                Transparent Interconnection of Lots of Links (TRILL):
                       Clarifications, Corrections, and Updates

Abstract

   Since the publication of the TRILL (Transparent Interconnection of
   Lots of Links) base protocol in 2011, active development and
   deployment of TRILL have revealed errata in RFC 6325 and areas that
   could use clarifications or updates. RFC 7177, RFC 7357, and an
   intended replacement of RFC 6439 provide clarifications and updates
   with respect to adjacency, the TRILL ESADI (End Station Address
   Distribution Information) protocol, and Appointed Forwarders,
   respectively. This document provides other known clarifications,
   corrections, and updates. It obsoletes RFC 7180 (the previous "TRILL
   clarifications, corrections, and updates" RFC), and it updates RFCs
   6325, 7177, and 7179.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc7780.

Copyright Notice

Table of Contents

1. Introduction (Changed)

   Since the TRILL base protocol [RFC6325] was published in 2011, active
   development and deployment of TRILL have revealed errors in the
   specification [RFC6325] and several areas that could use
   clarifications or updates.

   [RFC7177], [RFC7357], and [RFC6439bis] provide clarifications and
   updates with respect to adjacency, the TRILL ESADI (End Station
   Address Distribution Information) protocol, and Appointed Forwarders,
   respectively. This document provides other known clarifications,
   corrections, and updates to [RFC6325], [RFC7177], and [RFC7179].
   This document obsoletes [RFC7180] (the previous TRILL
   "clarifications, corrections, and updates" document), updates
   [RFC6325], updates [RFC7177] as described in Section 9, and updates
   [RFC7179] as described in Sections 10.2 and 10.3. The changes to
   these RFCs are summarized in Appendix C.

   Sections of this document are annotated as to whether they are "New"
   technical material, material that has been technically "Changed", or
   material that is technically "Unchanged", by the appearance of one of
   these three words in parentheses at the end of the section header. A
   section with only editorial changes is annotated as "(Unchanged)".
   If no such notation appears, then the first notation encountered on
   going to successively higher-level section headers (those with
   shorter section numbers) applies. Appendix C describes changes,
   summarizes material added, and lists material deleted.

1.1. Precedence (Changed)

   In the event of any conflicts between this document and [RFC6325],
   [RFC7177], or [RFC7179], this document takes precedence.

   In addition, Section 1.2 of [RFC6325] ("Normative Content and
   Precedence") is updated to provide a more complete precedence
   ordering of the sections of [RFC6325], as shown below, where sections
   to the left take precedence over sections to their right. There are
   no known conflicts between these sections; however, Sections 1 and 2
   are less detailed and do not mention every corner case, while
   subsequent sections of [RFC6325] are more detailed. This precedence
   is specified as a fallback in case some conflict is found in the
   future.

                     4 > 3 > 7 > 5 > 2 > 6 > 1

1.2. Changes That Are Not Backward Compatible (Unchanged)

   The change made by Section 3.4 below (unchanged from Section 3.4 of
   [RFC7180]) is not backward compatible with [RFC6325] but has
   nevertheless been adopted to reduce distribution tree changes
   resulting from topology changes.

   Several other changes herein that are fixes to errata for [RFC6325]
   -- [Err3002], [Err3003], [Err3004], [Err3052], [Err3053], and
   [Err3508] -- may not be backward compatible with previous
   implementations that conformed to errors in the specification.

1.3. Terminology and Acronyms (Changed)

   This document uses the acronyms defined in [RFC6325], some of which
   are repeated below for convenience, along with some additional
   acronyms and terms, as follows:

   BFD - Bidirectional Forwarding Detection.

   Campus - A TRILL network consisting of TRILL switches, links, and
      possibly bridges bounded by end stations and IP routers. For
      TRILL, there is no "academic" implication in the name "campus".

   CFI - Canonical Format Indicator [802].

   CSNP - Complete Sequence Number PDU.

   DEI - Drop Eligibility Indicator [802.1Q-2014].

   FGL - Fine-Grained Labeling [RFC7172].

   FS-LSP - Flooding Scope LSP.

   OOMF - Overload Originated Multi-destination Frame.

   P2P - Point-to-point.

   PDU - Protocol Data Unit.

   PSNP - Partial Sequence Number PDU.

   RBridge - Routing Bridge, an alternative name for a TRILL switch.

   RPFC - Reverse Path Forwarding Check.

   SNPA - Subnetwork Point of Attachment (for example, Media Access
      Control (MAC) address).

ToS - Type of Service.

TRILL - Transparent Interconnection of Lots of Links or Tunneled
   Routing in the Link Layer.

TRILL switch - A device implementing the TRILL protocol. An
   alternative name for an RBridge.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in
[RFC2119].

In this document, a "packet" usually refers to a TRILL Data packet or
TRILL IS-IS packet received from or sent to a TRILL switch, while a
"frame" usually refers to a native frame being received from or sent
to an end station. (The word "frame" also occurs in other contexts,
such as the "Frame Check Sequence" that is at the end of Ethernet
transmissions.)

2. Overloaded and/or Unreachable RBridges (Unchanged)

In this section, the term "neighbor" refers only to actual RBridges
and ignores pseudonodes.

RBridges may be in overload, as indicated by the [IS-IS] overload
flag in their LSPs (Link State PDUs). This means that either (1)
they are incapable of holding the entire link-state database and thus
do not have a view of the entire topology or (2) they have been
configured to have the overload bit set. Although networks should be
engineered to avoid actual link-state overload, it might occur under
various circumstances -- for example, if a very large campus included
one or more low-end TRILL switches.

It is a common operational practice to set the overload bit in an
[IS-IS] router (such as a TRILL switch) when performing maintenance
on that router that might affect its ability to correctly forward
packets; this will usually leave the router reachable for maintenance
traffic, but transit traffic will not be routed through it. (Also,
in some cases, TRILL provides for setting the overload bit in the
pseudonode of a link to stop TRILL Data traffic on an access link
(see Section 4.9.1 of [RFC6325]).)

[IS-IS] and TRILL make a reasonable effort to do what they can, even
if some TRILL switches/routers are in overload. They can do
reasonably well if a few scattered nodes are in overload. However,
actual least-cost paths are no longer assured if any TRILL switches
are in overload.

For the effect of overload on the appointment of forwarders, see
[RFC6439bis].

2.1. Reachability

Packets are not least-cost routed through an overloaded TRILL switch,
although they may originate or terminate at an overloaded TRILL
switch. In addition, packets will not be least-cost routed over
links with cost 2**24 - 1 [RFC5305]; such links are reserved for
traffic-engineered packets, the handling of which is beyond the scope
of this document.

As a result, a portion of the campus may be unreachable for
least-cost routed TRILL Data because all paths to it would be either
through a link with cost 2**24 - 1 or through an overloaded RBridge.
For example, an RBridge (TRILL switch) RB1 is not reachable by TRILL
Data if all of its neighbors are connected to RB1 by links with cost
2**24 - 1. Such RBridges are called "data unreachable".

The link-state database at an RBridge -- for example, RB1 -- can also
contain information on TRILL switches that are unreachable by IS-IS
link-state flooding due to link or RBridge failures. When such
failures partition the campus, the TRILL switches adjacent to the
failure and on the same side of the failure as RB1 will update their
LSPs to show the lack of connectivity, and RB1 will receive those
updates. As a result, RB1 will be aware of the partition. Nodes on
the far side of the partition are both IS-IS unreachable and data
unreachable from RB1. However, LSPs held by RB1 for TRILL switches
on the far side of the failure will not be updated and may stay
around until they time out, which could be tens of minutes or longer.
(The default in [IS-IS] is twenty minutes.)

2.2. Distribution Trees

An RBridge in overload cannot be trusted to correctly calculate
distribution trees or correctly perform the RPFC (Reverse Path
Forwarding Check). Therefore, it cannot be trusted to forward
multi-destination TRILL Data packets. It can only appear as a leaf
node in a TRILL multi-destination distribution tree. Furthermore, if
all the immediate neighbors of an RBridge are overloaded, then it is
omitted from all trees in the campus and is unreachable by
multi-destination packets.

When an RBridge determines what nicknames to use as the roots of the
distribution trees it calculates, it MUST ignore all nicknames held
by TRILL switches that are in overload or are data unreachable. When
calculating RPFCs for multi-destination packets, an RBridge such as
RB1 MAY, to avoid calculating unnecessary RPFC state information,

   ignore any trees that cannot reach RB1, even if other RBridges list
   those trees as trees that other TRILL switches might use. (However,
   see Section 3.)

2.3. Overloaded Receipt of TRILL Data Packets

   The receipt of TRILL Data packets by overloaded RBridge RB2 is
   discussed in the subsections below. In all cases, the normal
   Hop Count decrement is performed, and the TRILL Data packets are
   discarded if the result is less than one or if the Egress Nickname is
   illegal.

2.3.1. Known Unicast Receipt

   RB2 will not usually receive unicast TRILL Data packets unless it is
   the egress, in which case it egresses and delivers the data normally.
   If RB2 receives a unicast TRILL Data packet for which it is not the
   egress, perhaps because a neighbor does not yet know it is in
   overload, RB2 MUST NOT discard the packet because the egress is an
   unknown nickname, as it might not know about all nicknames due to its
   overloaded condition. If any neighbor other than the neighbor from
   which it received the packet is not overloaded, it MUST attempt to
   forward the packet to one of those neighbors selected at random
   [RFC4086]. If there is no such neighbor, the packet is discarded.

2.3.2. Multi-Destination Receipt

   If RB2 in overload receives a multi-destination TRILL Data packet,
   RB2 MUST NOT apply an RPFC because, due to overload, it might not do
   so correctly. RB2 egresses and delivers the frame locally where it
   is Appointed Forwarder for the frame's VLAN (or, if the packet is
   FGL, for the VLAN that FGL maps to at the port), subject to any
   multicast pruning. But because, as stated above, RB2 can only be the
   leaf of a distribution tree, it MUST NOT forward a multi-destination
   TRILL Data packet (except as an egressed native frame where RB2 is
   Appointed Forwarder).

2.4. Overloaded Origination of TRILL Data Packets

   Overloaded origination of unicast TRILL Data packets with known
   egress and of multi-destination packets is discussed in the
   subsections below.

2.4.1. Known Unicast Origination

   When RB2, an overloaded RBridge, ingresses or creates a known
   destination unicast data packet, it delivers it locally if the
   destination is local. Otherwise, RB2 unicasts it to any neighbor
   TRILL switch that is not overloaded. It MAY use what routing
   information it has to help select the neighbor.

2.4.2. Multi-Destination Origination

   Overloaded RBridge RB2 ingressing or creating a multi-destination
   data packet presents a more complex scenario than that of the known
   unicast case, as discussed below.

2.4.2.1. An Example Network

   For example, consider the network diagram below in which, for
   simplicity, end stations and any bridges are not shown. There is one
   distribution tree of which RB4 is the root, as represented by double
   lines. Only RBridge RB2 is overloaded.

```
          +-----+     +-----+      +-----+      +-----+
          | RB7 +====+ RB5 +=====+ RB3 +=====+ RB1 |
          +-----+     +--+--+      +-++--+      +--+--+
                         |           ||            |
                      +--+---+        ||            |
                   +------+RB2(ov)|======++         |
                   |      +-------+        ||        |
                   |                       ||        |
          +--+--+     +-----+   ++==++=++     +--+--+
          | RB8 +====+ RB6 +===++ RB4 ++=====+ RB9 |
          +-----+     +-----+   ++=====++     +-----+
```
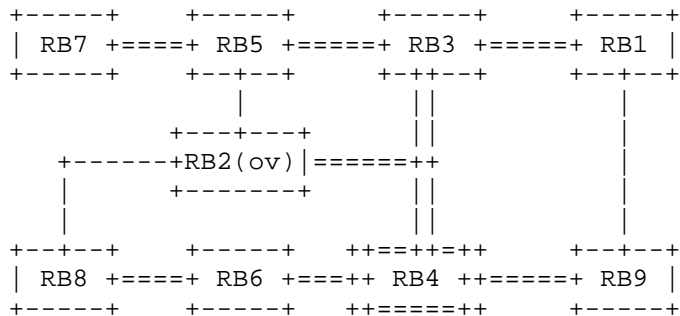
   Since RB2 is overloaded, it does not know what the distribution tree
   or trees are for the network. Thus, there is no way it can provide
   normal TRILL Data service for multi-destination native frames. So,
   RB2 tunnels the frame in a TRILL Data packet to a neighbor that is
   not overloaded if it has such a neighbor that has signaled that it is
   willing to offer this service. RBridges indicate this in their
   Hellos as described below. This service is called the OOMF (Overload
   Originated Multi-destination Frame) service.

   - The multi-destination frame MUST NOT be locally distributed in
     native form at RB2, because this would cause the frame to be
     delivered twice. Instead, it is tunneling to a neighbor as
     described in this section. For example, if RB2 locally distributed
     a multicast native frame and then tunneled it to RB5, RB2 would get
     a copy of the frame when RB3 transmitted it as a TRILL Data packet

      on the multi-access RB2-RB3-RB4 link. Since RB2 would, in general,
      not be able to tell that this was a frame it had tunneled for
      distribution, RB2 would decapsulate it and locally distribute it a
      second time.

   - On the other hand, if there is no neighbor of RB2 offering RB2 the
      OOMF service, RB2 cannot tunnel the frame to a neighbor. In this
      case, RB2 MUST locally distribute the frame where it is Appointed
      Forwarder for the frame's VLAN and optionally subject to multicast
      pruning.

2.4.2.2. Indicating OOMF Support

   An RBridge RB3 indicates its willingness to offer the OOMF service to
   RB2 in the TRILL Neighbor TLV in RB3's TRILL Hellos by setting a bit
   associated with the SNPA (Subnetwork Point of Attachment, also known
   as MAC address) of RB2 on the link (see the IANA Considerations
   section). Overloaded RBridge RB2 can only distribute
   multi-destination TRILL Data packets to the campus if a neighbor of
   RB2 not in overload offers RB2 the OOMF service. If RB2 does not
   have OOMF service available to it, RB2 can still receive
   multi-destination packets from non-overloaded neighbors, and if RB2
   should originate or ingress such a frame, it distributes it locally
   in native form.

2.4.2.3. Using OOMF Service

   If RB2 sees this OOMF (Overload Originated Multi-destination Frame)
   service advertised for it by any of its neighbors on any link to
   which RB2 connects, it selects one such neighbor by a means that is
   beyond the scope of this document. Assuming that RB2 selects RB3 to
   handle multi-destination packets it originates, RB2 MUST advertise in
   its LSP that it might use any of the distribution trees that RB3
   advertises so that the RPFC will work in the rest of the campus.
   Thus, notwithstanding its overloaded state, RB2 MUST retain this
   information from RB3 LSPs, which it will receive, as it is directly
   connected to RB3.

   RB2 then encapsulates such frames as TRILL Data packets to RB3 as
   follows: "M" bit = 0; Hop Count = 2; Ingress Nickname = a nickname
   held by RB2; and, since RB2 cannot tell what distribution tree RB3
   will use, Egress Nickname = a special nickname indicating an OOMF
   packet (see the IANA Considerations section). RB2 then unicasts this
   TRILL Data packet to RB3. (Implementation of Item 4 in Section 4
   below provides reasonable assurance that, notwithstanding its
   overloaded state, the ingress nickname used by RB2 will be unique
   within at least the portion of the campus that is IS-IS reachable
   from RB2.)

On receipt of such a packet, RB3 does the following:

- changes the Egress Nickname field to designate a distribution tree
  that RB3 normally uses,

- sets the "M" bit to one,

- changes the Hop Count to the value it would normally use if it were
  the ingress, and

- forwards the TRILL Data packet on that tree.

RB3 MAY rate-limit the number of packets for which it is providing
this service by discarding some such packets from RB2. The provision
of even limited bandwidth for OOMFs by RB3, perhaps via the slow
path, may be important to the bootstrapping of services at RB2 or at
end stations connected to RB2, such as supporting DHCP and ARP/ND
(Address Resolution Protocol / Neighbor Discovery). (Everyone
sometimes needs a little OOMF (pronounced "oomph") to get off the
ground.)

3. Distribution Trees and RPF Check (Changed)

Two corrections, a clarification, and two updates related to
distribution trees appear in the subsections below, along with an
alternative, stronger RPF (Reverse Path Forwarding) check. See also
Section 2.2.

3.1. Number of Distribution Trees (Unchanged)

In [RFC6325], Section 4.5.2, page 56, point 2, fourth paragraph, the
parenthetical "(up to the maximum of {j,k})" is incorrect [Err3052].
It should read "(up to k if j is zero or the minimum of (j, k) if j
is non-zero)".

3.2. Distribution Tree Update Clarification (Unchanged)

When a link-state database change causes a change in the distribution
tree(s), several possible types of change can occur. If a tree root
remains a tree root but the tree changes, then local forwarding and
RPFC entries for that tree should be updated as soon as practical.
Similarly, if a new nickname becomes a tree root, forwarding and RPFC
entries for the new tree should be installed as soon as practical.
However, if a nickname ceases to be a tree root and there is
sufficient room in local tables, the forwarding and RPFC entries for
the former tree MAY be retained so that any multi-destination TRILL
Data packets already in flight on that tree have a higher probability
of being delivered.

3.3. Multicast Pruning Based on IP Address (Unchanged)

   The TRILL base protocol specification [RFC6325] provides for, and
   recommends the pruning of, multi-destination packet distribution
   trees based on the location of IP multicast routers and listeners;
   however, multicast listening is identified by derived MAC addresses
   as communicated in the Group MAC Address sub-TLV [RFC7176].

   TRILL switches MAY communicate multicast listeners and prune
   distribution trees based on the actual IPv4 or IPv6 multicast
   addresses involved. Additional Group Address sub-TLVs are provided
   in [RFC7176] to carry this information. A TRILL switch that is only
   capable of pruning based on derived MAC addresses SHOULD calculate
   and use such derived MAC addresses from the multicast listener IPv4
   or IPv6 address information it receives.

3.4. Numbering of Distribution Trees (Unchanged)

   Section 4.5.1 of [RFC6325] specifies that, when building distribution
   tree number j, node (RBridge) N that has multiple possible parents in
   the tree is attached to possible parent number j mod p. Trees are
   numbered starting with 1, but possible parents are numbered starting
   with 0. As a result, if there are two trees and two possible
   parents, then in tree 1 parent 1 will be selected, and in tree 2
   parent 0 will be selected.

   This is changed so that the selected parent MUST be (j-1) mod p. As
   a result, in the case above, tree 1 will select parent 0, and tree 2
   will select parent 1. This change is not backward compatible with
   [RFC6325]. If all RBridges in a campus do not determine distribution
   trees in the same way, then for most topologies, the RPFC will drop
   many multi-destination packets before they have been properly
   delivered.

3.5. Link Cost Directionality (Unchanged)

   Distribution tree construction, like other least-cost aspects of
   TRILL, works even if link costs are asymmetric, so the cost of the
   hop from RB1 to RB2 is different from the cost of the hop from RB2 to
   RB1. However, it is essential that all RBridges calculate the same
   distribution trees, and thus all must use either the cost away from
   the tree root or the cost towards the tree root. The text in
   Section 4.5.1 of [RFC6325] is incorrect, as documented in [Err3508].
   The text says:

      In other words, the set of potential parents for N, for the tree
      rooted at R, consists of those that give equally minimal cost
      paths from N to R and ...

   but the text should say "from R to N":

      In other words, the set of potential parents for N, for the tree
      rooted at R, consists of those that give equally minimal cost
      paths from R to N and ...

3.6. Alternative RPF Check (New)

   [RFC6325] mandates a Reverse Path Forwarding (RPF) check on
   multi-destination TRILL Data packets to avoid possible multiplication
   and/or looping of multi-destination traffic during TRILL campus
   topology transients. This check is logically performed at each TRILL
   switch input port and determines whether it is arriving on the
   expected port based on where the packet started (the ingress
   nickname) and the tree on which it is being distributed. If not, the
   packet is silently discarded. This check is fine for point-to-point
   links; however, there are rare circumstances involving multi-access
   ("broadcast") links where a packet can be duplicated despite this
   RPF check and other checks performed by TRILL.

   Section 3.6.1 gives an example of the potential problem, and
   Section 3.6.2 specifies a solution. This solution is an alternative,
   stronger RPF check that TRILL switches can implement in place of the
   RPF check discussed in [RFC6325].

3.6.1. Example of the Potential Problem

   Consider this network:

           F--A--B--C--o--D
                       |
                       E
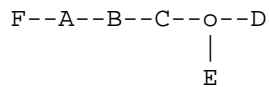
   All the links except the link between C, D, and E are point-to-point
   links. C, D, and E are connected over a broadcast link represented
   by the pseudonode "o". For example, they could be connected by a
   bridged LAN. (Bridged LANs are transparent to TRILL.)

   Although the choice of root is unimportant here, assume that D or F
   is chosen as the root of a distribution tree so that it is obvious
   that the tree looks just like the diagram above.

Now assume that a link comes up from A to the same bridged LAN. The
network then looks like this:

```
            +--------+
            |        |
       F--A--B--C--o--D
                  |
                  E
```
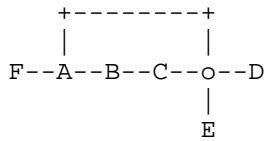
Let's say the resulting tree in steady state includes all links
except the B-C link. After the network has converged, a packet that
starts from F will go F->A. Then A will send one copy on the A-B
link and another copy into the bridged LAN from which it will be
received by C and D.

Now consider a transition stage where A and D have acted on the new
LSPs and programmed their forwarding plane, while B and C have not
yet done so. This means that B and C both consider the link between
them to still be part of the tree. In this case, a packet that
starts out from F and reaches A will be copied by A into the A-B link
and to the bridged LAN. D's RPF check says to accept packets on this
tree coming from F over its port on the bridged LAN, so it gets
accepted. D is also adjacent to A on the tree, so the tree adjacency
check, a separate check mandated by [RFC6325], also passes.

However, the packet that gets to B gets sent out by B to C. C's RPF
check still has the old state, and it thinks the packet is OK. C
sends the packet along the old tree, which sends the packet into the
bridged LAN. D receives one more packet, but the tree adjacency
check passes at D because C is adjacent to D in the new tree as well.
The RPF check also passes at D because D's port on the bridged LAN is
OK for receiving packets from F.

So, during this transient state, D gets duplicates of every
multi-destination packet ingressed at F (unless the packet gets
pruned) until B and C act on the new LSPs and program their
forwarding tables.

3.6.2. Solution and Discussion

The problem stems from the RPF check described in [RFC6325] depending
only on the port at which a TRILL Data packet is received, the
ingress nickname, and the tree being used, that is, a check if
{ingress nickname, tree, input port} is a valid combination according
to the receiving TRILL switch's view of the campus topology. A
multi-access link actually has multiple adjacencies overlaid on one
physical link, and to avoid the problem shown in Section 3.6.1, a
stronger check is needed that includes the Layer 2 source address of

the TRILL Data packet being received. (TRILL is a Layer 3 protocol, and TRILL switches are true routers that logically strip the Layer 2 header from any arriving TRILL Data packets and add the appropriate new Layer 2 header to any outgoing TRILL Data packet to get it to the next TRILL switch, so the Layer 2 source address in a TRILL Data packet identifies the immediately previous TRILL switch that forwarded the packet.)

What is needed, instead of checking the validity of the triplet {ingress nickname, tree, input port}, is to check that the quadruplet {ingress nickname, source SNPA, tree, input port} is valid (where "source SNPA" (Subnetwork Point of Attachment) is the Outer.MacSA for an Ethernet link). Although it is true that [RFC6325] also requires a check to ensure that a multi-destination TRILL Data packet is from a TRILL switch that is adjacent in the distribution tree being used, this check is separate from the RPF check, and these two independent checks are not as powerful as the single unified check for a valid quadruplet.

```
              _____
             /        \
          RB1 ------ o ----- RB2
             _____/
```
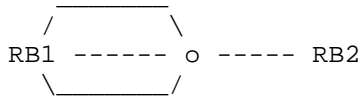
However, this stronger RPF check is not without cost. In the simple case of a multi-access link where each TRILL switch has only one port on the link, it merely increases the size of validity entries by adding the source SNPA (Outer.MacSA). However, assume that some TRILL switch RB1 has multiple ports attached to a multi-access link. In the figure above, RB1 is shown with three ports on the multi-access link. RB1 is permitted to load split multi-destination traffic it is sending into the multi-access link across those ports (Section 4.4.4 of [RFC6325]). Assume that RB2 is another TRILL switch on the link and RB2 is adjacent to RB1 in the distribution tree. The number of validity quadruplets at RB2 for ingress nicknames whose multi-destination traffic would arrive through RB1 is multiplied by the number of ports RB1 has on the access link, because RB2 has to accept such traffic from any such ports. Although such instances seem to be very rare in practice, the number of ports an RBridge has on a link could in principle be tens or even a hundred or more ports, vastly increasing the RPF check state at RB2 when this stronger RPF check is used.

Another potential cost of the stronger RPF check is increased transient loss of multi-destination TRILL Data packets during a topology change. For TRILL switch D, the new stronger RPF check is (tree->A, Outer.MacSA=A, ingress=A, arrival port=if1), while the old one was (tree->A, Outer.MacSA=C, ingress=A, arrival port=if1).

Suppose that both A and B have switched to the new tree for multicast forwarding but D has not updated its RPF check yet; the multicast packet will then be dropped at D's input port, because D still expects a packet from "Outer.MacSA=C". But we do not have this packet loss issue if the weaker triplet check (tree->A, ingress=A, arrival port=if1) is used. Thus, the stronger check can increase the RPF check discard of multi-destination packets during topology transients.

Because of these potential costs, implementation of this stronger RPF check is optional. The TRILL base protocol is updated to provide that TRILL switches MUST, for multi-destination packets, either implement the RPF and other checks as described in [RFC6325] or implement this stronger RPF check as a substitute for the [RFC6325] RPF and tree adjacency checks. There is no problem with a campus having a mixture of TRILL switches, some of which implement one of these RPF checks and some of which implement the other.

4. Nickname Selection (Unchanged)

   Nickname selection is covered by Section 3.7.3 of [RFC6325].
   However, the following should be noted:

   1. The second sentence in the second bullet item in Section 3.7.3 of
      [RFC6325] on page 25 is erroneous [Err3002] and is corrected as
      follows:

      o The occurrence of "IS-IS ID (LAN ID)" is replaced with
         "priority".

      o The occurrence of "IS-IS System ID" is replaced with "7-byte
         IS-IS ID (LAN ID)".

      The resulting corrected sentence in [RFC6325] reads as follows:

         If RB1 chooses nickname x, and RB1 discovers, through receipt
         of an LSP for RB2 at any later time, that RB2 has also chosen
         x, then the RBridge or pseudonode with the numerically higher
         priority keeps the nickname, or if there is a tie in priority,
         the RBridge with the numerically higher 7-byte IS-IS ID
         (LAN ID) keeps the nickname, and the other RBridge MUST select
         a new nickname.

   2. In examining the link-state database for nickname conflicts,
      nicknames held by IS-IS unreachable TRILL switches MUST be
      ignored, but nicknames held by IS-IS reachable TRILL switches
      MUST NOT be ignored even if they are data unreachable.

3. An RBridge may need to select a new nickname, either initially
   because it has none or because of a conflict. When doing so, the
   RBridge MUST consider as available all nicknames that do not
   appear in its link-state database or that appear to be held by
   IS-IS unreachable TRILL switches; however, it SHOULD give
   preference to selecting new nicknames that do not appear to be
   held by any TRILL switch in the campus, reachable or unreachable,
   so as to minimize conflicts if IS-IS unreachable TRILL switches
   later become reachable.

4. An RBridge, even after it has acquired a nickname for which there
   appears to be no conflicting claimant, MUST continue to monitor
   for conflicts with the nickname or nicknames it holds. It does so
   by monitoring any received LSPs that should update its link-state
   database for any occurrence of any of its nicknames held with
   higher priority by some other TRILL switch that is IS-IS reachable
   from it. If it finds such a conflict, it MUST select a new
   nickname, even when in overloaded state. (It is possible to
   receive an LSP that should update the link-state database but does
   not do so due to overload.)

5. In the very unlikely case that an RBridge is unable to obtain a
   nickname because all valid RBridge nicknames (0x0001 through
   0xFFBF inclusive) are in use with higher priority by IS-IS
   reachable TRILL switches, it will be unable to act as an ingress,
   egress, or tree root but will still be able to function as a
   transit TRILL switch. Although it cannot be a tree root, such an
   RBridge is included in distribution trees computed for the campus
   unless all its neighbors are overloaded. It would not be possible
   to send a unicast RBridge Channel message specifically to such a
   TRILL switch [RFC7178]; however, it will receive unicast RBridge
   Channel messages sent by a neighbor to the Any-RBridge egress
   nickname and will receive appropriate multi-destination RBridge
   Channel messages.

5. MTU (Maximum Transmission Unit) (Unchanged)

   MTU values in TRILL are derived from the originatingL1LSPBufferSize
   value communicated in the IS-IS originatingLSPBufferSize TLV [IS-IS].
   The campus-wide value Sz, as described in Section 4.3.1 of [RFC6325],
   is the minimum value of originatingL1LSPBufferSize for the RBridges
   in a campus, but not less than 1470. The MTU testing mechanism and
   limiting LSPs to Sz assure that the LSPs can be flooded by IS-IS and
   thus that IS-IS can operate properly.

If an RBridge knows nothing about the MTU of the links or the
originatingL1LSPBufferSize of other RBridges in a campus, the
originatingL1LSPBufferSize for that RBridge should default to the
minimum of the LSP size that its TRILL IS-IS software can handle and
the minimum MTU of the ports that it might use to receive or transmit
LSPs. If an RBridge does have knowledge of link MTUs or other
RBridge originatingL1LSPBufferSize, then, to avoid the necessity of
regenerating the local LSPs using a different maximum size, the
RBridge's originatingL1LSPBufferSize SHOULD be configured to the
minimum of (1) the smallest value that other RBridges are, or will
be, announcing as their originatingL1LSPBufferSize and (2) a value
small enough that the campus will not partition due to a significant
number of links with limited MTUs. However, as specified in
[RFC6325], in no case can originatingL1LSPBufferSize be less than
1470. In a well-configured campus, to minimize any LSP regeneration
due to resizing, all RBridges will be configured with the same
originatingL1LSPBufferSize.

Section 5.1 below corrects errata in [RFC6325], and Section 5.2
clarifies the meaning of various MTU limits for TRILL Ethernet links.

5.1. MTU-Related Errata in RFC 6325

Three MTU-related errata in [RFC6325] are corrected in the
subsections below.

5.1.1. MTU PDU Addressing

Section 4.3.2 of [RFC6325] incorrectly states that multi-destination
MTU-probe and MTU-ack TRILL IS-IS PDUs are sent on Ethernet links
with the All-RBridges multicast address as the Outer.MacDA [Err3004].
As TRILL IS-IS PDUs, when multicast on an Ethernet link, these
multi-destination MTU-probe and MTU-ack PDUs MUST be sent to the
All-IS-IS-RBridges multicast address.

5.1.2. MTU PDU Processing

   As discussed in [RFC6325] and (in more detail) [RFC7177], MTU-probe
   and MTU-ack PDUs MAY be unicast; however, Section 4.6 of [RFC6325]
   erroneously does not allow for this possibility [Err3003]. It is
   corrected by replacing Item 1 in Section 4.6.2 of [RFC6325] with the
   following text, to which TRILL switches MUST conform:

      1. If the Ethertype is L2-IS-IS and the Outer.MacDA is either
         All-IS-IS-RBridges or the unicast MAC address of the receiving
         RBridge port, the frame is handled as described in
         Section 4.6.2.1.

   The reference to "Section 4.6.2.1" in the above text is to that
   section in [RFC6325].

5.1.3. MTU Testing

   The last two sentences of Section 4.3.2 of [RFC6325] contain errors
   [Err3053]. They currently read as follows:

      If X is not greater than Sz, then RB1 sets the "failed minimum MTU
      test" flag for RB2 in RB1's Hello. If size X succeeds, and X >
      Sz, then RB1 advertises the largest tested X for each adjacency in
      the TRILL Hellos RB1 sends on that link, and RB1 MAY advertise X
      as an attribute of the link to RB2 in RB1's LSP.

   They should read as follows:

      If X is not greater than or equal to Sz, then RB1 sets the "failed
      minimum MTU test" flag for RB2 in RB1's Hello. If size X
      succeeds, and X >= Sz, then RB1 advertises the largest tested X
      for each adjacency in the TRILL Hellos RB1 sends on that link,
      and RB1 MAY advertise X as an attribute of the link to RB2 in
      RB1's LSP.

5.2. Ethernet MTU Values

   originatingL1LSPBufferSize is the maximum permitted size of LSPs
   starting with and including the IS-IS 0x83 "Intradomain Routeing
   Protocol Discriminator" byte. In Layer 3 IS-IS,
   originatingL1LSPBufferSize defaults to 1492 bytes. (This is because,
   in its previous life as DECnet Phase V, IS-IS was encoded using the
   SNAP SAP (Subnetwork Access Protocol Service Access Point) [RFC7042]
   format, which takes 8 bytes of overhead and 1492 + 8 = 1500, the
   classic Ethernet maximum. When standardized by ISO/IEC [IS-IS] to
   use Logical Link Control (LLC) encoding, this default could have been
   increased by a few bytes but was not.)

In TRILL, originatingL1LSPBufferSize defaults to 1470 bytes. This
allows 27 bytes of headroom or safety margin to accommodate legacy
devices with the classic Ethernet maximum MTU, despite headers such
as an Outer.VLAN.

Assuming that the campus-wide minimum link MTU is Sz, RBridges on
Ethernet links MUST limit most TRILL IS-IS PDUs so that PDUz (the
length of the PDU starting just after the L2-IS-IS Ethertype and
ending just before the Ethernet Frame Check Sequence (FCS)) does not
exceed Sz. The PDU exceptions are TRILL Hello PDUs, which MUST NOT
exceed 1470 bytes, and MTU-probe and MTU-ack PDUs that are padded by
an amount that depends on the size being tested (which may
exceed Sz).

Sz does not limit TRILL Data packets. They are only limited by the
MTU of the devices and links that they actually pass through;
however, links that can accommodate IS-IS PDUs up to Sz would
accommodate, with a generous safety margin, TRILL Data packet
payloads of (Sz - 24) bytes, starting after the Inner.VLAN and ending
just before the FCS.

Most modern Ethernet equipment has ample headroom for frames with
extensive headers and is sometimes engineered to accommodate 9 KB
jumbo frames.

6. TRILL Port Modes (Unchanged)

    Section 4.9.1 of [RFC6325] specifies four mode bits for RBridge ports
    but may not be completely clear on the effects of all combinations of
    bits in terms of allowed frame types.

The table below explicitly indicates the effects of all possible
combinations of the TRILL port mode bits. "*" in one of the first
four columns indicates that the bit can be either zero or one. The
remaining columns indicate allowed frame types. The "disable bit"
normally disables all frames; however, as an implementation choice,
some or all low-level Layer 2 control messages can still be sent or
received. Examples of Layer 2 control messages are those control
frames for Ethernet identified in Section 1.4 of [RFC6325] or PPP
link negotiation messages [RFC6361].

D i s a b l e	P 2 P	A c c e s s	T r u n k	Layer 2 Control	Native Ingress Native Egress	TRILL Data LSP SNP MTU	TRILL Hello	P2P Hello
0	0	0	0	Yes	Yes	Yes	Yes	No
0	0	0	1	Yes	No	Yes	Yes	No
0	0	1	0	Yes	Yes	No	Yes	No
0	0	1	1	Yes	No	No	Yes	No
0	1	0	*	Yes	No	Yes	No	Yes
0	1	1	*	Yes	No	No	No	Yes
1	*	*	*	Optional	No	No	No	No

The formal name of the "access bit" above is the "TRILL traffic
disable bit". The formal name of the "trunk bit" is the "end-station
service disable bit" [RFC6325].

7. The CFI/DEI Bit (Unchanged)

In May 2011, the IEEE promulgated IEEE Std 802.1Q-2011, which changed
the meaning of the bit between the priority and VLAN ID bits in the
payload of C-VLAN tags. Previously, this bit was called the CFI
(Canonical Format Indicator) bit [802] and had a special meaning in
connection with IEEE 802.5 (Token Ring) frames. After 802.1Q-2011
and in subsequent versions of 802.1Q -- the most current of which is

   [802.1Q-2014] -- this bit is now the DEI (Drop Eligibility Indicator)
   bit. (The corresponding bit in S-VLAN/B-VLAN tags has always been a
   DEI bit.)

   The TRILL base protocol specification [RFC6325] assumed, in effect,
   that the link by which end stations are connected to TRILL switches
   and the restricted virtual link provided by the TRILL Data packet are
   IEEE 802.3 Ethernet links on which the CFI bit is always zero.
   Should an end station be attached by some other type of link, such as
   a Token Ring link, [RFC6325] implicitly assumed that such frames
   would be canonicalized to 802.3 frames before being ingressed, and
   similarly, on egress, such frames would be converted from 802.3 to
   the appropriate frame type for the link. Thus, [RFC6325] required
   that the CFI bit in the Inner.VLAN, which is shown as the "C" bit in
   Section 4.1.1 of [RFC6325], always be zero.

   However, for TRILL switches with ports conforming to the change
   incorporated in the IEEE 802.1Q-2011 standard, the bit in the
   Inner.VLAN, now a DEI bit, MUST be set to the DEI value provided by
   the port interface on ingressing a native frame. Similarly, this bit
   MUST be provided to the port when transiting or egressing a TRILL
   Data packet. As with the 3-bit Priority field, the DEI bit to use in
   forwarding a transit packet MUST be taken from the Inner.VLAN. The
   exact effect on the Outer.VLAN DEI and priority bits, and whether or
   not an Outer.VLAN appears at all on the wire for output frames, may
   depend on output port configuration.

   TRILL campuses with a mixture of ports, some compliant with versions
   of 802.1Q from IEEE Std 802.1Q-2011 onward and some compliant with
   pre-802.1Q-2011 standards, especially if they have actual Token Ring
   links, may operate incorrectly and may corrupt data, just as a
   bridged LAN with such mixed ports and links would.

8. Other IS-IS Considerations (Changed)

   This section covers Extended Level 1 Flooding Scope (E-L1FS) support,
   control packet priorities, unknown PDUs, the Nickname Flags
   APPsub-TLV, graceful restart, and the Purge Originator
   Identification TLV.

8.1. E-L1FS Support (New)

   TRILL switches MUST support E-L1FS PDUs [RFC7356] and MUST include a
   Scope Flooding Support TLV [RFC7356] in all TRILL Hellos they send
   indicating support for this scope and any other FS-LSP scopes that
   they support. This support increases the number of fragments
   available for link-state information by over two orders of magnitude.
   (See Section 9 for further information on support of the Scope
   Flooding Support TLV.)

   In addition, TRILL switches MUST advertise their support of E-L1FS
   flooding in a TRILL-VER sub-TLV Capability Flag (see [RFC7176] and
   Section 12.2). This flag is used by a TRILL switch, say RB1, to
   determine support for E-L1FS by some remote RBx. The alternative of
   simply looking for an E-L1FS FS-LSP originated by RBx fails because
   (1) RBx might support E-L1FS flooding but is not originating any
   E-L1FS FS-LSPs and (2) even if RBx is originating E-L1FS FS-LSPs
   there might, due to legacy TRILL switches in the campus, be no path
   between RBx and RB1 through TRILL switches supporting E-L1FS
   flooding. If that were the case, no E-L1FS FS-LSP originated by RBx
   could get to RB1.

   E-L1FS will commonly be used to flood TRILL GENINFO TLVs and enclosed
   TRILL APPsub-TLVs [RFC7357]. For robustness, E-L1FS fragment zero
   MUST NOT exceed 1470 bytes in length; however, if such a fragment is
   received that is larger, it is processed normally. It is anticipated
   that in the future some particularly important TRILL APPsub-TLVs will
   be specified as being flooded in E-L1FS fragment zero. TRILL GENINFO
   TLVs MUST NOT be sent in LSPs; however, if one is received in an LSP,
   it is processed normally.

8.1.1. Backward Compatibility

   A TRILL campus might contain TRILL switches supporting E-L1FS
   flooding and legacy TRILL switches that do not support E-L1FS or
   perhaps do not support any [RFC7356] scopes.

   A TRILL switch conformant to this document can always tell which
   adjacent TRILL switches support E-L1FS flooding from the adjacency
   table entries on its ports (see Section 9). In addition, such a
   TRILL switch can tell which remote TRILL switches in a campus support
   E-L1FS by the presence of a TRILL version sub-TLV in that TRILL
   switch's LSP with the E-L1FS support bit set in the Capabilities
   field; this capability bit is ignored for adjacent TRILL switches for
   which only the adjacency table entry is consulted to determine E-L1FS
   support.

      TRILL specifications making use of E-L1FS MUST specify how situations
      involving a mixed TRILL campus of TRILL switches will be handled.

8.1.2. E-L1FS Use for Existing (Sub-)TLVs

      In a campus where all TRILL switches support E-L1FS, all TRILL
      sub-TLVs listed in Section 2.3 of [RFC7176], except the TRILL version
      sub-TLV, MAY be advertised by inclusion in Router Capability or
      MT-Capability TLVs in E-L1FS FS-LSPs [RFC7356]. (The TRILL version
      sub-TLV still MUST appear in an LSP fragment zero.)

      In a mixed campus where some TRILL switches support E-L1FS and some
      do not, then only the following four sub-TLVs of those listed in
      Section 2.3 of [RFC7176] can appear in E-L1FS, and then only under
      the conditions discussed below. In the following list, each sub-TLV
      is preceded by an abbreviated acronym used only in this section of
      this document:

         IV: Interested VLANs and Spanning Tree Roots sub-TLV
         VG: VLAN Group sub-TLV
         IL: Interested Labels and Spanning Tree Roots sub-TLV
         LG: Label Group sub-TLV

      An IV or VG sub-TLV MUST NOT be advertised by TRILL switch RB1 in an
      E-L1FS FS-LSP (and should instead be advertised in an LSP) unless the
      following conditions are met:

      - E-L1FS is supported by all of the TRILL switches that are data
        reachable from RB1 and are interested in the VLANs mentioned in the
        IV or VG sub-TLV, and

      - there is E-L1FS connectivity between all such TRILL switches in the
        campus interested in the VLANs mentioned in the IV or VG sub-TLV
        (connectivity involving only intermediate TRILL switches that also
        support E-L1FS).

      Any IV and VG sub-TLVs MAY still be advertised via core TRILL IS-IS
      LSPs by any TRILL switch that has enough room in its LSPs.

      The conditions for using E-L1FS for the IL and LG sub-TLVs are the
      same as for IV and VG, but with Fine-Grained Labels [RFC7172]
      substituted for VLANs.

         Note, for example, that the above would permit a contiguous subset
         of the campus that supported Fine-Grained Labels and E-L1FS to use
         E-L1FS to advertise IL and LG sub-TLVs, even if the remainder of
         the campus did not support Fine-Grained Labels or E-L1FS.

8.2. Control Packet Priorities (New)

   When deciding what packet to send out a port, control packets used to
   establish and maintain adjacency between TRILL switches SHOULD be
   treated as being in the highest-priority category. This includes
   TRILL IS-IS Hello and MTU PDUs, and possibly other adjacency
   [RFC7177] or link-technology-specific packets. Other control and
   data packets SHOULD be given lower priority so that a flood of such
   other packets cannot lead to loss of, or inability to establish,
   adjacency. Loss of adjacency causes a topology transient that can
   result in reduced throughput; reordering; increased probability of
   loss of data; and, in the worst case, network partition if the
   adjacency is a cut point.

   Other important control packets should be given second-highest
   priority. Lower priorities should be given to data or less important
   control packets.

   Based on the above, control packets can be ordered into priority
   categories as shown below, based on the relative criticality of these
   types of messages, where the most critical control packets relate to
   the core routing between TRILL switches and the less critical control
   packets are closer to "application" information. (There may be
   additional control packets, not specifically listed in any category
   below, that SHOULD be handled as being in the most nearly analogous
   category.) Although few implementations will actually treat these
   four categories with different priority, an implementation MAY choose
   to prioritize more critical messages over less critical. However, an
   implementation SHOULD NOT send control packets in a lower-priority
   category with a priority above those in a higher-priority category
   because, under sufficiently congested conditions, this could block
   control packets in a higher-priority category, resulting in network
   disruption.

      Priority
      Category   Description
      --------   --------------

      4.   Hello, MTU-probe, MTU-ack, and other packets critical
                 to establishing and maintaining adjacency. (Normally
                 sent with highest priority, which is priority 7.)

      3.   LSPs, CSNPs/PSNPs, and other important control packets.

      2.   Circuit scoped FS-LSPs, FS-CSNPs, and FS-PSNPs.

      1.   Non-circuit scoped FS-LSPs, FS-CSNPs, and FS-PSNPs.

8.3. Unknown PDUs (New)

   TRILL switches MUST silently discard [IS-IS] PDUs they receive with
   PDU numbers they do not understand, just as they ignore TLVs and
   sub-TLVs they receive that have unknown Types and sub-Types; however,
   they SHOULD maintain a counter of how many such PDUs have been
   received, on a per-PDU-number basis. (This is not burdensome, as the
   PDU number is only a 5-bit field.)

      Note: The set of valid [IS-IS] PDUs was stable for so long that
         some IS-IS implementations may treat PDUs with unknown PDU
         numbers as a serious error and, for example, an indication that
         other valid PDUs from the sender are not to be trusted or that
         they should drop adjacency to the sender if it was adjacent.
         However, the MTU-probe and MTU-ack PDUs were added by
         [RFC7176], and now [RFC7356] has added three more new PDUs.
         Although the authors of this document are not aware of any
         Internet-Drafts calling for further PDUs, the eventual addition
         of further new PDUs should not be surprising.

8.4. Nickname Flags APPsub-TLV (New)

   An optional Nickname Flags APPsub-TLV within the TRILL GENINFO TLV
   [RFC7357] is specified below.

```
                         1 1 1 1 1 1
         0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
        | Type = NickFlags (6)          |   (2 bytes)
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
        | Length = 4*K                  |   (2 bytes)
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
        |    NICKFLAG RECORD 1           (4 bytes)                     |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
         ...
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
        |    NICKFLAG RECORD K           (4 bytes)                     |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
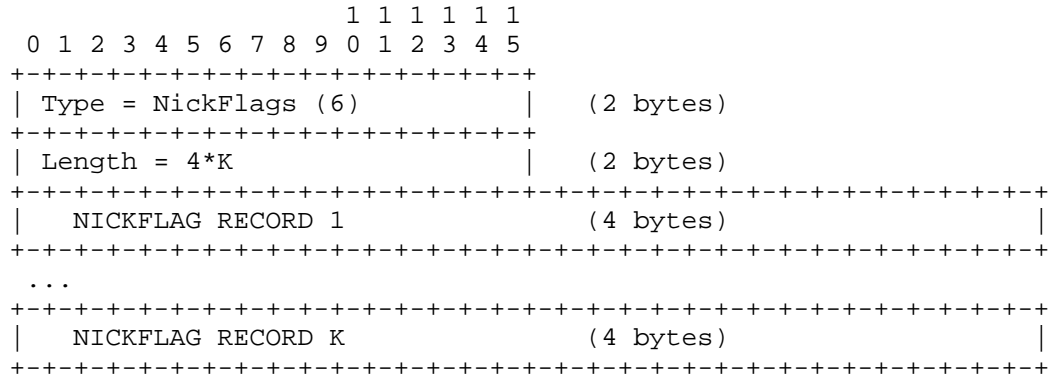
   where each NICKFLAG RECORD has the following format:

```
         0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
        +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
        |   Nickname                                    |
        +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
        |IN|     RESV                                   |
        +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```
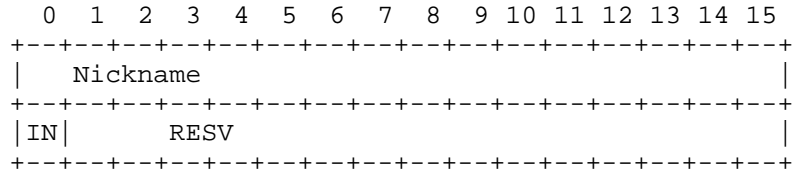
   o Type: NickFlags TRILL APPsub-TLV, set to 6 (NICKFLAGS).

   o Length: 4 times the number of NICKFLAG RECORDS present.

   o Nickname: A 16-bit TRILL nickname held by the advertising TRILL
      switch ([RFC6325] and Section 4).

   o IN: Ingress. If this flag is one, it indicates that the
      advertising TRILL switch may use the nickname in the NICKFLAG
      RECORD as the Ingress Nickname of TRILL Headers it creates. If
      the flag is zero, that nickname will not be used for that
      purpose.

   o RESV: Reserved for additional flags to be specified in the
      future. MUST be sent as zero and ignored on receipt.

The entire NickFlags APPsub-TLV is ignored if the Length is not a
multiple of 4. A NICKFLAG RECORD is ignored if the nickname it lists
is not a nickname owned by the TRILL switch advertising the enclosing
NickFlags APPsub-TLV.

If a TRILL switch intends to use a nickname in the Ingress Nickname
field of TRILL Headers it constructs, it can advertise this through
E-L1FS FS-LSPs (see Section 8.1) using a NickFlags APPsub-TLV entry
with the IN flag set. If it owns only one nickname, there is no
reason to do this because, if a TRILL switch advertises no NickFlags
APPsub-TLVs with the IN flag set for nicknames it owns, it is assumed
that the TRILL switch might use any or all nicknames it owns as the
Ingress Nickname in TRILL Headers it constructs. If a TRILL switch
advertises any NickFlags APPsub-TLV entries with the IN flag set,
then it MUST NOT use any other nickname(s) it owns as the Ingress
Nickname in TRILL Headers it constructs.

Every reasonable effort should be made to be sure that Nickname
sub-TLVs [RFC7176] and NickFlags APPsub-TLVs remain in sync. If all
TRILL switches in a campus support E-L1FS, so that Nickname sub-TLVs
can be advertised in E-L1FS FS-LSPs, then the Nickname sub-TLV and
any NickFlags APPsub-TLVs for any particular nickname SHOULD be
advertised in the same fragment. If they are not in the same
fragment, then, to the extent practical, all fragments involving
those sub-TLVs for the same nickname should be propagated as an
atomic action. If a TRILL switch sees multiple NickFlags APPsub-TLV
entries for the same nickname, it assumes that that nickname might be
used as the ingress in a TRILL Header if any of the NickFlags
APPsub-TLV entries have the IN bit set.

   It is possible that a NickFlags APPsub-TLV would not be propagated
   throughout the TRILL campus due to legacy TRILL switches not
   supporting E-L1FS. In that case, Nickname sub-TLVs MUST be
   advertised in LSPs, and TRILL switches not receiving NickFlags
   APPsub-TLVs having entries with the IN flag set will simply assume
   that the source TRILL switch might use any of its nicknames as the
   ingress in constructing TRILL Headers. Thus, the use of this
   optional APPsub-TLV is backward compatible with legacy lack of E-L1FS
   support.

   (Additional flags are assigned from those labeled RESV above and
   specified in [TRILL-L3-GW] and [Centralized-Replication].)

8.5. Graceful Restart (Unchanged)

   TRILL switches SHOULD support the features specified in [RFC5306],
   which describes a mechanism for a restarting IS-IS router to signal
   to its neighbors that it is restarting, allowing them to reestablish
   their adjacencies without cycling through the down state, while still
   correctly initiating link-state database synchronization. If this
   feature is not supported, it may increase the number of topology
   transients caused by a TRILL switch rebooting due to errors or
   maintenance.

8.6. Purge Originator Identification (New)

   To ease debugging of any purge-related problems, TRILL switches
   SHOULD include the Purge Originator Identification TLV [RFC6232] in
   all purge PDUs in TRILL IS-IS. This includes Flooding Scope LSPs
   [RFC7356] and ESADI LSPs [RFC7357].

9. Updates to RFC 7177 (Adjacency) (Changed)

   To support the E-L1FS flooding scope [RFC7356] mandated by
   Section 8.1 and backward compatibility with legacy RBridges not
   supporting E-L1FS flooding, this document updates [RFC7177] as
   follows:

   1. The list in the second paragraph of Section 3.1 of [RFC7177] is
      updated by adding the following item:

      o The Scope Flooding Support TLV.

      In addition, the sentence immediately after that list is updated
      by this document to read as follows:

         Of course, (a) the priority, (b) the Desired Designated VLAN,
         (c) the Scope Flooding Support TLV, and whether or not the
         (d) PORT-TRILL-VER sub-TLV and/or (e) BFD-Enabled TLV are
         included, and their value if included, could change on
         occasion. However, if these change, the new value(s) must
         similarly be used in all TRILL Hellos on the LAN port,
         regardless of VLAN.

   2. This document adds another bullet item to the end of Section 3.2
      of [RFC7177], as follows:

      o The value from the Scope Flooding Support TLV, or a null string
         if none was included.

   3. Near the bottom of Section 3.3 of [RFC7177], this document adds
      the following bullet item:

      o The variable-length value part of the Scope Flooding Support
         TLV in the Hello, or a null string if that TLV does not occur
         in the Hello.

   4. At the beginning of Section 4 of [RFC7177], this document adds a
      bullet item to the list, as follows:

      o The variable-length value part of the Scope Flooding Support
         TLV used in TRILL Hellos sent on the port.

   5. This document adds a line to Table 4 ("TRILL Hello Contents") in
      Section 8.1 of [RFC7177], as follows:

         LAN P2P   Number   Content Item
         --- ---   ------   --------------------------

          M   M   1   Scope Flooding Support TLV

10. TRILL Header Update (New)

   The TRILL Header has been updated from its original specification in
   [RFC6325] by [RFC7455] and [RFC7179] and is further updated by this
   document. The TRILL Header is now as shown in the figure below
   (which is followed by references for all of the fields). Those
   fields for which the reference is only to [RFC6325] are unchanged
   from that RFC.

```
                                 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                 | V |A|C|M| RESV  |F| Hop Count |
         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
         |    Egress Nickname          |     Ingress Nickname           |
         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
         :    Optional Flags Word                                       :
         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
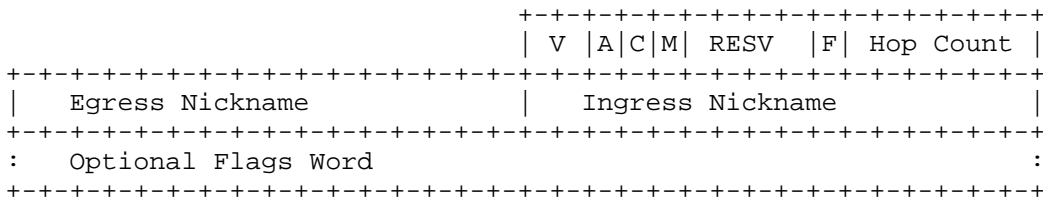
   In calculating a TRILL Data packet hash as part of equal-cost
   multipath selection, a TRILL switch MUST ignore the value of the
   "A" and "C" bits.

   In [RFC6325] and [RFC7179], there is a TRILL Header Extension Length
   field called "Op-Length", which is hereby changed to consist of the
   RESV field and "F" bit shown above.

   o V (Version): 2-bit unsigned integer. See Section 3.2
      of [RFC6325].

   o A (Alert): 1 bit. See [RFC7455].

   o C (Color): 1 bit. See Section 10.1.

   o M (Multi-destination): 1 bit. See Section 3.4 of [RFC6325].

   o RESV: 4 bits. These bits are reserved and MUST be sent as zero.
      Due to the previous use of these bits as specified in [RFC6325],
      most TRILL "fast path" hardware implementations trap and do not
      forward TRILL Data packets with these bits non-zero. A TRILL

   switch receiving a TRILL Data packet with any of these bits
   non-zero MUST discard the packet unless the non-zero bit or bits
   have some future use specified that the TRILL switch understands.

   o F: 1 bit. If this field is non-zero, then the optional flags word
      described in Section 10.2 is present. If it is zero, the
      flags word is not present.

   o Hop Count: 6 bits. See Section 3.6 of [RFC6325] and
      Section 10.2.1 below.

   o Egress Nickname: See Section 3.7.1 of [RFC6325].

   o Ingress Nickname: See Section 3.7.2 of [RFC6325].

   o Optional Flags Word: See [RFC7179] and Section 10.2.

10.1. Color Bit

   The Color bit provides an optional way by which ingress TRILL
   switches MAY mark TRILL Data packets for implementation-specific
   purposes. Transit TRILL switches MUST NOT change this bit. Transit
   and egress TRILL switches MAY use the Color bit for implementation-
   dependent traffic labeling, or for statistical analysis or other
   types of traffic study or analysis.

10.2. Flags Word Changes (Update to RFC 7179)

   When the "F" bit in the TRILL Header is non-zero, the first 32 bits
   after the Ingress Nickname field provide additional flags. These
   bits are as specified in [RFC7179], except as changed by the
   subsections below, in which the Extended Hop Count and Extended Color
   fields are described. See Section 10.3 for a diagram and summary of
   these fields.

10.2.1. Extended Hop Count

   The TRILL base protocol [RFC6325] specifies the Hop Count field in
   the header, to avoid packets persisting in the network due to looping
   or the like. However, the Hop Count field size (6 bits) limits the
   maximum hops a TRILL Data packet can traverse to 64. Optionally,
   TRILL switches can use a field composed of bits 14 through 16 in the
   flags word, as specified below, to extend this field to 9 bits. This
   increases the maximum Hop Count to 512. Except in rare
   circumstances, reliable use of Hop Counts in excess of 64 requires
   support of this optional capability at all TRILL switches along the
   path of a TRILL Data packet.

10.2.1.1. Advertising Support

   It may be that not all the TRILL switches support the Extended Hop
   Count mechanism in a TRILL campus and in that campus more than
   64 hops are required either for the distribution tree calculated path
   or for the unicast calculated path plus a reasonable allowance for
   alternate pathing. As such, it is required that TRILL switches
   advertise their support by setting bit 14 in the TRILL Version
   Sub-TLV Capabilities and Header Flags Supported field [RFC7176];
   bits 15 and 16 of that field are now specified as Unassigned (see
   Section 12.2.5).

10.2.1.2. Ingress Behavior

   If an ingress TRILL switch determines that it should set the
   Hop Count for a TRILL Data packet to 63 or less, then behavior is as
   specified in the TRILL base protocol [RFC6325]. If the optional
   TRILL Header flags word is present, bits 14, 15, and 16 and the
   critical reserved bit of the critical summary bits are zero.

   If the Hop Count for a TRILL Data packet should be set to some value
   greater than 63 but less than 512 and all TRILL switches that the
   packet is reasonably likely to encounter support Extended Hop Count,
   then the resulting TRILL Header has the flags word extension present,
   the high-order 3 bits of the desired Hop Count are stored in the
   Extended Hop Count field in the flags word, the low-order 5 bits are
   stored in the Hop Count field in the first word of the TRILL Header,
   and bit two (the critical reserved bit of the critical summary bits)
   in the flags word is set to one.

   For known unicast traffic (TRILL Header "M" bit zero), an ingress
   TRILL switch discards the frame if it determines that the least-cost
   path to the egress is (1) more than 64 hops and not all TRILL
   switches on that path support the Extended Hop Count feature or
   (2) more than 512 hops.

   For multi-destination traffic, when a TRILL switch determines that
   one or more tree paths from the ingress are more than 64 hops and not
   all TRILL switches in the campus support the Extended Hop Count
   feature, the encapsulation uses a total Hop Count of 63 to obtain at
   least partial distribution of the traffic.

10.2.1.3. Transit Behavior

   A transit TRILL switch supporting Extended Hop Count behaves like a
   base protocol [RFC6325] TRILL switch in decrementing the Hop Count,
   except that it considers the Hop Count to be a 9-bit field where the
   Extended Hop Count field constitutes the high-order 3 bits.

   To be more precise: a TRILL switch supporting Extended Hop Count
   takes the first of the following actions that is applicable:

   1. If both the Hop Count and Extended Hop Count fields are zero, the
      packet is discarded.

   2. If the Hop Count is non-zero, it is decremented. As long as the
      Extended Hop Count is non-zero, no special action is taken. If
      the result of this decrement is zero, the packet is processed
      normally.

   3. If the Hop Count is zero, it is set to the maximum value of 63,
      and the Extended Hop Count is decremented. If this results in the
      Extended Hop Count being zero, the critical reserved bit in the
      critical summary bits is set to zero.

10.2.1.4. Egress Behavior

   No special behavior is required when egressing a TRILL Data packet
   that uses the Extended Hop Count. The flags word, if present, is
   removed along with the rest of the TRILL Header during decapsulation.

10.2.2. Extended Color Field

   Flags word bits 27 and 28 are specified to be a 2-bit Extended Color
   field (see Section 10.3). These bits are in the non-critical
   ingress-to-egress region of the flags word.

   The Extended Color field provides an optional way by which ingress
   TRILL switches MAY mark TRILL Data packets for implementation-
   specific purposes. Transit TRILL switches MUST NOT change these
   bits. Transit and egress TRILL switches MAY use the Extended Color
   bits for implementation-dependent traffic labeling, or for
   statistical analysis or other types of traffic study or analysis.

   Per Section 2.3.1 of [RFC7176], support for these bits is indicated
   by the same bits (27 and 28) in the Capabilities and Header Flags
   Supported field of the TRILL version sub-TLV. If these bits are zero
   in those capabilities, Extended Color is not supported. A TRILL
   switch that does not support Extended Color will ignore the
   corresponding bits in any TRILL Header flags word it receives as part
   of a TRILL Data packet and will set those bits to zero in any TRILL
   Header flags word it creates. A TRILL switch that sets or senses the
   Extended Color field on transmitting or receiving TRILL Data packets
   MUST set the corresponding 2-bit field in the TRILL version sub-TLV
   to a non-zero value. Any difference in the meaning of the three
   possible non-zero values of this 2-bit capability field (0b01, 0b10,
   or 0b11) is implementation dependent.

10.3. Updated Flags Word Summary

   With the changes above, the 32-bit flags word extension to the TRILL
   Header [RFC7179], which is detailed in the "TRILL Extended Header
   Flags" registry on the "Transparent Interconnection of Lots of Links
   (TRILL) Parameters" IANA web page, is now as follows:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Crit.|   CHbH   |    NCHbH   |CRSV | NCRSV |    CItE   |  NCItE  |
|.....|.........|...........|.....|.......|...........|.........|
|C|C|C|         |C|N|        | Ext |       |           |Ext|     |
|R|R|R|         |R|C|        | Hop |       |           |Clr|     |
|H|I|R|         |C|C|        | Cnt |       |           |   |     |
|b|t|s|         |A|A|        |     |       |           |   |     |
|H|E|v|         |F|F|        |     |       |           |   |     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
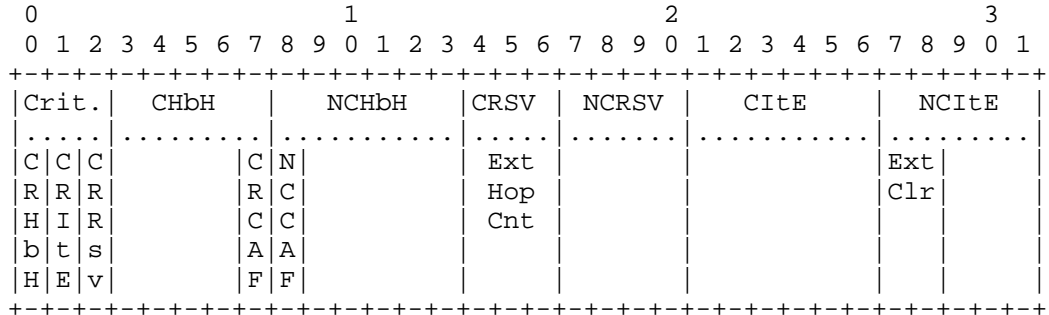
   Bits 0, 1, and 2 are the critical summary bits, as specified in
   [RFC7179], consisting of the critical hop-by-hop, critical
   ingress-to-egress, and critical reserved bits, respectively. The
   next two fields are specific critical and non-critical hop-by-hop
   bits -- CHbH and NCHbH, respectively -- containing the Critical and
   Non-critical Channel Alert flags as specified in [RFC7179]. The next
   field is the critical reserved bits (CRSV), which are specified
   herein to be the Extended Hop Count. The non-critical reserved bits
   (NCRSV) and the critical ingress-to-egress bits (CItE) as specified
   in [RFC7179] follow. Finally, there is the non-critical
   ingress-to-egress field, including bits 27 and 28, which are
   specified herein as the Extended Color field.

11. Appointed Forwarder Status Lost Counter (New)

   Strict conformance to the provisions of Section 4.8.3 of [RFC6325] on
   the value of the Appointed Forwarder Status Lost Counter can result
   in the splitting of Interested VLANs and Spanning Tree Roots sub-TLVs
   [RFC7176] (or the corresponding Interested Labels and Spanning Tree
   Roots sub-TLVs where a VLAN is mapped to an FGL) due to differences
   in this counter value for adjacent VLAN IDs (or 24-bit FGLs). This
   counter is a mechanism to optimize data-plane learning by trimming
   the expiration timer for learned addresses on a per-VLAN/FGL basis
   under some circumstances.

   The requirement to increment this counter by one whenever a TRILL
   switch loses Appointed Forwarder status on a port is hereby changed
   from the mandatory provisions of [RFC6325] to the enumerated
   provisions below. To the extent that this might cause the Appointed

Forwarder Status Lost Counter to be increased when [RFC6325]
indicates that it should not, this will cause data-plane address
learning timeouts at remote TRILL switches to be reduced. To the
extent that this might cause the Appointed Forwarder Status Lost
Counter to remain unchanged when [RFC6325] indicates that it should
be increased, this will defeat a reduction in such timeouts that
would otherwise occur.

(1) If any of the following apply, either data-plane address learning
    is not in use or Appointed Forwarder status is irrelevant. In
    these cases, the Appointed Forwarder Status Lost Counter MAY be
    left at zero or set to any convenient value such as the value of
    the Appointed Forwarder Status Lost Counter for an adjacent
    VLAN ID or FGL.

    (1a) The TRILL switch port has been configured with the
         "end-station service disable" bit (also known as the
         trunk bit) on.

    (1b) The TRILL switch port has been configured in IS-IS as an
         IS-IS point-to-point link.

    (1c) The TRILL switch is relying on ESADI [RFC7357] or Directory
         Assist [RFC7067] and not using data-plane learning.

(2) In cases other than those enumerated in point 1 above, the
    Appointed Forwarder Status Lost Counter SHOULD be incremented as
    described in [RFC6325]. Such incrementing has the advantage of
    optimizing data-plane learning. Alternatively, the value of the
    Appointed Forwarder Status Lost Counter can deviate from that
    value -- for example, to make it match the value for an adjacent
    VLAN ID (or FGL), so as to permit greater aggregation of
    Interested VLANs and Spanning Tree Roots sub-TLVs.

12. IANA Considerations (Changed)

   This section lists IANA actions previously completed and new IANA
   actions.

12.1. Previously Completed IANA Actions (Unchanged)

   The following IANA actions were completed as part of [RFC7180] and
   are included here for completeness, since this document obsoletes
   [RFC7180].

   1. The nickname 0xFFC1, which was reserved by [RFC6325], is allocated
      for use in the TRILL Header Egress Nickname field to indicate an
      OOMF (Overload Originated Multi-destination Frame).

   2. Bit 1 from the seven previously reserved (RESV) bits in the
      per-neighbor "Neighbor RECORD" in the TRILL Neighbor TLV [RFC7176]
      is allocated to indicate that the RBridge sending the TRILL Hello
      volunteers to provide the OOMF forwarding service described in
      Section 2.4.2 to such frames originated by the TRILL switch whose
      SNPA (MAC address) appears in that Neighbor RECORD. The
      description of this bit is "Offering OOMF service".

   3. Bit 0 is allocated from the capability bits in the PORT-TRILL-VER
      sub-TLV [RFC7176] to indicate support of the VLANs Appointed
      sub-TLV [RFC7176] and the VLAN inhibition setting mechanisms
      specified in [RFC6439bis]. The description of this bit is "Hello
      reduction support".

12.2. New IANA Actions (New)

   The following are new IANA actions for this document.

12.2.1. Reference Updated

   All references to [RFC7180] in the "Transparent Interconnection of
   Lots of Links (TRILL) Parameters" registry have been replaced with
   references to this document, except that the Reference for bit 0 in
   the PORT-TRILL-VER Sub-TLV Capability Flags has been changed to
   [RFC6439bis].

12.2.2. The "E" Capability Bit

   There is an existing TRILL version sub-TLV, sub-TLV #13, under both
   TLV #242 and TLV #144 [RFC7176]. This TRILL version sub-TLV contains
   a capability bits field for which assignments are documented in the
   "TRILL-VER Sub-TLV Capability Flags" registry on the TRILL Parameters
   IANA web page. IANA has allocated 4 from the previously reserved

   bits in this "TRILL-VER Sub-TLV Capability Flags" registry to
   indicate support of the E-L1FS flooding scope as specified in
   Section 8.1. This capability bit is referred to as the "E" bit. The
   following is the addition to the "TRILL-VER Sub-TLV Capability Flags"
   registry:

      Bit   Description   References
      ----   --------------------- ---------------
      4   E-L1FS FS-LSP support [RFC7356], RFC 7780

12.2.3. NickFlags APPsub-TLV Number and Registry

   IANA has assigned an APPsub-TLV number, as follows, under the TRILL
   GENINFO TLV from the range less than 255.

      Type   Name   References
      ----   ---------   -----------
      6   NICKFLAGS   RFC 7780

   In addition, IANA has created a registry on its TRILL Parameters web
   page for NickFlags bit assignments, as follows:

      Name: NickFlags Bits
      Registration Procedure: IETF Review [RFC5226]
      Reference: RFC 7780

       Bit   Mnemonic Description   Reference
      -----   -------- -----------   ---------
       0   IN   Used as ingress RFC 7780
      1-15   -   Unassigned   RFC 7780

12.2.4. Updated TRILL Extended Header Flags

   The "TRILL Extended Header Flags" registry has been updated as
   follows:

   Bits   Purpose   Reference
   -----   --------------------------------------- ------------

   14-16   Extended Hop Count   RFC 7780

   27-28   Extended Color   RFC 7780

   29-31   Available non-critical ingress-to-egress [RFC7179], RFC 7780
            flags

12.2.5. TRILL-VER Sub-TLV Capability Flags

   The "TRILL-VER Sub-TLV Capability Flags" registry has been updated as
   follows:

   Bit   Description   Reference
   -----   ------------------------- ----------------

      14   Extended Hop Count support   RFC 7780

   15-16   Unassigned   RFC 7780

   27-28   Extended Color support   RFC 7780

   29-31   Extended header flag support [RFC7179], RFC 7780

12.2.6. Example Nicknames

   As shown in the table below, IANA has assigned a block of eight
   nicknames for use as examples in documentation. Appendix B shows a
   use of some of these nicknames. The "TRILL Nicknames" registry has
   been updated by changing the previous "0xFFC2-0xFFFE Unassigned" line
   to the following:

      Name   Description   Reference
   ------------- --------------   -----------
   0xFFC2-0xFFD7 Unassigned
   0xFFD8-0xFFDF For use in documentation examples RFC 7780
   0xFFE0-0xFFFE Unassigned

13. Security Considerations (Changed)

   See [RFC6325] for general TRILL security considerations.

   This memo improves the documentation of the TRILL protocol; corrects
   six errata in [RFC6325]; updates [RFC6325], [RFC7177], and [RFC7179];
   and obsoletes [RFC7180]. It does not change the security
   considerations of those RFCs, except as follows:

   o E-L1FS FS-LSPs can be authenticated with IS-IS security [RFC5310],
      that is, through the inclusion of an IS-IS Authentication TLV in
      E-L1FS PDUs.

   o As discussed in Section 3.6, when using an allowed weaker RPF
      check under very rare topologies and transient conditions,
      multi-destination TRILL Data packets can be duplicated; this could
      have security consequences for some protocols.

14. References

14.1. Normative References

   [802.1Q-2014]
             IEEE, "IEEE Standard for Local and metropolitan area
             networks -- Bridges and Bridged Networks",
             DOI 10.1109/IEEESTD.2014.6991462, IEEE Std 802.1Q-2014.

   [IS-IS]   International Organization for Standardization,
             "Information technology -- Telecommunications and
             information exchange between systems -- Intermediate
             System to Intermediate System intra-domain routeing
             information exchange protocol for use in conjunction with
             the protocol for providing the connectionless-mode network
             service (ISO 8473)", ISO/IEC 10589:2002, Second Edition,
             November 2002.

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

   [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic
             Engineering", RFC 5305, DOI 10.17487/RFC5305,
             October 2008, <http://www.rfc-editor.org/info/rfc5305>.

   [RFC5306] Shand, M. and L. Ginsberg, "Restart Signaling for IS-IS",
             RFC 5306, DOI 10.17487/RFC5306, October 2008,
             <http://www.rfc-editor.org/info/rfc5306>.

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

   [RFC6232] Wei, F., Qin, Y., Li, Z., Li, T., and J. Dong, "Purge
             Originator Identification TLV for IS-IS", RFC 6232,
             DOI 10.17487/RFC6232, May 2011,
             <http://www.rfc-editor.org/info/rfc6232>.

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

   [RFC6361] Carlson, J. and D. Eastlake 3rd, "PPP Transparent
              Interconnection of Lots of Links (TRILL) Protocol Control
              Protocol", RFC 6361, DOI 10.17487/RFC6361, August 2011,
              <http://www.rfc-editor.org/info/rfc6361>.

   [RFC7172] Eastlake 3rd, D., Zhang, M., Agarwal, P., Perlman, R., and
              D. Dutt, "Transparent Interconnection of Lots of Links
              (TRILL): Fine-Grained Labeling", RFC 7172,
              DOI 10.17487/RFC7172, May 2014,
              <http://www.rfc-editor.org/info/rfc7172>.

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

   [RFC7177] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Yang, H., and
              V. Manral, "Transparent Interconnection of Lots of Links
              (TRILL): Adjacency", RFC 7177, DOI 10.17487/RFC7177,
              May 2014, <http://www.rfc-editor.org/info/rfc7177>.

   [RFC7179] Eastlake 3rd, D., Ghanwani, A., Manral, V., Li, Y., and C.
              Bestler, "Transparent Interconnection of Lots of Links
              (TRILL): Header Extension", RFC 7179,
              DOI 10.17487/RFC7179, May 2014,
              <http://www.rfc-editor.org/info/rfc7179>.

   [RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding
              Scope Link State PDUs (LSPs)", RFC 7356,
              DOI 10.17487/RFC7356, September 2014,
              <http://www.rfc-editor.org/info/rfc7356>.

   [RFC7455] Senevirathne, T., Finn, N., Salam, S., Kumar, D., Eastlake
              3rd, D., Aldrin, S., and Y. Li, "Transparent
              Interconnection of Lots of Links (TRILL): Fault
              Management", RFC 7455, DOI 10.17487/RFC7455, March 2015,
              <http://www.rfc-editor.org/info/rfc7455>.

14.2. Informative References

   [802]   IEEE 802, "IEEE Standard for Local and Metropolitan Area
              Networks: Overview and Architecture",
              DOI 10.1109/IEEESTD.2014.6847097, IEEE Std 802-2014.

   [Centralized-Replication]
              Hao, W., Li, Y., Durrani, M., Gupta, S., Qu, A., and T.
              Han, "Centralized Replication for BUM traffic in
              active-active edge connection", Work in Progress,
              draft-ietf-trill-centralized-replication-03,
              November 2015.

   [Err3002] RFC Errata, Erratum ID 3002, RFC 6325.

   [Err3003] RFC Errata, Erratum ID 3003, RFC 6325.

   [Err3004] RFC Errata, Erratum ID 3004, RFC 6325.

   [Err3052] RFC Errata, Erratum ID 3052, RFC 6325.

   [Err3053] RFC Errata, Erratum ID 3053, RFC 6325.

   [Err3508] RFC Errata, Erratum ID 3508, RFC 6325.

   [RFC792]   Postel, J., "Internet Control Message Protocol", STD 5,
              RFC 792, DOI 10.17487/RFC0792, September 1981,
              <http://www.rfc-editor.org/info/rfc792>.

   [RFC826]   Plummer, D., "Ethernet Address Resolution Protocol: Or
              Converting Network Protocol Addresses to 48.bit Ethernet
              Address for Transmission on Ethernet Hardware", STD 37,
              RFC 826, DOI 10.17487/RFC0826, November 1982,
              <http://www.rfc-editor.org/info/rfc826>.

   [RFC4086] Eastlake 3rd, D., Schiller, J., and S. Crocker,
              "Randomness Requirements for Security", BCP 106, RFC 4086,
              DOI 10.17487/RFC4086, June 2005,
              <http://www.rfc-editor.org/info/rfc4086>.

   [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an
              IANA Considerations Section in RFCs", BCP 26, RFC 5226,
              DOI 10.17487/RFC5226, May 2008,
              <http://www.rfc-editor.org/info/rfc5226>.

   [RFC6327] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Dutt, D., and
              V. Manral, "Routing Bridges (RBridges): Adjacency",
              RFC 6327, DOI 10.17487/RFC6327, July 2011,
              <http://www.rfc-editor.org/info/rfc6327>.

   [RFC6439] Perlman, R., Eastlake, D., Li, Y., Banerjee, A., and F.
              Hu, "Routing Bridges (RBridges): Appointed Forwarders",
              RFC 6439, DOI 10.17487/RFC6439, November 2011,
              <http://www.rfc-editor.org/info/rfc6439>.

   [RFC6439bis]
              Eastlake 3rd, D., Li, Y., Umair, M., Banerjee, A., and H.
              Fangwei, "TRILL: Appointed Forwarders", Work in Progress,
              draft-ietf-trill-rfc6439bis-01, January 2016.

   [RFC7042] Eastlake 3rd, D. and J. Abley, "IANA Considerations and
              IETF Protocol and Documentation Usage for IEEE 802
              Parameters", BCP 141, RFC 7042, DOI 10.17487/RFC7042,
              October 2013, <http://www.rfc-editor.org/info/rfc7042>.

   [RFC7067] Dunbar, L., Eastlake 3rd, D., Perlman, R., and I.
              Gashinsky, "Directory Assistance Problem and High-Level
              Design Proposal", RFC 7067, DOI 10.17487/RFC7067,
              November 2013, <http://www.rfc-editor.org/info/rfc7067>.

   [RFC7175] Manral, V., Eastlake 3rd, D., Ward, D., and A. Banerjee,
              "Transparent Interconnection of Lots of Links (TRILL):
              Bidirectional Forwarding Detection (BFD) Support",
              RFC 7175, DOI 10.17487/RFC7175, May 2014,
              <http://www.rfc-editor.org/info/rfc7175>.

   [RFC7178] Eastlake 3rd, D., Manral, V., Li, Y., Aldrin, S., and D.
              Ward, "Transparent Interconnection of Lots of Links
              (TRILL): RBridge Channel Support", RFC 7178,
              DOI 10.17487/RFC7178, May 2014,
              <http://www.rfc-editor.org/info/rfc7178>.

   [RFC7180] Eastlake 3rd, D., Zhang, M., Ghanwani, A., Manral, V., and
              A. Banerjee, "Transparent Interconnection of Lots of Links
              (TRILL): Clarifications, Corrections, and Updates",
              RFC 7180, DOI 10.17487/RFC7180, May 2014,
              <http://www.rfc-editor.org/info/rfc7180>.

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

   [TRILL-L3-GW]
              Hao, W., Li, Y., Qu, A., Durrani, M., Sivamurugan, P., and
              L. Xia, "TRILL Distributed Layer 3 Gateway", Work in
              Progress, draft-ietf-trill-irb-10, January 2016.

Appendix A. Life Cycle of a TRILL Switch Port (New)

   Text from <http://www.ietf.org/mail-archive/web/trill/
   current/msg06355.html> is paraphrased in this informational appendix.

   Question:
      Suppose we are developing a TRILL implementation to run on
      different machines. Then what happens first? Is LSP flooding or
      ESADI started first? -> Link-state database creation ->
      Designated RBridge election (How to set priority? Any fixed
      process that depends on user settings?) -> etc.

   Answer:
      The first thing that happens on a port/link is any link setup that
      is needed. For example, on a PPP link [RFC6361], you need to
      negotiate that you will be using TRILL. However, if you have
      Ethernet links [RFC6325], which are probably the most common type,
      there isn't any link setup needed.

      As soon as the port is set up, it can ingress or egress native
      frames if end-station service is being offered on that port.
      Offering end-station service is the default. However, if the port
      trunk bit (end-station service disable) is set or the port is
      configured as an IS-IS point-to-point link port, then end-station
      service is not offered; therefore, native frames received are
      ignored, and native frames are not egressed.

      TRILL IS-IS Hellos then get sent out the port to be exchanged with
      any other TRILL switches on the link [RFC7177]. Only the Hellos
      are required; optionally, you might also exchange MTU-probe/ack
      PDUs [RFC7177], BFD PDUs [RFC7175], or other link test packets.

      TRILL doesn't send any TRILL Data or TRILL IS-IS packets out the
      port to the link, except for Hellos, until the link gets to the
      2-Way or Report state [RFC7177].

      If a link is configured as a point-to-point link, there is no
      Designated RBridge (DRB) election. By default, an Ethernet link
      is considered a LAN link, and the DRB election occurs when the
      link is in any state other than Down. You don't have to configure
      priorities for each TRILL switch (RBridge) to be the DRB. Things
      will work fine with all the RBridges on a link using default
      priority. But if the network manager wants to control this, there
      should be a way for them to configure the priority to be the DRB
      of the TRILL switch ports on the link.

(To avoid complexity, this appendix generally describes the
life cycle for a link that only has two TRILL switches on it. But
TRILL works fine as currently specified on a broadcast link with
multiple TRILL switches on it -- actually, multiple TRILL switch
ports -- since a TRILL switch can have multiple ports connected to
the same link. The most likely way to get such a multi-access
link with current technology and the existing TRILL standards is
to have more than two TRILL switch Ethernet ports connected to a
bridged LAN. The TRILL protocol operates above all bridging; in
general, the bridged LAN looks like a transparent broadcast link
to TRILL.)

When a link gets to the 2-Way or Report state, LSPs, CSNPs, and
PSNPs will start to flow on the link (as well as FS-LSPs,
FS-CSNPs, and FS-PSNPs for E-L1FS (see Section 8.1)).

When a link gets to the Report state, there is adjacency. The
existence of that adjacency is flooded (reported) to the campus in
LSPs. TRILL Data packets can then start to flow on the link as
TRILL switches recalculate the least-cost paths and distribution
trees to take the new adjacency into account. Until it gets to
the Report state, there is no adjacency, and no TRILL Data packets
can flow over that link (with the minor corner case exception that
an RBridge Channel message can, for its first hop only, be sent on
a port where there is no adjacency (Section 2.4 of [RFC7178]).
(Although this paragraph seems to be talking about link state, it
is actually port state. It is possible for different TRILL switch
ports on the same link to temporarily be in different states. The
adjacency state machinery runs independently on each port.)

ESADI [RFC7357] is built on top of the regular TRILL Data routing.
Since ESADI PDUs look, to transit TRILL switches, like regular
TRILL Data packets, no ESADI PDUs can flow until adjacencies are
established and TRILL Data is flowing. Of course, ESADI is
optional and is not used unless configured.

Question:
   Does it require TRILL Full Headers at the time TRILL LSPs start
   being broadcast on a link? Because at that time it's not defined
   egress and ingress nicknames.

Answer:
   TRILL Headers are only for TRILL Data packets. TRILL IS-IS
   packets, such as TRILL LSPs, are sent in a different way that does
   not use a TRILL Header and does not depend on nicknames.

   Probably, in most implementations, a TRILL switch will start up
   using the same nickname it had when it shut down or last got
   disconnected from a campus. If you want, you can implement TRILL
   to come up initially not reporting any nickname (by not including
   a Nickname sub-TLV in its LSPs) until you get the link-state
   database or most of the link-state database, and then choose a
   nickname no other TRILL switch in the campus is using. Of course,
   if a TRILL switch does not have a nickname, then it cannot ingress
   data, cannot egress known unicast data, and cannot be a tree root.

   TRILL IS-IS PDUs such as LSPs, and the link-state database, all
   work based on the 7-byte IS-IS System ID (sometimes called the
   LAN ID [IS-IS]). Since topology determination uses System IDs,
   which are always unique across the campus, it is not affected by
   the nickname assignment state. The nickname system is built on
   top of that.

Appendix B. Example TRILL PDUs (New)

   This appendix shows example TRILL IS-IS PDUs. The primary purpose of
   these examples is to clarify issues related to bit ordering.

   The examples in this appendix concentrate on the format of the packet
   header and trailer. There are frequently unspecified optional items
   or data in the packet that would affect header or trailer fields like
   the packet length or checksum. Thus, an "Xed out" placeholder is
   used for such fields, where each X represents one hex nibble.

B.1. LAN Hello over Ethernet

   A TRILL Hello sent from a TRILL switch (RBridge) with 7-byte
   System ID 0x30033003300300 holding nickname 0xFFDE over Ethernet from
   a port with MAC address 0x00005E0053DE on VLAN 1 at priority 7.
   There is one neighbor that is the DRB. The neighbor's port MAC is
   0x00005E0053E3, and the neighbor's System ID is 0x44444444444400.

      Ethernet Header
        Outer.MacDA, Outer.MacSA
          0x0180C2000041   All-IS-IS-RBridges Destination MAC Address
          0x00005E0053DE   Source MAC Address
        Outer VLAN Tag (optional)
          0x8100   C-VLAN Ethertype [802.1Q-2014]
          0xE001   Priority 7, Outer.VLAN
        IS-IS
          0x22F4   L2-IS-IS Ethertype
      IS-IS Payload
        Common Header
          0x83   Intradomain Routeing Protocol Discriminator
          0x08   Header Length
          0x01   IS-IS Version Number
          0x06   ID Length of 6 Bytes
          0x0F   PDU Type (Level 1 LAN Hello)
          0x01   Version
          0x00   Reserved
          0x01   Maximum Area Addresses
        Hello PDU Specific Fields
          0x01   Circuit Type (Level 1)
          0x30033003300300 Source System ID
          0x0009   Holding Time
          0xXXXX   PDU Length
          0x40   Priority to be DRB
          0x44444444444400 LAN ID
        TLVs (the following order of TLVs or of sub-TLVs in a TLV
          is not significant)

```
      Area Addresses TLV
         0x01               Area Addresses Type
         0x02               Length of Value
         0x01               Length of Address
         0x00               The fixed TRILL Area Address
      MT Port Capabilities TLV
         0x8F               MT Port Capabilities Type
         0x0011             Length of Value
         0x0000             Topology
           Special VLANs and Flags Sub-TLV
              0x01               Sub-TLV Type
              0x08               Length
              0x0123             Port ID
              0xFFDE             Sender Nickname
              0x0001             Outer.VLAN
              0x0001             Designated VLAN
           Enabled VLANs Sub-TLV (optional)
              0x02               Sub-TLV Type
              0x03               Length
              0x0001             Start VLAN 1
              0x80               VLAN 1
      TRILL Neighbor TLV
         0x91               Neighbor Type
         0x0A               Length of Value
         0xC0               S Flag = 1, L Flag = 1, SIZE field 0
           NEIGHBOR RECORD
           0x00               Flags
           0x2328             MTU = 9 KB
           0x00005E0053E3  Neighbor MAC Address
      Scope Flooding Support TLV
      0xF3               Scope Flooding Support Type
      0x01               Length of Value
      0x40               E-L1FS Flooding Scope
      More TLVs (optional)
         ...
   Ethernet Trailer
      0xXXXXXXXX       Ethernet Frame Check Sequence (FCS)
```

B.2. LSP over PPP

   Here is an example of a TRILL LSP sent over a PPP link by the same
   source TRILL switch as the example in Appendix B.1.

```
      PPP Header
        0x405D                 PPP TRILL Link State Protocol
      IS-IS Payload
        Common Header
          0x83                 Intradomain Routeing Protocol Discriminator
          0x08                 Header Length
          0x01                 IS-IS Version Number
          0x06                 ID Length of 6 Bytes
          0x12                 PDU Type (Level 1 LSP)
          0x01                 Version
          0x00                 Reserved
          0x01                 Maximum Area Addresses
        LSP Specific Fields
          0xXXXX               PDU Length
          0x0123               Remaining Lifetime
          0x3003300330030009 LSP ID (fragment 9)
          0x00001234           Sequence Number
          0xXXXX               Checksum
          0x01                 Flags = Level 1
        TLVs (the following order of TLVs or of sub-TLVs in a TLV
          is not significant)
        Router Capability TLV
          0xF2                 Router Capability Type
          0x0F                 Length of Value
          0x00                 Flags
            Nickname Sub-TLV
              0x06             Sub-TLV Type
              0x05             Length of Value
            NICKNAME RECORD
              0x33             Nickname Priority
              0x1234           Tree Root Priority
              0xFFDE           Nickname
          TRILL Version Sub-TLV
              0x0D             Sub-TLV Type
              0x05
              0x00             Max Version
              0x40000000       Flags = FGL Support
        More TLVs (optional
          ...
      PPP Trailer
        0xXXXXXX               PPP Frame Check Sequence (FCS)
```

B.3. TRILL Data over Ethernet

   Below is an IPv4 ICMP Echo [RFC792] sent in a TRILL Data packet from
   the TRILL switch that sent the Hello in Appendix B.1 to the neighbor
   TRILL switch on the link used in Appendix B.1.

```
      Ethernet Header
        Outer.MacDA, Outer.MacSA
          0x00005E0053E3  Destination MAC Address
          0x00005E0053DE  Source MAC Address
        Outer VLAN Tag (optional)
          0x8100          C-VLAN Ethertype [802.1Q-2014]
          0x0001          Priority 0, Outer.VLAN 1
         TRILL
          0x22F3          TRILL Ethertype
      TRILL Header
          0X000E          Flags, Hop Count 14
          0XFFDF          Egress Nickname
          0xFFDC          Ingress Nickname
      Inner Ethernet Header
        Inner.MacDA, Inner.MacSA
          0x00005E005322  Destination MAC Address
          0x00005E005344  Source MAC Address
        Inner VLAN Tag
          0x8100          C-VLAN Ethertype
          0x0022          Priority 0, Inner.VLAN 34
        Ethertype
          0x0800          IPv4 Ethertype
      IP Header
          0x4500          Version 4, Header Length 5, ToS 0
          0xXXXX          Total Length
          0x3579          Identification
          0x0000          Flags, Fragment Offset
          0x1101          TTL 17, ICMP = Protocol 1
          0xXXXX          Header Checksum
          0xC0000207      Source IP 192.0.2.7
          0xC000020D      Destination IP 192.0.2.13
          0x00000000      Options, Padding
      ICMP
          0x0800          ICMP Echo
          0xXXXX          Checksum
          0x87654321      Identifier, Sequence Number
          ...             Echo Data
      Ethernet Trailer
        0xXXXXXXXX        Ethernet Frame Check Sequence (FCS)
```

B.4.   TRILL Data over PPP

   Below is an ARP Request [RFC826] sent in a TRILL Data packet from the
   TRILL switch that sent the Hello in Appendix B.1 over a PPP link.

```
      PPP Header
        0x005D            PPP TRILL Network Protocol
      TRILL Header
          0X080D            Flags (M = 1), Hop Count 13
          0xFFDD            Distribution Tree Root Nickname
          0xFFDC            Ingress Nickname
      Inner Ethernet Header
        Inner.MacDA, Inner.MacSA
          0xFFFFFFFFFFFF  Destination MAC Address
          0x00005E005344  Source MAC Address
        Inner VLAN Tag
          0x8100          C-VLAN Ethertype
          0x0022          Priority 0, Inner.VLAN 34
        Ethertype
          0x0806          ARP Ethertype
      ARP
          0x0001            Hardware Address Space = Ethernet
          0x0001            Protocol Address Space = IPv4
          0x06              Size of Hardware Address
          0x04              Size of Protocol Address
          0x0001            OpCode = Request
          0x00005E005344  Sender Hardware Address
          0xC0000207        Sender Protocol Address 192.0.2.7
          0x000000000000  Target Hardware Address
          0xC000020D        Target Protocol Address 192.0.2.13
      PPP Trailer
        0xXXXXXX            PPP Frame Check Sequence (FCS)
```

Appendix C. Changes to Previous RFCs (New)

C.1. Changes to Obsoleted RFC 7180

   This section summarizes the changes, augmentations, and excisions
   this document specifies for [RFC7180], which it obsoletes and
   replaces.

C.1.1. Changes

   For each section header in this document ending with "(Changed)",
   this section summarizes the changes that are made by this document:

   Section 1 ("Introduction"): Numerous changes to reflect the overall
   changes in contents.

   Section 1.1 ("Precedence"): Changed to add mention of [RFC7179].

   Section 1.3 ("Terminology and Acronyms"): Numerous terms added.

   Section 3 ("Distribution Trees and RPF Check"): Changed by the
   addition of the new material in Section 3.6. See Appendix C.1.2,
   Item 1.

   Section 8 ("Other IS-IS Considerations"): Changed by the addition of
   Sections 8.1, 8.2, 8.3, and 8.4. See Appendix C.1.2 -- Items 2, 3,
   4, and 5, respectively.

   Section 9 ("Updates to RFC 7177 (Adjacency)": Changes and additions
   to [RFC7177] to support E-L1FS. See Appendix C.1.2, Item 2.

   Section 12 ("IANA Considerations"): Changed by the addition of
   material in Section 12.2. See Appendix C.1.2, Item 7.

   Section 13 ("Security Considerations"): Minor changes in the RFCs
   listed.

C.1.2. Additions

   This document contains the following material not present in
   [RFC7180]:

   1. Support for an alternative Reverse Path Forwarding Check (RPFC),
       along with considerations for deciding between the original
       [RFC6325] RPFC and this alternative RPFC. This alternative RPFC
       was originally discussed on the TRILL WG mailing list in
       <http://www.ietf.org/mail-archive/web/trill/current/
       msg01852.html> and subsequent messages (Section 3.6).

   2. Mandatory E-L1FS [RFC7356] support (Sections 8.1 and 9).

   3. Recommendations concerning control packet priorities
       (Section 8.2).

   4. Implementation requirements concerning unknown IS-IS PDU types
       (Section 8.3).

   5. Specification of an optional Nickname Flags APPsub-TLV and an
       ingress flag within that APPsub-TLV (Section 8.4).

   6. Update to the TRILL Header to allocate a Color bit
       (Section 10.1), and update to the optional TRILL Header Extension
       flags word to allocate a 2-bit Extended Color field
       (Section 10.2).

   7. Some new IANA Considerations in Section 12.2, including
       reservation of nicknames for use as examples in documentation.

   8. A new "Appointed Forwarder Status Lost Counter" section
       (Section 11 of this document) that loosens the mandatory update
       requirements specified in [RFC6325].

   9. Informative Appendix A on the life cycle of a TRILL port.

   10. A new Appendix B containing example TRILL PDUs.

   11. Recommendation to use the Purge Originator Identification TLV
       (Section 8.6).

C.1.3. Deletions

   This document omits the following material that was present in
   [RFC7180]:

   1. All updates to [RFC6327] that occurred in [RFC7180]. These have
       been rolled into [RFC7177], which obsoletes [RFC6327]. However,
       new updates to [RFC7177] are included (see Appendix C.3).

   2. All updates to [RFC6439]. These have been rolled into
       [RFC6439bis], which is intended to obsolete [RFC6439].

C.2. Changes to RFC 6325

   This document contains many normative updates to [RFC6325], some of
   which were also in [RFC7180], which this document replaces. These
   changes include the following:

   1. Changing nickname allocation to ignore conflicts with RBridges
       that are IS-IS unreachable.

   2. Fixing errors: [Err3002], [Err3003], [Err3004], [Err3052],
       [Err3053], and [Err3508].

   3. Changing the requirement to use the RPF check described in
       [RFC6325] for multi-destination TRILL Data packets by providing
       an alternative stronger RPF check.

   4. Adoption of the change of the CFI bit, which was required to be
       zero in the inner frame, to the DEI bit, which is obtained from
       inner frame ingress or creation.

   5. Requiring that all RBridges support E-L1FS FS-LSP flooding.

   6. Reducing the variable-length TRILL Header extensions area to one
       optional flags word. The Extension Length field (called
       "Op-Length" in [RFC6325]) is reduced to 1 bit that indicates
       whether the flags word is present. The rest of that Length field
       is now reserved.

   7. Changing the mandatory Appointed Forwarder Status Lost Counter
       increment provisions, as specified in Section 11.

C.3. Changes to RFC 7177

   All of the updates to [RFC7177] herein are in Section 9. Basically,
   this document requires that a Scope Flooding Support TLV [RFC7356]
   appear in all Hellos and that TRILL switches retain in their
   adjacency state the information received in that TLV.

C.4. Changes to RFC 7179

   The updates to [RFC7179] herein are in Sections 10.2 and 10.3.

Authors' Addresses

   Donald Eastlake 3rd
   Huawei Technology
   155 Beaver Street
   Milford, MA 01757
   United States

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

   Mingui Zhang
   Huawei Technologies
   No. 156 Beiqing Rd., Haidian District
   Beijing 100095
   China

   Email: zhangmingui@huawei.com

   Radia Perlman
   EMC
   2010 256th Avenue NE, #200
   Bellevue, WA 98007
   United States

   Email: radia@alum.mit.edu

   Ayan Banerjee
   Cisco

   Email: ayabaner@cisco.com

Anoop Ghanwani
Dell
5450 Great America Parkway
Santa Clara, CA 95054
United States

Email: anoop@alumni.duke.edu

Sujay Gupta
IP Infusion
RMZ Centennial
Mahadevapura Post
Bangalore 560048
India

Email: sujay.gupta@ipinfusion.com