellos tagged with all its Enabled VLAN tags and any non-DRB RBridge RBn will send Hellos tagged with the Designated VLAN, if enabled, and tagged with all VLANs for which RBn is an appointed forwarder. (It is possible to send even more Hellos. In particular, non-DRB RBridges could send Hellos on enabled VLANs for which they are not an appointed forwarder and which are not the Designated VLAN. This would not cause harm, other than a further communications and processing burden.)

When an RBridge port comes up, until it has heard a Hello from a higher priority RBridge, it considers itself to be DRB on that port and sends Hellos on that basis. Similarly, even if it has at some time recognized some other RBridge on the link as DRB, if on that port it receives no Hellos from an RBridge with higher priority as DRB for a long enough time, as specified by IS-IS, it will revert to believing itself DRB. Note that an RBridge RBn does not defer to the DRB listed in a Hello, even if that claimed DRB is higher priority, if the Hello was sent by an RBridge with lower priority than RBn.

4.2.3.1.2 TRILL IS-IS Hello Contents

A TRILL IS-IS Hello includes the following information, in addition to the standard IS-IS Hello header information. The actual encoding of this information and the IS-IS Type or sub-Type values for the TLV or sub-TLV data elements are specified in a separate document.

1. The VLAN ID of the Designated VLAN for the link.

2. In connection with VLAN mapping (see Section 4.2.3.1.3):
   2.a A copy of the Outer.VLAN ID with which the Hello was tagged.
   2.b A flag which, if set, indicates that the sender has detected VLAN mapping on the link, within the past five Hello times.

3. The set of VLANs for which end station service is enabled on the port. If this is missing or null, it implies that the port is configured as a trunk port (see Section 4.7.1). This MAY be omitted if the sender is DRB.

4. A flag which, if set, indicates that the sender believes it is
appointed forwarder for the VLAN and port on which the Hello was sent.

5. A flag which, if set, indicates that the sender’s port was configured as an access port. The value of this flag for the DRB controls and when it is asserted by the DRB all ports on that link which recognize the DRB act as access ports.

6. If the sender is DRB, the Rbridges (including itself) that it appoints as forwarders for that link and the VLANs for which it appoints them.

7. TRILL connectivity over which TRILL data, ESADI, and non-Hello TRILL IS-IS frames will be sent is only established on the Designated VLAN. Establishing such connectivity requires exchange of Hellos containing the IS Neighbor TLV, so that TLV MUST be included in Hellos sent on the Designated VLAN. The Neighbor TLV MAY be send on other VLANs but neighbor status is only established and updated based on Hellos on the Designated VLAN.

It is anticipated that many links between Rbridges will be point-to-point in which case a pseudonode merely adds to the complexity. If the DRB specifies the pseudonode ID as all zeros, this indicates that the Rbridges on the link are just to report their adjacencies as point-to-point. This has no effect on how LSPs are flooded on a link. It only affects what LSPs are generated.

For example, if RB1 and RB2 are the only Rbridges on the link and RB1 is DRB, then if RB1 creates a pseudonode there are 3 LSPs: for, say, RB1.25 (the pseudonode), RB1, and RB2, where RB1.25 reports connectivity to RB1 and RB2, and RB1 and RB2 each just say they are connected to RB1.25. Whereas if DRB R1 declares no pseudonode, then there will be only 2 LSPs: RB1 and RB2 each reporting connectivity to each other.

A DRB SHOULD NOT create a pseudonode for its link unless it has seen at least two simultaneous adjacencies on the link at some point since it last re-booted or in certain cases for access ports (see Section 4.7.1).

4.2.3.1.3 VLAN Mapping Within a Link

IEEE [802.1Q] does not provide for bridges changing the C-tag VLAN ID for a tagged frame they receive, that is, mapping VLANs. Nevertheless, some bridge products provide this capability and, in any case, bridged LANs can be configured to display this behavior. For example, a bridge port can be configured to strip certain VLAN tags on output and send the resulting untagged frames onto a link leading to another bridge’s port configured to tag these frames with a different VLAN. Although each port’s configuration is legal under
in the aggregate they perform manipulations not permitted within a single customer 802.1Q bridge. Since RBridge ports have the same VLAN capabilities as customer 802.1Q bridges, this can occur even in the absence of bridges. (VLAN mapping is referred to in IEEE 802 as VID translation.)

R Bridges include the Outer.VLAN ID inside a TLV within each Hello message. When a Hello is received, they compare this saved copy with the Outer.VLAN ID information associated with the received frame. If these differ and the VLAN ID inside the Hello is X and the Outer.VLAN is Y, it can be assumed that VLAN ID X is being mapped into VLAN ID Y.

When a VLAN mapping is detected, the RBridge detecting it sets a flag in all Hellos it sends on the link for the subsequent five Hellos times. This notifies the DRB if the detecting RBridge is not the DRB. The DRB then assures that only one RBridge (either the DRB itself or some RBridge it appoints) is appointed forwarder for any VLANs on the link. This avoids loops and duplication of frames with different VLAN tags.

4.2.3.2 Designated RBridge

TRILL IS-IS elects one RBridge for each link to be the Designated RBridge (DRB), that is, to have special duties. The Designated RBridge:

- Chooses, for the link, and announces in its Hellos, the Designated VLAN ID to be used for inter-RBridge communication. This VLAN is used for all TRILL data and ESADI frames and all non-Hello TRILL IS-IS frames. TRILL IS-IS Hellos are sent on this VLAN but are usually also sent on others (see Section 4.2.3.1.1).

- If the link is represented in the IS-IS topology as a pseudonode, chooses a pseudonode ID and announces that in its Hello and issues an LSP on behalf of the pseudonode.

- Issues CSNPs.

- For each VLAN-x appearing on the link, chooses an RBridge on the link to be the appointed VLAN-x forwarder (the DRB MAY choose itself to be the appointed VLAN-x forwarder for all or some of the VLANs).

- Before appointing a VLAN-x forwarder (including appointing itself), wait at least 5 Hello intervals (to ensure it is DRB).
o Continues sending IS-IS Hellos on all its enabled VLANs that have been configured in the Announcing VLANs set of the DRB, which defaults to all enabled VLANs.

4.2.3.3 Appointed VLAN-x Forwarder

The appointed VLAN-x forwarder for a link is responsible for the following:

o Loop avoidance:
  - Inhibiting itself for a configurable time from 30 to zero seconds, which defaults to 30 seconds, after it sees a root bridge change on the link (see Section 4.7.3.2). An inhibited appointed forwarder for a port drops any native frames it receives and does not transmit and native frames in the VLAN for which it is appointed.
  - Inhibiting itself, as described above, for VLAN-x if, within the past five hello times, it has received a Hello on VLAN-x in which the sender asserts that it is appointed forwarder for VLAN-x.
  - Optionally, not decapsulating a frame from ingress RBridge RBm unless it has RBm's LSP, and the root bridge on the link it is about to forward onto is not listed in RBm's list of root bridges for VLAN-x. This is known as the "decapsulation check" or "root bridge collision check".

o Unless inhibited (see above), receiving VLAN-x native traffic from the link and, as appropriate, forwarding it in native and/or encapsulated form.

o Receiving VLAN-x traffic for the link and, if uninhibited, transmitting it in native form after decapsulating it as appropriate.

o Learning the MAC address of local VLAN-x nodes by looking at the source address of VLAN-x frames from the link.

o Optionally learning the port of local VLAN-x nodes based on any sort of Layer 2 registration protocols such as IEEE 802.11 association and authentication.

o Keeping track of the { egress RBridge, VLAN, MAC address } of distant VLAN-x end nodes, learned by looking at the fields { ingress RBridge, Inner.VLAN ID, Inner.MacSA } from VLAN-x frames being received for decapsulation onto the link.
o Optionally observe native IGMP [RFC3376], MLD [RFC2710], and MRD [RFC4286] frames to learn the presence of local multicast listeners and multicast routers.

o Optionally listening to the messages in the TRILL IS-IS end station address distribution instance (ESADI) for VLAN-x to learn { egress RBridge, VLAN-x, MAC address } triplets and the confidence level of such explicitly advertised end nodes.

o Optionally advertising VLAN-x end nodes, on links for which it is appointed VLAN-x forwarder, in a TRILL IS-IS ESADI.

o Listening to BPDUs on the common spanning tree to learn the root bridge, if any, for that link and to report in its LSP the complete set of root bridges seen on any of its links for which it is appointed forwarder for VLAN-x.

o Including a "port number" in its Hellos, and if it sees its own Hello on port p, where the port number in the received Hello is "q", then if q>p, not forwarding traffic to/from port p, as already provided in IS-IS.

4.2.3.4 Core TRILL IS-IS LSP Information

The information in the TRILL IS-IS LSP for the mandatory core instance is listed below. The actual encoding of this information and the IS-IS Type or sub-Type values for any new IS-IS TLV or sub-TLV data elements are specified in a separate document.

1. The IS-IS IDs of neighbors (pseudonodes as well as RBridges) of RBridge RBn, and the cost of the link to each of those neighbors except for neighbors reached through certain access ports (see Section 4.7.1).

2. If an RBn is to be able to act as an ingress, egress, or tree root, the nickname of RBridge RBn (2 octets) and the unsigned 8-bit priority for RBn to have that nickname (see Section 3.7.3).

3. The TRILL Header Versions supported by RBridge RBn (4 bits).

4. The priority of RBn for becoming a distribution tree root and the number of additional distribution trees it wants computed for the campus (each 16 bits, see Section 4.3).

5. The list of RBridge nicknames that RBn might select for a distribution tree root when RBn injects a multi-destination frame into the campus. These tree roots MUST be from the set of roots for the distribution trees which all RBridges in the campus are
computing (see Section 4.3). Using this field RBridges can efficiently build receipt filters to avoid multicast loops (see Section 4.3.1). If the list is empty or not provided, RBn MUST only select the highest priority distribution tree root for the campus.

6. The list of VLAN IDs of VLANs directly connected to RBn for links on which RBn is the appointed forwarder for that VLAN. (Note: an RBridge may advertise that it is connected to additional VLANs in order to receive additional frames to support certain VLAN based features beyond the scope of this specification as mentioned in Section 4.6.3.) In addition, the LSP contains the following information on a per-VLAN basis:

6.1 Per VLAN Multicast Router attached flags: This is two bits of information that indicate whether there is an IPv4 and/or IPv6 multicast router attached to the Rbridge on that VLAN. An RBridge which does not do IP multicast control snooping MUST set both of these bits (see Section 4.3.3). This information is used because IGMP [RFC3376] and MLD [RFC2710] Membership Reports MUST be transmitted to all links with IP multicast routers, and SHOULD NOT be transmitted to links without such routers. Also, all frames for IP-derived multicast addresses MUST be transmitted to all links with IP multicast routers (within a VLAN), in addition to links from which an IP node has explicitly asked to join the group the frame is for, except for some IP multicast addresses that MUST be treated as broadcast.

6.2 Per VLAN Other Multicast flag. This is a flag bit that indicates that the RBridge wishes to receive non-IP derived multicast for that VLAN. It defaults to true (one). Within each VLAN, all non-IP derived multicast traffic MUST be sent to an RBridge that asserts this flag.

6.3 Per VLAN mandatory announcement of the set of IDs of Root bridges for any of RBn’s links on which RBn is forwarder for that VLAN. Where MSTP (Multiple Spanning Tree Protocol) is running on a link, this is the root bridge of the CIST (Common and Internal Spanning Tree). This is to quickly detect cases where two Layer 2 clouds accidentally get merged, and where there might otherwise temporarily be two DRBs for the same VLAN on the same link. (See Section 4.2.3.3.)

6.4 Optionally, per VLAN Layer 2 multicast addresses derived from IPv4 IGMP or IPv6 MLD notification messages received from attached end nodes on that VLAN, indicating the location of listeners for these multicast addresses (see Section 4.3.4).

6.5 Per VLAN ESADI participation flag, priority, and holding time.
If this flag is one, it indicates that the RBridge wishes to receive such TRILL ESADI frames (see Section 4.2.4.1).

6.6 Per VLAN appointed forwarder status lost counter (see Section 4.6.2).

7. Optionally, a list of VLAN groups where address learning is shared across that VLAN group (see Section 4.6.3). Each VLAN group is a list of VLAN IDs, with the first VLAN ID listed in a group, if present, is the "primary" and the others are "secondary". This is to detect misconfiguration of features outside the scope of this document. RBridges that do not support features such as "shared VLAN learning" ignore this field.

4.2.4 TRILL ESADI IS-IS

RBridges that are the appointed VLAN-x forwarder for a link MAY participate in the TRILL end station address distribution instance (ESADI) of IS-IS for that VLAN. But all transit RBridges MUST properly forward TRILL ESADI frames as if they were multicast TRILL data frames.

Because of this forwarding, it appears to an IS-IS ESADI at an RBridge that it is directly connected by a shared virtual link to all other RBridges in the campus running that instance. RBridges that do not implement that ESADI or are not appointed forwarder for that VLAN do not decapsulate or locally process any TRILL ESADI frames they receive for that VLAN. In other words, these frames are transparently tunneled through transit RBridges. Such transit RBridges treat them exactly as multicast TRILL data frames and no IS-IS processing is invoked due to such forwarding.

4.2.4.1 TRILL ESADI Participation

An RBridge participating in an end station address distribution instance (ESADI) does not send any additional Hellos. The information available in the core TRILL IS-IS link state database is sufficient to determine the DRB on the virtual link for each VLAN's ESADI. In particular, the core link state database information for each RBridge includes the VLANs, if any, for which that RBridge is participating in an ESADI, its priority for being selected as DRB for each of those instances, its holding time, and its IS-IS system ID for breaking ties in priority.

The DRB sends TRILL ESADI CSNP frames on the ESADI virtual link. A participating RBridge that determines that some other RBridge should
be DRB on such a virtual link and has not received or sent a CSNP in at least the DRB holding time MAY also send a CSNP on the virtual link. A participating RBridge that determines that no other RBridges are participating in an ESADI for a particular VLAN SHOULD NOT send TRILL ESADI LSPs or CSNPs on the virtual link.

4.2.4.2 TRILL ESADI Information

The information in the LSP for a optional TRILL ESADI is the list of local end station MAC addresses known to the originating RBridge and for each such address a one octet unsigned "confidence" rating in the range 0-254 (see Section 4.6). In order to make it practical to optionally provide for customer VLAN ID translation, as specified in a separate document, TRILL ESADI frames MUST NOT contain the VLAN ID in the body of the frame after the Inner.VLAN tag.

4.3 Distribution Trees

RBridges use distribution trees to forward multi-destination frames (see Section 2.2.2). Distribution Trees are bidirectional. Although a single tree is logically sufficient for the entire campus, the computation of additional distribution trees is warranted for the following reasons: it enables multipathing of multi-destination frames and enables the choice of a tree root closer to or, in the limit, identical with the ingress RBridge. Such a closer tree root reduces out-of-order delivery when a unicast address transitions between unknown and known and improves the efficiency of the delivery of multi-destination frames that are being delivered to a subset of the links in the campus.

Each RBridge RBn may advertise in the core instance link state database its priority to be chosen as a tree root and the number of additional distribution trees it specifies that every RBridge in the campus must compute if RBn is the highest priority tree root. The priority is a 16-bit unsigned integer that defaults, for a zero configuration RBridge to 0x8000. The number of distribution trees to be computed is a 16-bit unsigned integer giving the number of trees to be computed in addition to the one rooted at the highest priority root. The number of additional trees defaults, for a zero configuration RBridge, to one.

The RBridge with highest priority to be tree root is determined by the numerically lowest priority field or, if priority fields are equal, by the numerically lowest system ID. A tree is always calculated rooted at this highest priority RBridge and that RBridge specifies to all RBridges in the campus the total number of
additional distribution trees to be calculated. If it indicates that K-1 additional trees are to be calculated, then they are rooted at the 2nd through the Kth highest priority R Bridges. Thus every R Bridge calculates the same set of K distribution trees.

Every R Bridge RBn defaults, in the zero configuration case, to using a single distribution tree for multi-destination frames. For this purpose, it orders the trees being computed for the campus in order of increasing cost from RBn to the root R Bridge of that tree and, if cost is equal, by decreasing priority to be a tree root and selects the first tree in that ordering.

If RBn is to multi-path multi-destination frames, it can be configured with the number of different trees it would like to use, say J. RBn selects the first trees in its priority-of-use ordering, up to the minimum of J and K number of trees. However, RBn MUST announce, in its LSP, an intention to use any particular trees by listing the tree root, unless it is content to only use the highest priority tree root in the campus.

4.3.1 Distribution Tree Calculation and Checks

R Bridges do not use the spanning tree protocol to calculate distribution trees. Instead, distribution trees are calculated based on the link state information, selecting a particular R Bridge as the root. Each R Bridge RBn independently calculates a tree rooted at RBi by performing the SPF (Shortest Path First) calculation with RBi as the root without requiring any additional exchange of information.

When a node RBn has two or more minimal equal cost paths toward the Root RBi, a deterministic tiebreaker is needed to guarantee that all R Bridges calculate the same distribution tree. This is obtained by selecting the path that goes to the parent that has the lower 7-octet IS-IS ID.

Each R Bridge RBn keeps a set of adjacencies ( { port, neighbor} pairs ) for each distribution tree it is calculating. One of these adjacencies is toward the tree root RBi and the others are toward the leaves. Once the adjacencies are chosen, it is irrelevant which ones are towards the root RBi, and which are away from RBi. Let's suppose that RBn has calculated that adjacencies a, c, and f are in the RBi tree. A multi-destination frame for the distribution tree RBi is received only from one of the adjacencies a, c, or f (otherwise it is discarded) and forwarded to the other two adjacencies.

To further avoid temporary multicast loops during topology changes, R Bridges MUST do a sanity check that a multi-destination frame arrives on the expected link. This is called the Reverse Path...
Forwarding Check and is done as follows. When RBn calculates the RBi tree, for each adjacency in the RBi tree, RBn lists the possible ingress RBridge nicknames on that adjacency. The only ingress RBridges that appear on any of the adjacencies are RBridges that have explicitly stated, in their LSP, that they may select RBi as a distribution tree root or ingress RBridges that list no roots on adjacencies for the distribution tree with the highest priority root. If a multi-destination frame is received on a particular adjacency, marked as the RBi-tree, then RBn MUST NOT forward it if the ingress RBridge is not listed in the allowed list of ingress RBridges for that adjacency for that tree.

When a topology change occurs (including apparent changes during start up), an RBridge MUST adjust its input distribution tree filters no later than it adjusts its output forwarding.

4.3.2 Pruning the Distribution Tree

Each distribution tree SHOULD be pruned per-VLAN, eliminating branches that have no potential receivers downstream. Multi-destination TRILL data frames SHOULD only be forwarded on branches that are not pruned.

Further pruning SHOULD be done in several cases: (1) IGMP [RFC3376], MLD [RFC2710], and MRD [RFC4286] messages, where these are to be delivered only to links with IP Multicast routers; (2) other multicast frames derived from an IP multicast address that should be delivered only to links that have registered listeners, plus links which have IP Multicast routers, except for IP multicast addresses which must be broadcast; and (3) other multicast traffic not derived from an IP address that is only delivered to links for which the appointed forwarder has the Other Multicast requested flag set. All of these cases are scoped per-VLAN.

Let’s assume that RBridge RBn knows that adjacencies (a, c, and f) are in the RBi-distribution tree. RBn marks pruning information for each of the adjacencies in the RBi-tree. For each adjacency and for each tree, RBn marks:

- the set of VLANs reachable downstream,
- for each one of those VLANs, flags indicating whether there are IPv4 or IPv6 multicast routers downstream, and whether there are one or more RBridges downstream with the Other Multicast flag set, and
- the set of Layer 2 multicast addresses derived from IP multicast groups for which there are receivers downstream.
4.3.3 Tree Distribution Optimization

RBridges MUST determine the VLAN associated with all native frames they ingress and properly enforce VLAN rules on the emission of native frames at egress RBridge ports according to how those ports are configured. They SHOULD also prune the distribution tree of multi-destination frames according to VLAN. But, since they are not required to do such pruning, they may receive TRILL data frames that should have been VLAN pruned earlier in the tree distribution. They silently discard such frames. A campus may contain some Rbridges that prune on VLAN and some that do not.

The situation is more complex for multicast. RBridges SHOULD analyze IP derived native multicast frames, and learn and announce listeners and IP multicast routers for such frames as discussed in Section 4.5 below. And they SHOULD prune the distribution of IP derived multicast frames based on such learning and announcements. But, they are not required to prune based on IP multicast listener and router attachment state. And, unlike VLANs, where VLAN attachment state of ports MUST be maintained and honored, RBridges are not required to maintain IP multicast listener and router attachment state.

An RBridge that does not examine native IGMP [RFC3376], MLD [RFC2710], and MRD [RFC4286] frames that it ingresses MUST advertise that it has IPv4 and IPv6 IP multicast routers attached for all the VLANs for which it is an appointed forwarder. It need not advertise any IP derived multicast listeners. This will cause all IP derived multicast traffic to be sent to this RBridge for those VLANs. It then egresses that traffic onto the links for which it is appointed forwarder where the VLAN of the traffic matches the VLAN for which it is appointed forwarder on that link. (This may cause the suppression of certain IGMP membership report messages from end stations but that is not significant as any multicast traffic such reports would be requesting will be sent to such end stations under these circumstances.)

A campus may contain a mixture of Rbridges with different levels of IP derived multicast optimization. An RBridge may receive IP derived multicast frames that should have been pruned earlier in the tree distribution. It silently discards such frames.

See also "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches" [RFC4541].
4.3.4 Forwarding Using a Distribution Tree

With full optimization, forwarding a multi-destination data frame is done as follows:

- The RBridge RBn receives a multi-destination TRILL data frame with inner VLAN-x and a TRILL header indicating that the selected tree is the RBi-tree;

- if the adjacency from which the frame was received is not one of the adjacencies in the RBi-tree for the specified ingress RBridge, the frame is dropped (see Section 4.3.1);

- else, if the frame is an IGMP or MLD announcement message or an MRD query message, then the frame is forwarded onto adjacencies in the RBi-tree that indicate there are downstream VLAN-x IPv4 or IPv6 multicast routers as appropriate;

- else, if the frame is for a Layer 2 multicast address derived from an IP multicast group, but its IP address is not the range of IP multicast addresses that must be treated as broadcast, the frame is forwarded onto adjacencies in the RBi-tree that indicate there are downstream VLAN-x IP multicast routers of the corresponding type (IPv4 or IPv6), as well as adjacencies that indicate there are downstream VLAN-x receivers for that group address;

- else, if the frame is for a Layer 2 multicast address not derived from an IP multicast group, then the frame is forwarded onto adjacencies in the RBi-tree that indicate there are downstream RBridges in VLAN-x with the Other Multicast flag set;

- else (the inner frame is for an unknown destination or broadcast or an IP multicast address which is required to be treated as broadcast) the frame is forwarded onto an adjacency if and only if that adjacency is in the RBi-tree, and marked as reaching VLAN-x links.

For each link for which RBn is appointed forwarder, RBn additionally checks to see if it should decapsulate the frame and send it to the link, or process the frame.

TRILL ESADI frames will be delivered only to RBridges that are appointed forwarders for their VLAN. Such frames will be multicast throughout the campus, like other non-IP-derived multicast data frames, on the distribution tree chosen by the RBridge which created the TRILL ESADI frame, and pruned according to the Inner.VLAN ID. Thus all the RBridges that are appointed forwarders for a link in that VLAN receive them unless the RBridge has cleared its Other Multicast bit for that VLAN and has no appointed forwarders downstream in the tree with the Other Multicast bit set.
4.4 Frame Processing Behavior

This section describes RBridge behavior for all varieties of received frames, including how they are forwarded when appropriate. Section 4.4.1 covers native frames, Section 4.4.2 covers TRILL frames, and section 4.4.3 covers control frames. Processing may be organized or sequenced in a different way than described here as long as the result is the same.

Frames with a bad FCS are discarded on receipt. Source address information ( { VLAN, Outer.MacSA, port } ) is learned from any frame with a unicast sources address (see Section 4.6).

4.4.1 Receipt of a Native Frame

If end station service is disabled on the port, the frame is discarded.

The ingress RBridge RB1 determines the VLAN ID for a native frame according to the same rules as IEEE 802.1Q bridges do (see Appendix D and Section 4.7.2). Once the VLAN is determined, if RB1 is not the appointed forwarder for that VLAN on the port where the frame was received, the frame is discarded. If it is appointed forwarder for that VLAN and is not inhibited (see Section 4.2.3.3), then the native frame is forwarded according to 4.4.1.1 if it is unicast and according to 4.4.1.2 if it is multicast or broadcast.

4.4.1.1 Native Unicast Case

If the destination MAC address of the native frame is a unicast address, the following steps are performed.

The Layer 2 destination address and VLAN are looked up in the ingress RBridge’s Encapsulation Database to find the egress RBridge RBm or the local egress port or to discover that the destination is unknown. One of the following three cases will apply:

1. If the destination is known to be on the same link from which the native frame was received, the RBridge silently discards the frame, since the destination should already have received it.

2. If the destination is known to be on a different local link for which RBm is the appointed forwarder, then RB1 converts the native frame to a TRILL data frame with an Outer.MacSA of RB1 and an Outer.MacDA of the next hop RBridge towards RBm, a TRILL header with M = 0, the ingress nickname for RB1, and the egress nickname
for RBm. (RBm may be RB1 in which case processing then proceeds as in 4.4.2.2.1.)

3. If a unicast destination address is unknown, RB1 handles the frame as described in Section 4.4.1.2 for a broadcast frame except that the Inner.MacDA is the original native frame’s unicast destination address.

4.4.1.2 Native Multicast and Broadcast Frames

If the frame is a native IGMP [RFC3376], MLD [RFC2710], or MRD [RFC4286] frame, then RB1 SHOULD analyze it, learn any group membership or IP multicast router presence indicated, and announce that information for the appropriate VLAN in its LSP (see Section 4.5).

For all multi-destination native frames, RB1 forwards the frame in native form to its links where it is appointed forwarder for the frame’s VLAN subject to further pruning and inhibition. In addition, it converts the native frame to a TRILL data frame with Outer.MacSA of RB1 and the All-Rbridges multicast address as Outer.MacDA, a TRILL header with the multi-destination bit M = 1, the ingress nickname for RB1, and the egress nickname for the root of the distribution tree it decides to use. It then forwards the frame on the pruned distribution tree (see Section 4.3).

The default is for RB1 to write into the egress nickname field the nickname for the distribution tree, from the set of distribution trees being computed by each RBridge in the campus, whose root is least cost from RB1. However, RB1 MAY choose a different distribution tree if RB1 has been configured to path-split multicast. In that case RB1 MUST select a tree by specifying an RBridge that is a distribution tree root (see Section 4.3). Also, RB1 MUST select a tree that RB1 has announced (in RB1’s own LSP) to be one of those that RB1 may choose as a distribution tree or the tree with the highest priority root if none is announced.

Although the Outer.MacDA is normally the All-Rbridges multicast address if, for any particular frame sent out a particular port, there is only one next hop RBridge of interest, the frame MAY be sent with the unicast Outer.MacDA of the target RBridge. (Using a unicast Outer.MacDA is of no benefit on a point-to-point link but may result in substantial savings if the link is actually a bridged LAN with many bridged branches and end stations, to all of which the frame may get flooded if a multicast destination is used.)
4.4.2 Receipt of a TRILL Frame

A TRILL frame has either a multicast Outer.MacDA allocated to TRILL (see Section 7.2) or is a non-control frame with an outer TRILL Ethertype. Processing proceeds in the following order:

If the Outer.MacDA is All-IS-IS-RBridges, the frame is handled as described in Section 4.4.2.1.

If the Outer.MacDA is All-ESADI-RBridges, the frame is discarded.

If the Ethertype is not TRILL, the frame is discarded.

If the Outer.MacDA is a unicast address, the frame is discarded unless that address is the address of the receiving Rbridge. (Such discarded frames are most likely addressed to another Rbridge on a multi-access link and that other Rbridge will handle them.)

After the above checks, further processing of TRILL frames is independent of the Outer.MacDA.

If the Version field in the TRILL Header is greater than 0, the frame is discarded. The Inner.MacDA is then tested. If it is the All-ESADI-Rbridges multicast address and RBn implements the ESADI feature, processing proceeds as in Section 4.4.2.2 below. If it is any other address or RBn does not implement the ESADI feature, processing proceeds as in Section 4.4.2.3.

4.4.2.1 TRILL IS-IS Frames

If the frame protocol is not IS-IS NSAP, it is discarded. Otherwise, it is processed by the core IS-IS instance on RBn and is not forwarded.

4.4.2.2 TRILL ESADI Frames

The port on which the frame was received is first checked and the frame discarded if there is no TRILL core IS-IS adjacency on that port.

If M == 0, the frame is silently discarded. The egress nickname designates the distribution tree. In this case, the frame is forwarded as described in Section 4.4.2.3.2. In addition, if the forwarding Rbridge is an appointed forwarder for a link in the specified VLAN and implements a TRILL ESADI for that VLAN and that instance is enabled, the inner frame is decapsulated and provided to
that local IS-IS instance.

4.4.2.3 TRILL Data Frames

The port on which the frame was received is first checked and the frame discarded if there is no TRILL core IS-IS adjacency on that port.

The egress nickname in the TRILL header is examined and, if it is unknown or reserved, the frame is discarded.

The M flag is then checked. If it is zero, processing continues as described in Section 4.4.2.3.1, if it is one, processing continues as described in Section 4.4.2.3.2.

4.4.2.3.1 Known Unicast TRILL Data Frames

If the egress RBridge indicated is the RBridge performing the processing (RBn), the frame being forwarded is decapsulated to native form. The Inner.MacDA is checked: if it is not unicast, the frame is silently discarded; if it is unicast, the frame is then either (1) sent onto the link containing the destination if the RBridge is appointed forwarder for that link for the frame’s VLAN and is not inhibited, (2) locally processed if the RBridge itself is the destination, or (3) processed as in the following paragraph.

A known unicast TRILL data frame can arrive at the egress RBridge only to find that the Inner.MacDA is not actually known by that RBridge. One way this can happen is that the Inner.MacDA may have timed out in the egress RBridge MAC address cache. In this case, the egress RBridge sends the native frame out on all links that are in the frame’s VLAN for which the RBridge is appointed forwarder and has not been inhibited, except that it MAY refrain from sending the frame on links where it knows there cannot be an end station with the destination MAC address, for example the link port is configured as a trunk (see Section 4.7.1).

If RBn is a transit RBridge and the hop count is zero, the frame is silently discarded. Otherwise the hop count is decremented by one and the frame forwarded to the next hop RBridge towards the egress RBridge, using the Forwarding Database. The Inner.VLAN is not examined by a transit RBridge forwarding a know unicast TRILL data frame.

4.4.2.3.2 Multi-Destination TRILL Data Frames

The Outer.MacSA is checked and the frame discarded if it is not a tree adjacency for the tree indicated by the egress RBridge nickname.
or the reverse path forwarding check fails (see Section 4.3.1).

If the RBridge is an appointed forwarder for the VLAN of the frame, a copy of the frame is decapsulated, sent in native form on those links in its VLAN for which the RBridge is appointed forwarder subject to additional pruning and inhibition as described in Section 4.2.3.3, and/or locally processed as appropriate.

If the hop count in the frame is zero, it is then silently discarded. If non-zero, it is decreased (see Section 3.6) and the frame forwarded down the tree specified by the egress RBridge nickname pruned as described in Section 4.3.

In the forwarded frame, the Outer.MacSA is set to that of the port on which the frame is being transmitted and the Outer.MacDA is normally the All-Rbridges multicast address; however, if for any particular frame transmitted on a particular port there is only one next hop RBridge of interest, the frame MAY be sent with a unicast Outer.MacDA of that next hop RBridge. (Using a unicast Outer.MacDA is of no benefit on a point-to-point link but may result in substantial savings if the link is actually a bridged LAN with many bridged branches and end stations, to all of which the frame may get flooded if a multicast destination is used.)

4.4.3 Receipt of a Control Frame

Low-level control frames received by an RBridge are handled within the port where they are received as described in Section 4.7.

There are two types of high-level control frames, distinguished by their destination address, which are handled as described in the sections referenced below.

<table>
<thead>
<tr>
<th>Name</th>
<th>Section</th>
<th>Destination Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPDU</td>
<td>4.7.3</td>
<td>01-80-C2-00-00-00</td>
</tr>
<tr>
<td>VRP</td>
<td>4.7.4</td>
<td>01-80-C2-00-00-21</td>
</tr>
</tbody>
</table>

4.5 IGMP, MLD, and MRD Learning

Ingress RBridges SHOULD learn, based on seeing native IGMP [RFC3376], MLD [RFC2710], and MRD [RFC4286] frames, which IP derived multicast messages should be forwarded onto which links. Such frames are also, in general, encapsulated as TRILL data frames and distributed as described below and in Section 4.3.
An IGMP or MLD membership report received in native form from a link indicates a multicast group listener for that group on that link. An IGMP or MLD query or an MRD advertisement received in native form from a link indicates the presence of an IP multicast router on that link.

IP multicast group membership reports have to be sent throughout the campus and delivered to all IP multicast routers, distinguishing IPv4 and IPv6. All IP-derived multicast traffic must also be sent to all IP multicast routers for the same version of IP.

IP multicast data SHOULD only be sent on links where there is either an IP multicast router for that IP type (IPv4 or IPv6) or an IP multicast group listener for that IP multicast derived MAC address, unless the IP multicast address is in the range required to be treated as broadcast.

RBridges do not need to announce themselves as listeners to the All-Snoopers multicast group (the group used for MRD reports [RFC4541]), because the IP multicast address for that group is in the range where all frames sent to that IP multicast addresses must be broadcast.

See also "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches" [RFC4541].

4.6 End Station Address Details

RBridges have to learn the MAC addresses and VLANs of their locally attached end stations for link/VLAN pairs for which they are the appointed forwarder so they can

- forward the native form of incoming known unicast TRILL data frames onto the correct link and

- decide, for an incoming native unicast frame from a link, where the RBridge is the appointed forwarder for the frame’s VLAN, whether the frame is
  - known to have been destined for another end station on the same link, so the RBridge need do nothing, or
  - has to be converted to a TRILL data frame and forwarded.

RBridges need to learn the MAC addresses, VLANs, and remote RBridges of remotely attached end stations for VLANs for which they and the remote RBridge are an appointed forwarder, so they can efficiently direct native frames they receive which are unicast to those...
addresses and VLANs.

4.6.1 Learning End Station Addresses

There are five independent ways an RBridge can learn end station addresses as follows:

1. From the observation of VLAN-x frames received on ports where it is appointed VLAN-x forwarder, learning the { source MAC, VLAN, port } triplet of received frames.

2. The { source MAC, VLAN, ingress RBridge nickname } triplet of any native frames that it decapsulates.

3. By Layer 2 registration protocols learning the { source MAC, VLAN, port } of end stations registering at a local port.

4. By running one or more TRILL ESADIs that receive remote address information and transmit local address information.

5. By management configuration.

RBridges MUST implement capabilities 1 and 2 above. RBridges use these capabilities unless configured, for one or more particular VLANs and/or ports, to not learn from either received frames or from decapsulating native frames to be transmitted or both.

RBridges MAY implement capabilities 3 and 4 above. If capability 4 is implemented, such ESADIs are run only when the RBridge is configured to do so on a per-VLAN basis.

RBridges SHOULD implement capability 5.

Entries in the table of learned MAC addresses and associated information also have a one octet unsigned confidence level associated with each entry. Such information learned from the observation of data has a confidence of 0x20 unless configured to have a different confidence. This confidence level can be configured on a per RBridge basis separately for information learned from local native frames and that learned from remotely originated encapsulated frames. Such information received via TRILL ESADI is accompanied by a confidence level in the range 0 to 254. Such information configured by management defaults to a confidence level of 255 but may be configured to have another value.

The table of learned MAC addresses includes (1) { confidence, VLAN, MAC address, local port } for addresses learned from local native frames and local registration protocols, (2) { confidence, VLAN, MAC
address, egress RBridge nickname } for addresses learned from remote encapsulated frames and ESADI link state databases, and (3) additional information to implement timeout of learned addresses, statically configured addresses, and the like.

When a new learned address and related information are to be entered into the local database there are three possibilities:

A. If this is a new { address, VLAN } pair, the information is entered accompanied by the confidence level.

B. If there is already an entry for this { address, VLAN } pair with the same accompanying delivery information, the confidence level in the local database is set to the maximum of its existing confidence level and the confidence level with which it is being learned. In addition, if the information is being learned with the same or a higher confidence level than its existing confidence level, timer information is reset.

C. If there is already an entry for this { address, VLAN } pair with different information, the learned information replaces the older information only if it is being learned with higher or equal confidence than that in the database entry. If it replaces older information, timer information is also reset.

4.6.2 Forgetting End Station Addresses

While RBridges need to learn end station addresses as described above, it is equally important that they be able to forget such information. Otherwise, frames for end stations that have moved to a different part of the campus could be indefinitely black holed by RBridges with stale information as to the link to which the end station is attached.

For end station address information locally learned from frames received, the time out from the last time a native frame was received or decapsulated with the information conforms to the recommendations of [802.1Q]. It is referred to as the "Aging Time" and is configurable per RBridge with a range of from 10 seconds to 1,000,000 seconds and a default value of 300 seconds.

The situation is different for end station address information acquired via TRILL ESADI. It is up to the originating RBridge to decide when to remove such information from the ESADI LSP (or up to IS-IS timeouts if the originating RBridge becomes inaccessible).

When an RBridge ceases to be appointed forwarder for VLAN-x on a port, it forgets all end station address information learned from the
observation of VLAN-x native frames received on that port. It also increments a per VLAN counter of the number of times it lost appointed forwarder status for that VLAN.

When, for all of its ports, RBridge RBn is no longer appointed forwarder for VLAN-x, it forgets all end station address information learned from decapsulating VLAN-x native frames. Also, if RBn is participating in TRILL ESADI for VLAN-x, it ceases to so participate after sending a final LSP nulling out the end station address information for that VLAN which it had been originating. In addition, all other RBridges that are VLAN-x forwarder on at least one of their ports notice that the link state data for RBn has changed to show that it no longer has a link on VLAN-x. In response, they forget all end station address information they have learned from decapsulating VLAN-x frames that show RBn as the ingress RBridge.

When the appointed forwarder lost counter for RBridge RBn for VLAN-x is observed to increase via the core TRILL IS-IS link state database but RBn continues to be an appointed forwarder for VLAN-x on at least one of its ports, every other RBridge that is an appointed forwarder for VLAN-x modifies the aging of all the addresses it has learned by decapsulating native frames in VLAN-x from ingress RBridge RBn as follows: The time remaining for each entry is adjusted to be no larger than a per RBridge configuration parameter called (to correspond to [802.1Q]) "Forward Delay". This parameter is in the range of 4 to 30 seconds with a default value of 15 seconds.

4.6.3 Shared VLAN Learning

RBridges can map VLAN IDs into a smaller number of identifiers for purposes of address learning, as [802.1Q] bridges can. Then, when a lookup is done in learned address information, this identifier is used for matching in place of the VLAN ID. If the ID of the VLAN on which the address was learned is not retained, then there are the following consequences:

- The RBridge no longer has the information needed to participate in a TRILL ESADI for the VLANs who’s ID is not being retained.

- In cases where 4.6.2 above requires the discarding of learned address information based on a particular VLAN, when the VLAN ID is not available for entries under a shared VLAN identifier, instead the time remaining for each entry under that shared VLAN identifier is adjusted to be no longer than the RBridge’s "Forward Delay".

Although outside the scope of this specification, there are some Layer 2 features in which a set of VLANs has shared learning, where
one of the VLANs is the "primary" and the other VLANs in the group are "secondaries". An example of this is where traffic from different communities are separated using VLAN tags, and yet some resource (such as an IP router or DHCP server) is to be shared by all the communities. A method of implementing this feature is to give a VLAN tag, say Z, to a link containing the shared resource, and have the other VLANs, say A, C, and D, be part of the group \( \{ \text{primary} = Z, \text{secondaries} = A, C, D \} \). An RBridge, aware of this grouping, attached to one of the secondary VLANs in the group also claims to be attached to the primary VLAN. So an RBridge attached to A would claim to also be attached to Z. An RBridge attached to the primary would claim to be attached to all the VLANs in the group.

This document does not specify how VLAN groups might be used. Only RBridges that participate in a VLAN group will be configured to know about the VLAN group. However, to detect misconfiguration, an RBridge configured to know about a VLAN group SHOULD report the VLAN group in its LSP.

4.7 RBridge Ports

Section 4.7.1 below describes the several general RBridge port configuration bits, Section 4.7.2 give a logical port structure in terms of frame processing, and Sections 4.7.3 and 4.7.4 describe the handling of high-level control frames.

4.7.1 RBridge Port Configuration

There are three per port configuration bits as follows:

- Disable port bit. When this bit is set, all frames received are discarded and no frames are sent, with the possible exception of some low-level control frames that may be received and processed locally.

- End station service disable (trunk port) bit. When this bit is set, all native frames received on the port and all native frames that would have been sent on the port are discarded. (See Appendix B.)
  (Note that, for this document, "native frames" does not include control frames.)

A port with end station service disabled reports, in the Hellos it sends out that port, that it has no VLANs for which it is provides end station support. As a result, such a port will not be appointed forwarder for any VLAN. Thus a port with end station service disabled cannot contribute to the VLANs which the RBridge
reports in its LSP as being "connected" to that RBridge. Unless there is at least one port on an RBridge for which VLAN-x is appointed forwarder, that RBridge does not normally advertise itself in the link state as connected to VLAN-x. As a consequence, it will not normally receive any traffic for VLAN-x except as TRILL data frames to forward as a transit RBridge.

- TRILL traffic disable (access port) bit. If this bit is set, the goal is to avoid sending TRILL frames, except TRILL Hellos, on the port since it is intended for end station traffic (see Appendix B). This bit is reported in Hellos sent out the port and the bit for the DRB controls the link. If the DRB asserts the access port bit in its Hello, then the link is an RBridge access link all RBridge ports which acknowledge that DRB act as access ports. If there are no TRILL IS-IS adjacencies on the access port, no special action need be taken.

If there are adjacencies, and the RBridge is DRB, it normally does not create a pseudonode for the link. In that case, no adjacencies over the access link are reported in their LSPs by any of the RBridges connected to the link. In this case no TRILL frames, except Hellos, are sent out the access ports.

Alternatively, the DRB MAY choose to creates a pseudonode for the access link. If it does create a pseudonode, it sets the IS-IS overflow bit in the pseudonode. This will cause IS-IS routing to avoid sending transit data on the link if any other path is available but it will still be available as a path of last resort; however, TRILL IS-IS frames will be sent over the link. In addition, if a pseudonode is created by the DRB for an access link, all the RBridges on the access link report connectivity to the pseudonode as usual and the DRB reports connectivity in the LSP it creates for the pseudonode.

4.7.2 RBridge Port Structure

An RBridge port can be modeled as having a structure, in its lower levels, similar to that of an [802.1Q] bridge port as shown in Figure 4.3. In this figure, the double lines represent the general flow of the frames and information while single lines represent information flow only. The dashed lines in connection with VRP (GVRP/MVRP) are to show that VRP support is optional. An actual RBridge port implementation may be structured in any way that provides the correct behavior.
Low-level control frames are handled in the lower level port/link control logic in the same way as in an 802.1Q bridge. This can optionally include a variety of 802.1 or link specific protocols such as link layer discovery, link aggregation (Clause 43 of [802.3]), MAC security [802.1AE], or port based access control [802.1X]. While handled at a low level, these frames may affect higher level processing. For example, a layer 2 registration protocol may affect the confidence in learned addresses. The upper interface to this...
lower level port control logic corresponds to the Internal Sublayer Service (ISS) in 802.1Q.

High-level control frames (BPDUs and, if supported, VRP frames) are not VLAN tagged. Although they extend through the ISS interface, they are not subject to port VLAN processing. Behavior on receipt of a VLAN tagged BPU or VLAN tagged VRP frame, is unspecified. If a VRP is not implemented, then all VRP frames are discarded. Handling of BPDUs is described in Section 4.7.3. Handling of VRP frames is described in Section 4.7.4.

Non-control frames, that is, all TRILL and native frames, are subject to Port VLAN and priority processing which is the same as for an 802.1Q bridge. The upper interface to the port VLAN processing corresponds to the Extended Internal Sublayer Service (EISS) in 802.1Q.

Incoming native frames are only accepted if the RBridge is an uninhibited appointed forwarder for the frame’s VLAN after which they are normally encapsulated and forwarded. Outgoing native frames are obtained by decapsulation but only output if the RBridge is an uninhibited appointed forwarder for the frame’s VLAN.

TRILL Hellos are handled per port and never forwarded. They can affect the appointed forwarder and inhibition logic as well as the RBridge’s LSP.

TRILL IS-IS frames, other than Hellos, and TRILL data and ESADI frames are pass up to higher level RBridge processing on receipt and transmitted on creation or forwarding. Note that these frames are never blocked due to the appointed forwarder and inhibition logic but there are additional filters on some of them such as the Reverse Path Forwarding check.

4.7.3 BPDU Handling

If RBridge campus topology were static, RBridges would be simply end stations from a bridging perspective, terminating but not otherwise interacting with spanning tree. However, there are reasons for RBridges to listen to and sometimes to transmit BPDUs as described below. Even when RBridges listen to and transmit BPDUs, these are a local RBridge port activity. The ports of a particular RBridge never interact so as to make the RBridge as a whole a spanning tree node.
4.7.3.1 Receipt of BPDUs

Rbridges MUST listen to spanning tree BPDUs received on a port and keep track of the root bridge, if any, on that link. If MSTP is running on the link, this is the CIST root. This information is reported per VLAN by the RBridge in its LSP. In addition, the receipt of spanning tree BPDUs is used as an indication that a link is a bridged LAN, which can affect the RBridge transmission of BPDUs.

An RBridge MUST NOT encapsulate or forward any BPDU frame it receives.

Rbridges discard any topology change BPDUs they receive, but note Section 4.7.3.3.

4.7.3.2 Root Bridge Changes

A change in the root bridge seen out a port may indicate a change in bridged LAN topology including the possibility of the merger of two bridged LANs or the like. During topology transients, bridges may go into pre-forwarding states that block TRILL IS-IS Hellos. For these reasons, when an RBridge sees a root bridge change on a port for which it is appointed forwarder for one or more VLANs, it is inhibited (discards all native frames received from or which it would otherwise have sent to the link) for a period of time between 30 and zero seconds. This time period is configurable per RBridge and defaults to 30 seconds.

For example, consider two bridged LANs carrying multiple VLANs, each with various appointed forwarders. Should they become merged, due to a cable being plugged in or the like, those R Bridges attached to the original bridged LAN with the lower priority root will see a root bridge change while those attached to the other original bridged LAN will not. Thus all appointed forwarders in the first set will be inhibited for the time period while things are sorted out by BPDUs within the merged bridged LAN and TRILL IS-IS Hellos between all the R Bridges involved.

4.7.3.3 Transmission of BPDUs

When an RBridge ceases to be appointed forwarder for one or more VLANs out a particular port it SHOULD, as long as it continues to receive spanning tree BPDUs on that port, send topology change BPDUs until it sees the topology change acknowledged in a spanning tree BPDU.
RBrigdes MAY support a capability for sending spanning tree BPDUs for the purpose of attempting to force a bridged LAN to partition as discussed in Section A.3.3. Except for this optional capability, RBrigdes MUST NOT send spanning tree BPDUs.

4.7.4 Dynamic VLAN Registration

Dynamic VLAN registration provides a means for bridges (and less commonly end stations) to request that VLANs be enabled or disabled on ports leading to the requestor. This is done by VLAN registration protocol (VRP) frames: GVRP or MVRP. RBrigdes MAY implement GVRP and/or MVRP as described below.

VRP frames are never encapsulated as TRILL frames between RBrigdes or forwarded in native form by an RBridge. If an RBridge does not implement a VRP, it discards any VRP frames received and sends none.

RBridge ports may have VLANs whose enablement is dynamic. If an RBridge supports a VRP, the actual enablement of dynamic VLANs is determined by GVRP/MVRP frames received at the port as it would be for an [802.1Q] / [802.1ak] bridge.

An RBridge that supports a VRP sends GVRP/MVRP frames as an [802.1Q] / [802.1ak] bridge would send on each port that is not configured as an RBridge trunk port. For this purpose, it sends VRP frames to request traffic in the VLANs for which it is appointed forwarder and traffic in the Designated VLAN, unless the Designated VLAN is disabled on the port, and to not request traffic in any other VLAN.
5. RBridge Addresses, Parameters, and Constants

IS-IS requires each RBridge to have a unique 48-bit (6-octet) System ID. This is easily obtainable, for example, as any one of the MAC-48 addresses owned by that RBridge.

A new Ethertype must be assigned to indicate a TRILL encapsulated frame.

Three Layer 2 multicast addresses must be assigned:

- All-RBridges for use as Outer.MacDA in TRILL ESDADI and multi-destination TRILL data frames.
- All-IS-IS-RBridges for use as the Outer.MacDA for TRILL IS-IS frames.
- All-ESADI-RBridges for use as the Inner.MacDA for TRILL ESADI frames.

The following per RBridge parameters may be configured:

- A nickname and nickname selection priority.
- Priority to be a distribution tree root and desired number of additional distribution trees for the campus, as discussed in Section 4.3.
- The per RBridge parameters Aging Timer and Forward Delay, as described in Section 4.6.

RBridges may be configured to have ESADIs (end station address distribution instances) of TRILL IS-IS and to send and/or learn end station address information via such instances. Static end address information and priority of such end station information statically configured and learned in various ways can also be configured.

The per RBridge per VLAN Other Multicast bit, which defaults to true, to request the receipt of non-IP derived multicast traffic.

The following RBridge per port parameters:

- The same parameters as for an 802.1Q port in terms of VLAN C-tags and frame priority code points.
- Three per-port configuration bits: disable port, disable end station service (trunk), and access port (see Section 4.7.1).
- Configuration for the optional send-BPDUs solution to the wiring closet topology problem (see Section A.3.3) consists of
System ID of the RBridge with lowest System ID. If RB1 and RB2 are part of a wiring closet topology, both need to be configured to know about this, and that RB1 is the ID that should be used in the spanning tree protocol on the specified port.
6. Security Considerations

Layer 2 bridging is not inherently secure. It is, for example, subject to spoofing of source addresses and bridging control messages. A goal for TRILL is that RBridges do not add new issues beyond those existing in current bridging technology.

Countermeasures are available such as to configure the TRILL IS-IS instances to use IS-IS security [RFC5304] and ignore unauthenticated TRILL IS-IS and ESADI frames received on a port. Since such authentication requires configuration, RBridges using it are no longer zero configuration.

IEEE 802.1 port admission and link security mechanisms, such as [802.1X] and [802.1AE], can also be used. These are best thought of as being implemented within a port and are outside the scope of TRILL (just as they are generally out of scope for bridging standards [802.1D] and [802.1Q]); however, TRILL can make use of secure registration through the confidence level communicated in the optional TRILL ESADIs (see Section 4.6).

TRILL encapsulates native frames inside the TRILL Ethertype while they are in transit between that frame’s ingress RBridge and egress RBridge(s). Thus, TRILL ignorant devices with firewall features and which cannot be detected by RBridges as end stations will generally not be able to inspect the content of such frames for security checking purposes. This may render them ineffective. Routers and hosts appear to RBridges to be end stations and such frames will be decapsulated before being sent to such devices. Thus they will not see the TRILL Ethertype. Firewall devices which do not appear to an RBridge to be an end station, for example bridges with co-located firewalls, should be modified to understand TRILL encapsulation.

6.1 VLAN Security Considerations

TRILL supports VLANs. These provide logical separation of traffic but care should be taken in using VLANs for security purposes. To have reasonable assurance of such separation, all the RBridges and links in a campus must be secured and configured so as to prohibit end stations from using dynamic VLAN registration frames or otherwise gaining access to any VLAN carrying traffic for which they are not authorized to read and/or inject.

Furthermore, if VLANs were used to keep some information off links where it might be observed, this will no longer work with TRILL; with encapsulation and a different outer VLAN tag, the data will travel the least cost transit path regardless of VLAN. Appropriate counter measures are to use end-to-end encryption or an appropriate TRILL
security option should one be specified.

RBridges do not prevent nodes from impersonating other nodes, for instance, by issuing bogus ARP/ND replies. However, RBridges do not interfere with any schemes that would secure neighbor discovery.
7. Assignment Considerations

This section discusses IANA and IEEE 802 assignment considerations. See [RFC5226].

7.1 IANA Considerations

A new IANA registry is created for TRILL Versions, Nicknames, and Version 0 Header Reserved bits.

The initial contents of the TRILL Version Registry is as follows:

<table>
<thead>
<tr>
<th>Version</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>As specified in RFC5226</td>
</tr>
<tr>
<td>1-3</td>
<td>Available for allocation by IETF Standards Action</td>
</tr>
</tbody>
</table>

The initial contents of the Version 0 Header Reserved Bits Registry is as follows:

<table>
<thead>
<tr>
<th>Bit</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x2</td>
<td>Available for allocation by IETF Standards Action</td>
</tr>
<tr>
<td>0x1</td>
<td>Available for allocation by IETF Standards Action</td>
</tr>
</tbody>
</table>

The initial contents of the TRILL Nicknames Registry is as follows:

<table>
<thead>
<tr>
<th>Nickname</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x0000</td>
<td>Reserved to indicate no nickname specified</td>
</tr>
<tr>
<td>0x0001-0xFFBF</td>
<td>Dynamically allocated within each TRILL campus</td>
</tr>
<tr>
<td>0xFFC0-0xFFFE</td>
<td>Available for allocation by IETF Review</td>
</tr>
<tr>
<td>0xFFFF</td>
<td>Permanently reserved</td>
</tr>
</tbody>
</table>

7.2 IEEE Registration Authority Considerations

The Ethertype <tbd> is assigned by the IEEE Registration Authority to the TRILL Protocol.

The Layer 2 multicast MAC addresses <tbd1>, <tbd2>, and <tbd3> are assigned by the IEEE Registration Authority for "All-Rbridges", "All-IS-IS-Rbridges", and "All-ESADI-RBridges" respectively.
8. Normative References


[802.3] "IEEE Standard for Information technology / Telecommunications and information exchange between systems / Local and metropolitan area networks / Specific requirements Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications", 802.3-2005, 9 December 2005


9. Informative References


[RP1999] Perlman, R., "Interconnection: Bridges, Routers, Switches,
Appendix A: Incremental Deployment Considerations

Some aspects of partial RBridge deployment are described below for link cost determination (Section A.1) and possible congestion due to appointed forwarder bottlenecks (Section A.2). A particular example of a problem related to the TRILL use of a single appointed forwarder per link per VLAN (the "wiring closet topology") is explored in detail in Section A.3.

A.1 Link Cost Determination

With an RBridged campus having no bridges or repeaters on the links between RBridges, the RBridges can accurately determine the number of physical hops involved in a path and the line speed of each hop, assuming this is reported by their port logic. With intervening devices, this is no longer possible. For example, as shown in Figure A.1, the two bridges B1 and B2 can completely hide a slow link so that both RBridges RB1 and RB2 incorrectly believe the link is faster.

```
+-----+        +----+        +----+        +-----+
|     |  Fast  |    |  Slow  |    |  Fast  |     |
| RB1 +--------+ B1 +--------+ B2 +--------+ RB2 |
|     |  Link  |    |  Link  |    |  Link  |     |
+-----+        +----+        +----+        +-----+
```

Figure A.1: Link Cost of a Bridged Link

Even in the case of a single intervening bridge, two RBridges may know they are connected but each see the link as a different speed from how it is seen by the other.

However, this problem is not unique to RBridges. For example, routers can encounter similar situations due to links hidden by bridges, repeaters or RBridges.

A.2 Appointed Forwarders and Bridged LANs

With partial RBridge deployment, the RBridges may partition a bridged LAN into a relatively small number of relatively large remnant bridged LANs or possibly not partition it at all, so a single bridged LAN remains. Then two potential problems can occur as follows:

1. The requirement that native frames enter and leave a link via the appointed forwarder for the link and VLAN of the frame can cause congestion or suboptimal routing. (Similar problems can occur
within a bridged LAN due to the spanning tree algorithm.) The extent to which such a problem will occur is highly dependent on the network topology. For example, if a bridged LAN had a star-like structure with core bridges that connected only to other bridges and peripheral bridges that connected to end stations and are singly connected to a core bridge, the replacement of all of the core bridges by R Bridges without replacing the peripheral bridges would generally improve performance without inducing any appointed forwarder congestion. Solutions to this problem are discussed below and a particular example explored in Section A.3.

2. TRILL traffic sent to the All-Rbridges or All-IS-IS-Rbridges multicast addresses will typically be flooded throughout a bridged LAN, which may create a greater burden than necessary. In cases where there is actually only one RBridge next hop recipient of interest, this problem can be eliminated by using the option of unicasting the TRILL data or ESADI frame to that recipient rather than multicasting it.

Inserting R Bridges so that all the bridged portions of the LAN stay connected to each other and have multiple RBridge connections is generally the least efficient arrangement.

There are four techniques that may help if problem 1 above occurs and which can, to some extent, be used in combination:

1. Replace more IEEE 802.1 bridges with R Bridges so as to minimize the size of the remnant bridged LANs between R Bridges. This requires no configuration of the R Bridges unless the bridges they replace required configuration.

2. Re-arrange network topology to minimize the problem. If the bridges and R Bridges involved are configured, this may require changes in their configuration.

3. Configure the R Bridges and bridges so that end stations on a remnant bridged LAN are separated into different VLANs that have different appointed forwarders. If the end stations were already assigned to different VLANs, this is straightforward (see Section 4.2.3.2). If the end stations were on the same VLAN and have to be split into different VLANs, this technique may lead to connectivity problems between end stations.

4. Configure the R Bridges such that their ports which are connected to the bridged LAN send BPDUs (see Section A.3.3) in such a way as to force the partition of the bridged LAN. (Note: A spanning tree is never formed through an RBridge but always terminates at RBridge ports.) To use this technique, the R Bridges must support this optional feature, and would need to be configured to make use of it, but the bridges involved would rarely have to be
configured. Warning: This technique makes the bridged LAN unavailable for TRILL through traffic because the bridged LAN partitions.

Conversely to item 3 above, there may be bridged LANs which use VLANs, or use more VLANs than would otherwise be necessary, to support the Multiple Spanning Tree Protocol or otherwise evade the congestion that can be caused by a single spanning tree. Replacing the IEEE 802.1 bridges in such LANs with RBridgees may enable a reduction in or elimination of VLANs and configuration complexity.

A.3 Wiring Closet Topology

If 802.1 bridges are present and RBridgees are not properly configured, the bridge spanning tree or the DRB may make inappropriate decisions. Below is a specific example of the more general problem that can occur when a bridged LAN is connected to multiple RBridgees.

In cases where there are two (or more) groups of end nodes, each attached to a bridge (say B1 and B2 respectively), and each bridge is attached to an RBridge (say RB1 and RB2 respectively), with an additional link connecting B1 and B2 (see Figure A.2), it may be desirable to have the B1-B2 link only as a backup in case one of RB1 or RB2 or one of the links B1-RB1 or B2-RB2 fail.

![Figure A.2: Wiring Closet Topology](image)

For example, B1 and B2 may be in a wiring closet and it may be easy to provide a short, high bandwidth, low cost link between them while
RB1 and RB2 are at a distant data center such that the RB1-B1 and RB2-B2 links are slower and more expensive.

Default behavior might be that one of RB1 or RB2 (say RB1) would become DRB for the bridged LAN including B1 and B2 and appoint itself forwarder for the VLANs on that bridged LAN. As a result, RB1 would forward all traffic to/from the link, so end nodes attached to B2 would be connected to the campus via the path B2-B1-RB1, rather than the desired B2-RB2. This wastes the bandwidth of the B2-RB2 path and cuts available bandwidth between the end stations and the data center in half. The desired behavior would be to make use of both the RB1-B1 and RB2-B2 links.

Three solutions to this problem are described below.

A.3.1 The RBridge Solution

Of course, if B1 and B2 are replaced with RBridges, the right thing will happen with zero configuration (other than VLAN support), but this may not be immediately practical if bridges are being incrementally replaced by RBridges.

A.3.2 The VLAN Solution

If the end stations attached to B1 and B2 are already divided among a number of VLANs, RB1 and RB2 could be configured so that which ever becomes DRB for this link will appoint itself forwarder for some of these VLANs and the other RBridge for the remaining VLANs. Should either of the RBridges fail or become disconnected, the other will have only itself to appoint as forwarder for all the VLANs.

If the end stations are all on a single VLAN, then it would be necessary to arbitrarily assign them between at least two VLANs to use this solution. This may lead to connectivity problems that might require further measures to rectify.

A.3.3 The Spanning Tree Solution

Another solution is to configure RB1 and RB2 to be part of a "wiring closet group", with a configured System ID RBx (which may be RB1 or RB2’s System ID). Both RB1 and RB2 emit BPDUs on their configured ports as highest priority root RBx. This causes the spanning tree to logically partition the bridged LAN as desired by blocking the B1-B2 link at one end or the other (unless one of the bridges is configured
to also have highest priority and has a lower ID, which we consider to be a misconfiguration). With the B1-B2 link blocked, RB1 and RB2 cannot see each other’s TRILL Hellos via that link and each acts as Designated RBridge and appointed forwarder for its respective partition. Of course, with this partition, no TRILL through traffic can flow over the RB1-B1-B2-RB2 path.

In the spanning tree BPDU, the Root is "RBx" with highest priority, cost to Root is 0, Designated Bridge ID is "RB1" when RB1 transmits and "RB2" when RB2 transmits, and port ID is a value chosen independently by each of RB1 and RB2 to distinguish each of its own ports. (If RB1 and RB2 were actually bridges on the same shared medium with no bridges between them, the result would be that the one with the larger ID sees "better" BPDUs (because of the tiebreaker on the third field: the ID of the transmitting RBridge), and would turn off its port.)

Should either RB1 or the RB1-B1 link or RB2 or the RB2-B2 link fail, the spanning tree algorithm will stop seeing one of the RBx roots and will unblock the B1-B2 link maintaining connectivity of all the end stations with the data center.

If the link RB1-B1-B2-RB2 is on the cut set of the campus and RB2 and RB1 have been configured to believe they are part of a wiring closet group, the campus becomes partitioned as the link is blocked.

### A.3.4 Comparison of Solutions

Replacing all 802.1 bridges with RBridges is usually the best solution with the least amount of configuration required, possibly none.

The VLAN solution works well with a relatively small amount of configuration if the end stations are already divided among a number of VLANs. If they are not, it becomes more complex and problematic.

The spanning tree solution does quite well in this particular case. But it depends on both RB1 and RB2 having implemented the optional feature of being able to configure a port to emit BPDUs as described in Section A.3.3 above. It also makes the bridged LAN whose partition is being forced unavailable for through traffic. Finally, while in this specific example it neatly breaks the link between the two bridges B1 and B2, if there were a more complex bridged LAN, instead of exactly two bridges, there is no guarantee that it would partition into roughly equal pieces. In such a case, you might end up with a highly unbalanced load on the RB1 link and the RB2 link.
Appendix B: Trunk and Access Port Configuration

Many modern bridged LANs are organized into a core and access model, The core bridges have only point-to-point links to other bridges while the access bridges connect to end stations, core bridges, and possibly other access bridges. It seems likely that some RBridge campuses will be organized in a similar fashion.

An RBridge port can be configured as a trunk port, that is, a point-to-point link to another RBridge (or RBridges), by configuring it to disable end station support. There is no reason for such a port to have more than one VLAN enabled and in its Announcing Set on the port. Of course, the RBridge (or RBridges) to which it is connected must have the same VLAN enabled. There is no reason for this VLAN to be other than the default VLAN 1 unless, perhaps, the link is actually over carrier Ethernet facilities that provide some other specific VLAN or the like. Such configuration minimizes wasted TRILL Hellos and eliminates useless decapsulation and transmission of multi-destination traffic in native form onto the link. (see Sections 4.2.3.1 and 4.7.1)

An RBridge access port would be expected to lead to a link with end stations and possibly one or more bridges. Such a link might also have more than one RBridge connected to it to provide more reliable service to the end stations. It would be a goal to minimize or eliminate transit traffic on such a link as it is intended for end station native traffic. This can be accomplished by turning on the access port configuration bit for the RBridge port or ports connected to the link as further detailed in Section 4.7.1.

When designing RBridge configuration user interfaces, consideration should be given to making it convenient to configure trunk and access ports.
Appendix C: Multipathing

Rbridges support multipathing of both known unicast and multi-destination traffic. Implementation of multipathing is optional.

Multi-destination traffic can be multipathed by using different distribution tree roots for different frames. For example, assume that in Figure C.1 end stations attached to RBy are the source of various multicast streams each of which has multiple listeners attached to various of RB1 through RB9. Assuming equal bandwidth links, the distribution tree rooted at RBy will predominantly use the vertical links among RB1 through RB9 while that rooted at RBz will predominantly use the horizontal. If RBy chooses itself as the distribution tree root for half of this traffic and RBz the root for the other half, it may be able to substantially increase the aggregate bandwidth by making use of both the vertical and horizontal links among RB1 through RB9.

Since the distribution trees an RBridge must calculate are the same for all R Bridges and transit R Bridges MUST respect the tree root specified by the ingress RBridge, a campus will operate correctly with a mix of R Bridges some of which use different roots for different multi-destination frames and some of which use a single root for all such frames.

```
+---+
| RBy | ---------------+
+---+
 /   /   /
 /   /   /
 +--++-++--+++
 |  |  |  |  |
 +---++---++---+
 |  |  |  |  |
 +---++---++---+
 |  |  |  |  |
 +---++---++---+
 |  |  |  |  |
 +---++---++---+
 |  |  |  |  |
 +---++---++---+
```

Figure C.1: Multi-Destination Multipath

Known unicast equal cost multipathing (ECMP) can occur if, instead of using a tie-breaker criterion when building an SPF path between ingress and egress R Bridges, information about equal cost paths is retained. Different unicast frames can then be sent via different equal cost paths. For example, in Figure C.2, there are three equal
cost paths between RB1 and RB2 and two equal cost paths between RB2 and RB5.

A transit RBridge receiving a known unicast frame forwards it towards the egress RBridge and is not concerned with whether it believes itself to be on any particular path from the ingress RBridge or a previous transit RBridge. Thus a campus will operate correctly with a mix of RBridges some of which implement ECMP and some of which do not.

As an alternative to multipathing, it might be possible to combine the three paths between RB1 and RB2 into one logical link through the "link aggregation" feature of 802.3 (see Clause 43 of [802.3]). Rbridges MAY implement link aggregation. However, link aggregation requires multiple single hop equal bandwidth links (no intervening bridges). Equal cost multipathing is more general in that there can be multiple hops with intervening bridges and RBridges and links of different costs as long as the path cost is the same. (Generally, the default estimate of the cost of a link is proportional to the reciprocal of its line speed.)

```
+----+       double line = 10 Gbps
  -----      ===|RB3|---     single line = 1 Gbps
   /     \     //
+----+ +----+ +----+  \\
===|RB1|-----|RB2|   |RB5|===
      +----+ +----+  \\
   \     /     \    +----+   //
  -----     ----|RB4|===
       +---+
```

Figure C.2: Known Unicast Multipath

When multipathing is used, frames that follow different paths will be subject to different delays and may be re-ordered. While some traffic may be order/delay insensitive, typically most traffic consists of flows of frames such the re-ordering within a flow is damaging. How to determine flows or what granularity flows should have is beyond the scope of this document but, as an example, under many circumstances it would be safe to consider all the frames flowing between a particular pair of end station ports to be a flow.
Appendix D: Determination of VLAN and Priority

A high level informative summary of how VLAN ID and priority are determined for incoming native frames, omitting some details, is given in the bulleted items below:

- When an untagged native frame arrives, a zero configuration RBridge associates the default priority zero and the VLAN ID 1 with it. It actually sets the VLAN for the untagged frame to be the "port VLAN ID" associated with that port. The port VLAN ID defaults to VLAN ID 1 but may be configured to be any other VLAN ID. An RBridge may also be configured on a per port basis to discard such frames or to associate a different priority code point with them. Determination of the VLAN ID associated with an incoming untagged non-control frame may also be made dependent on the Ethertype or NSAP (referred to in 802.1 as the Protocol) of the arriving frame, the source MAC address, or other local rules.

- When a priority tagged native frame arrives, a zero configuration RBridge associates with it both the port VLAN ID, which defaults to 1, and the priority code point provided in the priority tag in the frame. An RBridge may be configured on a per port basis to discard such frames or to associate them with a different VLAN ID as described in the point immediately above. It may also be configured to map the priority code point provided in the frame by specifying, for each of the eight possible values that might be in the frame, what actual priority code point will be associated with the frame by the RBridge.

- When a C-tagged (formerly called Q-tagged) native frame arrives, a zero configuration RBridge associates with it the VLAN ID and priority in the C-tag. An RBridge may be configured on a per port per VLAN basis to discard such frames. It may also be configured on a per port basis to map the priority value as specified above for priority tagged frames.

In 802.1, the process of associating a priority code point with a frame, including mapping a priority provided in the frame to another priority, is referred to as priority "regeneration".
Appendix Z: Revision History

RFC Editor: Please delete this Appendix Z before publication.

Changes from -03 to -04

1. Divide IANA Considerations section into IANA and IEEE parts. Add IANA considerations for TRILL Header variations and reserved bit and normative references to RFCs 2434 and 4020.

2. Add note on the terms Rbridge and TRILL to section 1.2.

3. Remove IS-IS marketing text.

4. Split Section 3 into Sections 3 and 4. Add a new top level section "5. Pseudo Code", renumbering following sections. Move pseudo code that was in old Section 3 into Section 5 and make section 3 more textural. This idea is that Section 3 and 4 have more readable text descriptions with some corner cases left out for simplicity while section 5 has more structured and complete coverage.

5. Revised and extended Security Considerations section.

6. Move multicast router attachment bit and IGMP membership report information from the per-VLAN IS-IS instance to the core IS-IS instance so the information can be used by core R Bridges to prune distribution trees.

7. Remove ARP/ND optimization.

8. Change TRILL Header to add option feature. Add option section.

9. Change TRILL Header to expand Version field to the Variation field. Add TRILL message variations (8 bits) supported to the per R Bridge link state information.

10. Distinguish TRILL data and IS-IS messages by using Variation = 0 and 1.

11. Consistently state that VLAN pruning and IP derived multicast pruning of distribution trees are SHOULD.

12. Add text and pseudo code to discard TRILL Ethertype data frames received on a port that does not have an IS-IS adjacency on it.

13. Add end station address learning section. Specify end station address learning from decapsulated native frames.
14. Add nickname allocation priority and optional nickname configuration. Reserve nickname values zero and 0xFFFF.

15. Explain about multiple Designated RBridges because of multiple VLANS.

16. Add Incremental Deployment Considerations Section incorporating expanded Wiring Closet Topology Section.

17. Add more detail on VLAN tag information and material on frame priority.

18. Miscellaneous minor editing and terminology updates.

Changes from -04 to -05

NOTE: Section 5 was NOT updated as indicated below but the remainder of the draft was so updated.

1. Mention optional VLAN and multicast optimization in Abstract.

2. Change to distinguish TRILL IS-IS from TRILL data frames based on the Inner.MacDA instead of a TRILL Header bit.

3. Split IP multicast router attached bit in two so you can separately indicate attachment of IPv4 and IPv6 routers. Provide that these bits must be set if an RBridge does not actually do multicast control snooping on ingressed traffic.

4. Add the term "port VLAN ID" (PVID).

5. Drop references to PIM. Improve discussions of IGMP, MLD, and MRD messages.

6. Move M bit over one and create two bit pruning field at the bottom of the "V" combined field.

7. Add pruning control values of V and discussion of same.

8. Permit optional unicast transmission of multi-destination frames when there is only one received out a port.

9. Miscellaneous minor editing and terminology updates.
Changes from -05 to -06

1. Revise Section 2 discussion of DRB determination in the presence of VLANs and move it to Section 2.2. Adjust VLAN handling description.

2. Change "V" field to be a 2-bit version fields followed by 2 reserved bits. Make corresponding changes to eliminate the inclusion in the header of frame analysis indicating type of multi-destination pruning which is proper for frame. Thus all non-ingress RBridges that wish to perform such pruning are forced to do full frame analysis. Make further corresponding changes in IANA Considerations.

3. The Inner.MacDA for TRILL IS-IS frames is changed to a second multicast address: All-IS-IS-RBridges. IEEE Allocation Considerations, etc., are correspondingly changed.

4. Note in Section 6 that bridges can hide slow links and generally make it harder from RBridges to determine the cost of an RBridge to RBridge hop that is a bridged LAN.

5. Add material noting that replacement of bridges by RBridges can cause connectivity between previously isolated islands of the same VLAN.

6. Expand Security Considerations by mentioning RFC 3567 and indicating that TRILL enveloping may reduce the effectiveness of TRILL-ignorant firewall functionality.

7. Extensive updates to pseudo code.

8. Change to one DRB per physical link which dictates the inter-RBridge VLAN for the link, appoints forwarders per-VLAN, can be configured to send Hellos on multiple VLANs, etc.

9. Add a minimal management by SNMP statement to Section 2.

10. Delete explicit requirement to process TRILL frames arriving on a port even if the port implements spanning tree and is in spanning tree blocked state.

11. Miscellaneous minor editing and terminology updates.

Changes from -06 to -07

[WARNING: Section 5 of draft -07 was not fully updated to incorporate the changes below.]
1. Drop recommendation to set "bridge" flags in some 802.1AB frame fields.

2. Add Section 2.5 giving an informative description of zero configuration behavior for 802.1D and 802.1Q bridges and RBridges.

3. Add Section 4.7 (renumbering the former 4.7 to be 4.9) on the receipt, handing, and transmission of MVRP and other MRP frames by RBridges. Add references to 802.1ak.

4. Add Section 4.8 on Multipathing.

5. Partial changes to Section 5 to correspond with changes elsewhere in the draft.

6. Addition of frame category definitions in Section 1.2.

7. Addition of Section 10, Acronyms.

8. Add note in Section 6.2 that difficult in link cost determination due to intervening devices is not confined to RBridges.

9. Re-ordered some sections in Section 6.

10. Added a paragraph about taking care if trying to use VLANs for security to Security Considerations Section and re-ordered paragraphs in that section.

11. Added mention of being able to configure a port so that native frames are not send and are dropped on receipt. Probably need to say more about this.

12. Remove material about pseudo node suppression.

13. Fix a few cases where hop count was off by one.

14. Add option critical bits when option area length non-zero.

16. Miscellaneous minor editing and terminology updates. Changed Figure numbers to be relative to major section. Added Table captions.

Changes from -07 to -08

1. Add "low" and "high" level control frame definitions to Section 1.2 and note concerning frames which would qualify as both
"TRILL" and "control" frames. Utilize these defined frame types more consistently through the document.

2. Move substantial areas of tutorial, motivational, and informational text to Appendices, or a separate document, including Sections numbered 2.5, 4.8, 6.3, and 6.4 in version -07. Remove pseudo-code (Section 5 in version -07).

3. Move link Hellos / VLAN specification and discussion to a new subsection of Section 4.

4. Replace distribution tree root flag per RBridge with new logic which orders all RBridges in a campus as to their priority to be a distribution tree root and provides for the highest priority distribution tree root to dictate the numbers of trees in the campus. RBridges use the tree with least cost from themselves to the tree root for multi-destination frame distribution, or the n such trees if they multi-path multi-destination traffic.

5. Add "Access" port configuration bit and Appendix on Trunk and Access Links.

6. Add statement that use of S-tags in TRILL is outside the scope of this document.

7. Add new section on RBridge port structure (Section 4.7) which includes discussion of RBridge interactions with BPDUs and revised interactions with VRP frames. Make provisions for dynamic VLAN registration a "MAY" implement and agnostic between GVRP and MVRP. Remove references to 802.1ak. Simplify text related to VRP. Remove related configuration option.

8. Add requirement to adjust input filters no later than output forwarding.

9. Add requirement for configurable (default 30 second) prohibition on RBridge decapsulation out a port if a root bridge change has just been observed on that port.

10. Add provisions for propagating topology change to attached bridged LAN when an RBridge is de-appointed forwarder. Also other end station addressing forgetting details including per VLAN forwarding status dropped counter.

11. Delete requirement that appointed forwarder wait until it has received all the LSPs listed in the first CSNP (if any) it has received from its neighbors before forwarding frames off a link.

12. Add explicit criterion for when an RBridge port defers to the DRB indicated in a Hello it receives even if that Hello is not from
the DRB or even from an RBridge in direct communication with the DRB.


14. Update reference to RFC 2434 to be to RFC 5226.

15. Miscellaneous minor editing and terminology updates. Add Figures index after Table of Contents.

Changes from -08 to -09

1. Specify SHOULD as the implementation requirement for SNMPv3 management.

2. Change default confidence level to 0x20 for addresses learned from observing locally received native frames and from decapsulating TRILL data frames. This provides more space for lower confidence levels.

3. Add security consideration for observation of traffic no longer constrained to links in its Inner.VLAN due to TRILL encapsulation.

4. Updated bridge configuration assumptions in Section 2.3.1.

5. Use "inhibited" to describe the status of an appointed forwarder when it is temporarily discarding all received native frames and not sending any native frames.

6. In Section 4.3, there was an implication that the priority to be a tree root and the number of trees to be computed had not only default values for a zero configuration RBridge but could also be individually present or absent in the LSP for the RBridge. This tends to lead to a variable-length sub-TLV or multiple sub-TLVs in the LSP which leads to additional code paths to test. So various "if advertised" conditional clauses have been removed.

7. Reserve nicknames 0xFFC0 through 0xFFFFE as well as 0x0000 and 0xFFFFF and provide IANA Considerations for their allocation.

8. Improve Figure 4.1, "TRILL Data Encapsulation over Ethernet" by generalizing it and adding an RBridge diagram.

9. Add "access port" bit to Hello. Extend and clarify behavior for access ports and for the occurrence of the IS Neighbor TLV in TRILL Hellos.
10. Miscellaneous minor editing.

Changes from -09 to -10

1. Split Section 2.4 into two subsections inserting 2.4.1 with a simplified RBridge port diagram and discussion of how RBridges mostly use the mechanisms of IEEE 802.1Q bridges below the EISS layer.

2. Remove the "SHOULD" requirement that the hop count for multi-destination frames not be set by the ingress RBridge in excess of the distance through the distribution tree to the most remote RBridge.

3. Remove any implication that addresses received by ESADI are always better than those learned from the data plane.

4. Rephrase language concerning the case where a known unicast native frame in receive by an RBridge to be output in native form on another link of that RBridge so that instead of describing this as logically forwarding the frame in native form it is described as logically encapsulating and then decapsulating the frame.

5. Remove language saying that a TRILL Ethertype frame with a broadcast outer destination address MAY be treated as if its outer destination address was All-RBridges.

6. Clarify that all TRILL data frames with unknown or reserved egress nicknames are discarded.

7. Substantially expand Figure 4.3 at the upper port layers and correspondingly expand the accompanying text which is now Section 4.7.2.

8. Change TRILL IS-IS frames so they are no longer encapsulated but have the All-IS-IS-RBridges Outer.MacDA. Change the Inner.MacDA of ESADI frames to be the new All-ESADI-RBridges multicast address.

9. Update reference to RFC 3567 to be to RFC 5304.

10. Miscellaneous minor editing changes.
Disclaimer

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Additional IPR Provisions

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.
Authors’ Addresses

Radia Perlman
Sun Microsystems
16 Network Circle
Menlo Park, CA 94025

Phone: +1-650-960-1300
Email: Radia.Perlman@sun.com

Donald E. Eastlake, 3rd
Eastlake Enterprises
155 Beaver Street
Milford, MA 01757 USA

Phone: +1-508-634-2066
Email: d3e3e3@gmail.com

Dinesh G. Dutt
Cisco Systems
170 Tasman Drive
San Jose, CA 95134-1706 USA

Phone: +1-408-527-0955
Email: dduitt@cisco.com

Silvano Gai
Nuova Systems
2600 San Tomas Expressway
Santa Clara, CA 95051 USA

Phone: +1-408-387-6123
Email: sgai@nuovasystems.com

Anoop Ghanwani
Brocade
1745 Technology Drive
San Jose, CA 95110 USA

Phone: +1-408-333-7149
Email: anoop@brocade.com