Rbridges: TRILL Header Options

<draft-eastlake-trill-rbridge-options-01.txt>

Status of This Document

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

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

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

Abstract

The TRILL base protocol specification, draft-ietf-trill-rbridge-protocol-10.txt, specifies minimal hooks for options. This draft more fully describes the format for options and an initial set of options.
Table of Contents

Status of This Document........................................1
Abstract........................................................................1

Table of Contents..................................................2

1. Introduction....................................................3
1.1 Conventions used in this document.........................3

2. TRILL Header Options........................................4
2.1 RBridge Option Handling Requirements..................4
2.2 No Surprises..................................................5
2.3 Options Format.............................................5
2.3.1 Flag Options and Summary Bits.........................6
2.3.2 TLV Option Format......................................6
2.3.3 The Padding Option......................................7
2.3.4 Marshalling of Options................................8

3. Specific Flag Options.........................................9
3.1 ECN Flag Option............................................9

4. Specific TLV Options........................................10
4.1 Additional Flags TLV Option..............................10
4.2 Flow ID TLV Option.......................................11
4.3 Port ID TLV Option.......................................11
4.4 TRILL Security TLV Option..............................12

4. Additions to IS-IS...........................................14
4.1 Additions to Link State.....................................14
4.2 Additiona to Port Capabilities............................14

5. IANA Considerations.........................................15
6. Security Considerations......................................15
7. Acknowledgement.............................................15

8. Normative References.......................................16
9. Informative References......................................16

Disclaimer........................................................17
Additional IPR Provisions.........................................17
Author’s Address..................................................17
1. Introduction

The base TRILL protocol specification appears in [Protocol]. That specification provides an options feature and describes minimal hooks to incorporate that feature. But it does not specify the structure of options or the details of any particular options.

1.1 Conventions used in this document

The terminology and acronyms for [Protocol] are used herein with the same meaning.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.
2. TRILL Header Options

The TRILL Protocol includes an option capability in the TRILL Header (see [Protocol] Section 3.5). The Op-Length header field gives the length of the options in units of 4 octets which allows up to 124 octets of options area. If Op-Length is zero there are no options present; else, the options follow immediately after the Ingress Rbridge Nickname field in the TRILL Header.

As described below, provision is made for both hop-by-hop options, which could affect any RBridge which received a TRILL frame, and ingress-to-egress options, which would necessarily normally affect only the RBridge(s) where a TRILL frame is decapsulated. Provision is also made for both "critical" and "non-critical" options. An RBridge potentially affected by a critical option that it does not understand MUST discard the frame as it is unsafe to process the frame without understanding the option. Non-critical options can be safely ignored. Options also indicate whether the value associated with them can change (mutable options) or not (immutable options).

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

2.1 RBridge Option Handling Requirements

The requirements given in this section are in additional to all option handling requirements in [Protocol].

All RBridge MUST be able to detect whether there are any critical options present that are applicable to their processing of the frame as detailed below. If they do not implement all critical options present, they MUST discard the frame.

Transit RBridges MUST transparently forward any immutable ingress-to-egress options in frames they forward. Any changes made by a transit RBridge to a mutable ingress-to-egress option value MUST be a change permitted by the specification of that option. Note: Even though a transit RBridge might not examine or act on an ingress-to-egress option, the presence of that option may cause the frame to suffer from slow path processing.

In addition, a transit RBridge
  o MAY add a hop-by-hop option to a frame,
INTERNET-DRAFT                                      TRILL Header Options

- MAY add a padding option if none is present,
- MAY remove an unnecessary padding option,
- MAY adjust the length of an existing padding option,
- MAY remove a hop-by-hop option as specified for that option,
MAY change the value and Length of a mutable option as permitted by
  that option’s specification, but
- MUST NOT add or remove an ingress-to-egress option.
For any of these changes which alter the overall length of the TRILL
Header options area, the transit RBridge also adjusts the Header Op-
Length field.

2.2 No Surprises

RBridges advertise the options which they support in the core TRILL
IS-IS instance. An RBridge is not required to support any options; however, an RBridge which supports any other option MUST also support
the padding option.

No RBridge will receive a frame with a critical TRILL Header option
it must apply unless it advertised support for that option, except
due to errors or transient conditions. Should an RBridge receive a
frame with an applicable critical option it does not implement, it
MUST discard the frame.

If an RBridge is about to send a TRILL frame and the next hop
destination RBridge (or any of the next hop destination RBridges if
the frame is multi-destination) would not understand a critical
option in the frame that the next hop RBridge(s) might be required to
apply, it is the responsibility of the transmitting RBridge to remove
the option and make any necessary other adjustments to the frame
before transmission or drop the frame. (The transmitting RBridge
should understand the option or else it would not have received or
generated that critical option.)

TRILL options are generally inappropriate for any "extension" which
all RBridges in a campus would be required to understand or a
critical hop-by-hop option which cannot be backed out as described
immediately above. The addition of such an "extension" would likely
be a major change to the protocol and should probably be handled by a
revision to the TRILL protocol version number.

2.3 Options Format

If any options are present in a TRILL header, as indicated by a non-
zero Op-Length field, the first two octets of the options area
consist of two summary bits and 14 flag option bits as described in
Section 2.3.1. Section 2.3.2 specifies the format of an individual TLV option. TLV options appear in the options area after the first two octets. Further details on the padding option are specified in Section 2.3.3. Section 2.3.4 describes the marshalling of options.

2.3.1 Flag Options and Summary Bits

|   0      1    2  3  4  5  6  7| 8  9 10 11 12 13 14 15|  
|--------------------------------|
| CHbH | CItE |                 |                       |
+--------------------------------+

Figure 1: Options Area Initial Flags Octets

The following summary bit description text is copied from [Protocol] for convenience:

If the CHbH (Critical Hop by Hop) bit is one, one or more critical hop-by-hop options are present so transit RBridges that support no options MUST drop the frame. If the CHbH bit is zero, the frame is safe, from the point of view of options, for a transit RBridge. A transit RBridge that supports no options and forwards a frame MUST transparently forward the options area.

If the CItE (Critical Ingress to Egress) bit is a one, one or more critical ingress-to-egress options are present. If it is zero, no such options are present. If either CHbH or CItE is non-zero, egress RBridges that support no options MUST drop the frame. If both CHbH and CItE are zero, the frame is safe, from the point of view of options, for any egress RBridge to process.

2.3.2 TLV Option Format

Except for flag bit options described above, all other options to the TRILL Header are TLV (type, length, value) encoded, with some flag bits in the Type and Length octets, in the format show in Figure 2.

| 0  1  2  3  4  5  6  7| 0 1-7 |  
|----------------------------|
|IE|NC|      Type       |MT|       Length       |value...|
+----------------------------+

Figure 2. Option TLV Structure

The highest order bit of the first octet (IE) is zero for hop-by-hop options and one for ingress-to-egress options and the padding option.
Hop-by-hop options are potentially applicable to every RBridge which receives the frame. Ingress-to-egress options are only added at the ingress RBridge and are potentially applicable only at egress RBridges. Ingress-to-egress options MAY also be examined and acted upon by transit RBridges.

The next to highest order bit of the first octet (NC) is zero for critical options and one for non-critical options and the padding option. Non-critical options are those which can be safely ignored. Critical options are those which it is unsafe to ignore, for example options which indicate a change in the format of the remainder of the frame after the TRILL Header, such that attempts to parse this remainder could fail without understanding the critical option.

The bottom six bits of the first octet give the option Type code. The option Type may constrain the values of the IE, NC, and MT bits.

The highest order bit of the second octet (MT) is zero for options with immutable values, that is where the value and Length will not change. It is one for such options that have a mutable value. The IE, NC, Type, and MT fields themselves are always immutable.

The Length field is the unsigned length of the option value in octets. It gives the amount of option value data, if any, beyond the initial two octets. The Length field MUST NOT be such that the option value extends beyond the end of the total options area as specified by the TRILL Header Op-Length. Thus, the value of Length can vary from zero to 122. The meaning of "Length" values of 123 through 127 is reserved and, when such values are detected, they cause the frame to be discarded.

2.3.3 The Padding Option

The padding option is used for padding at the end of the options area of a TRILL Header. A padding option is required if the total length of other options present is not an exact multiple of 4 octets or otherwise falls short of the space indicated by Op-Length.

The padding option is Type 0x37 and MUST have the IE, NC, and MT bits equal to one (although it is not an ingress to egress option). An option with Type 0x37 where any of the IE, NC, and MT bits are zero is invalid and, if detected, causes the frame to be discarded.

A padding option MAY be included even if the length of the other options present is an exact multiple of 4 octets. Where padding is needed, it MAY be larger than strictly necessary; for example, an ingress RBridge might choose to round Op-Length up to an even value and pad any options it includes in a TRILL Header up to an exact
multiple of 8 octets to retain 64-bit alignment for the inner frame. All value octets in a padding option may be any value and need not be preserved by transit RBridges.

2.3.4 Marshalling of Options

In a TRILL Header with options, those options start immediately after the Ingress RBridge Nickname and completely fill the options area whose overall length is given in the Op-Length field.

TLV options MUST appear in ascending order by the value of their first octet considered as an unsigned 8 bit integer. As a result, all hop-by-hop options MUST be placed before all ingress-to-egress options and, within each of those two categories, all critical options MUST appear before all non-critical options. The padding option, if present, MUST appear last. A particular option first octet value MUST NOT appear more than once in a TRILL Header. Frames which violate this paragraph are erroneous, will produce unspecified results, and MAY be discarded. ("MAY" is chosen to minimize the format checking burden required of transit RBridges.)

Options are 16-bit aligned. Should an option consist of an odd number of octets, the option is padded at the end with one octet which MUST be zero. Should the total length of the options (other than the padding option) in a frame not fill the area indicated by the TRILL Header Op-Length, a padding option MUST be used to exactly fill the remaining space. This space will be 4*N or 2+4*N octets depending on whether the non-padding options present fill an even or odd number of double octets.

If any options are present present, those options, both flag and TLV, MUST be correctly summarized into the HbH and ItE bits
3. Specific Flag Options

The table below shows the state of TRILL Header Flag option bit assignments. See Section 7 for IANA Considerations.

<table>
<thead>
<tr>
<th>Bit</th>
<th>Purpose</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>Summary</td>
<td>2.3</td>
</tr>
<tr>
<td>2-7</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td>ECN</td>
<td>3.1</td>
</tr>
<tr>
<td>10-15</td>
<td>available</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Flag Options

3.1 ECN Flag Option

TBD - Explicit Congestion Notification
4. Specific TLV Options

The table below shows the state of TRILL Header TLV option Type assignment. See Section 6 for IANA Considerations.

<table>
<thead>
<tr>
<th>Type</th>
<th>Purpose</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00</td>
<td>reserved</td>
<td></td>
</tr>
<tr>
<td>0x01-0x03</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x04</td>
<td>Security</td>
<td>4.4</td>
</tr>
<tr>
<td>0x05-0x0F</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x10</td>
<td>Flags</td>
<td>4.1</td>
</tr>
<tr>
<td>0x11</td>
<td>Flow ID</td>
<td>4.2</td>
</tr>
<tr>
<td>0x12-0x2F</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x30</td>
<td>Port ID</td>
<td>4.3</td>
</tr>
<tr>
<td>0x31-0x3E</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x3F</td>
<td>Padding</td>
<td>2.3.3</td>
</tr>
</tbody>
</table>

Table 2. TLV Option Types

The following subsections specify particular TRILL options.

4.1 Additional Flags TLV Option

The option provides a means of adding a variety of additional flags to the TRILL Header beyond the limited number of flag option bits available in the first two octets of the options area.

The value of the flags option consists of additional flags, eight per octet, numbered from the high-order to the low-order bit. Thus flag 1 is the 0x80 bit of the first octet, flag 8 is the 0x01 bit of that octet, flag 9 is the 0x80 bit of the second octet, etc. The number of additional flags that can be defined is bounded only by the options space that can be available. All flags not present, because the would be in value octets beyond those specified by the option Length, are considered zero.

This option can appear up to four times in a frame to provide independent sets of all combinations of ingress-to-egress, hop-by-hop, non-critical, and critical flags. To simplify canonicalization for security, this option MUST NOT be included if all of the flag bits would be zero and the value MUST NOT have any trailing zero octets. Thus its Length MUST be at least 1 and at least the last octet of the value present MUST be non-zero.

The option fields and flags are as follows:

- Type is 0x10.
INTERNET-DRAFT

TRILL Header Options

- Length is variable with a minimum value of 1.
- IE and NC are variable producing, in effect, four versions of this option.
- MT MUST be zero. This is an immutable option.

4.2 Flow ID TLV Option

In connection with multi-pathing of frames, it is desirable that frames which are part of the same flow follow the same path. Methods to determine flows are beyond the scope of TRILL however, it may be useful, once the flow of a frame has been determined, to preserve and transmit that information for use by subsequent RBridges.

This is a non-critical option. It is considered hop-by-hop because it can be added by a transit RBridge. It can also affect transit RBridge behavior. Because the ingress RBridge or even the originating end station, which may have some way of signaling the ingress RBridge beyond the scope of this document, may know the most about a frame, it is expected that this option would most commonly be added at the ingress RBridge. Once in a frame, the option SHOULD NOT be removed or changed unless, for example, a campus is divided into regions such that different flow IDs would make the most sense in different regions.

The option fields and flags are as follows:

- Type is 0x11.
- Length is variable with a minimum value of 1. [Should a fixed flow ID size be specified?]
- IE MUST be zero. This is a hop-by-hop option.
- NC and MT MUST be one. This is a non-critical mutable option.

4.3 Port ID TLV Option

The purpose of the Port ID option is to avoid the destination MAC address to physical port mapping lookup at the egress RBridge. This might be beneficial for extremely high speed applications.

This option provides a 2 octet logical destination port and a 2 octet logical source port which, in some ways, could be considered extensions to the 6 octet inner destination and source MAC addresses in a frame. These logical port designators are local to the destination and source RBridges and may be any values those RBridges find convenient to efficiently map to their physical ports; however, the value 0x0000 is used to indicate that a logical port designator is unknown and the value 0xFFFF is reserved and MUST NOT appear in a
port ID option.

RBRidges that implement this option learn the Port ID for a remote
MAC address from the source Port ID field in the Port ID option, if
present, in frames they decapsulate in the same way they can learn
the egress RBridge and VLAN. This information is timed out in the
same manner as remote MAC address information. Such RBRidges include
their local Port ID in the source field of a Port ID option when
encapsulating a frame if inclusion of this option is indicated by
their local policy.

For known unicast TRILL data frames, one would expect ingress
RBRidges implementing this option to include it if sending to egress
RBRidges that also implement the option. For multi-destination TRILL
data frames, inclusion of a Port ID option with a source port ID may
make sense but the destination port ID is meaningless and ignored by
egress RBRidges.

The option fields and flags are as follows:

- Type is 0x30.
- Length MUST be 4.
- IE and NC MUST be one. This is an ingress-to-egress non-
critical option.
- MT MUST be zero. This is an immutable option.

4.4 TRILL Security TLV Option

TRILL provides a security option which builds on the IS-IS security
keying and can be applied to any TRILL frame.

The first octet of the option value is the same algorithm selection
code as for IS-IS. The value length for the option is variable and
depends on the algorithm in the same way as the value in the IS-IS
security TLV. Algorithm zero indicates a plain text password which
must be configured in code which generates and checks this TLV and is
NOT RECOMMENDED. Thus far, other algorithms have indicated HMAC
signing of a canonical form of the message using a shared secret
which must likewise be configured.

This option can appear up to twice in a frame, once for ingress-to-
egress security and once for hop-by-hop security.

[the following material needs more work]

For algorithms which depend on the value of the frame (i.e., all
confidentiality algorithms and all strong authentication algorithms),
the frame must be canonicalized before the authentication code is
computed or verified. This is logically done by copying the frame starting with the TRILL Header and, in the copy, setting the TRILL Header Hop Count to zero, clearing the octets of the Authentication Option after the algorithm selection code, and, for all mutable options, setting the option Length to zero and deleted any value octets. In addition, if an ingress-to-egress authentication code is being computed, since hop-by-hop options can be added or deleted in transit, all hop-by-hop options must be removed from the frame copy. Penultimately, any needed padding option must be reduced to its minimal length, that is, no padding option if the preceding options are an even multiple of 4 octets, or the minimum padding option of 0xFF80 if they are an odd multiple of 4 octets. Finally, the TRILL Header Op-Length must be adjusted downward as necessary to make it correct for the adjusted copy frame. The authentication code is then calculated using this copy and either inserted into the real frame for transmission or compared against the authentication code in the real frame for verification.

The option fields and flags are as follows:

- Type is 0x04.
- Length MUST be at least 1.
- IE is variable. There may be an ingress-to-egress or hop-by-hop security option in a frame or both.
- NC and MT MUST be zero. This is a critical immutable option.
4. Additions to IS-IS

RBridges use IS-IS PDUs to inform other RBridges which options they support.

4.1 Additions to Link State

Rbridges indicate in their link state which ingress-to-egress TLV option Types they support. In addition, if they support the ingress-to-egress Additional Flags TLV option, they indicate which critical ingress-to-egress Additional Flags TLV option flags they support.

4.2 Additions to Port Capabilities

Rbridges indicate in their Hellos which Hop-by-Hop TLV option Types they support. In addition, if they support the Hop-by-Hop Additional Flags TLV option, they indicate which critical hop-by-hop Additional Flags TLV option flags they support.
5. IANA Considerations

IANA will create two subregistry within the TRILL registry. One for TRILL Header flag optional which is initially populated as specified in Table 1 in Section 3. And a second for TRILL TLV Option Types which is initially populated as specified in Table 2 in Section 4. New TRILL Option types are allocated by an IETF Standards action as modified by [RFC4020].

IANA will create a third subregistry within the TRILL registry for flags in each of the four variations of the Flags option (the four combinations of critical and non-critical, ingress-to-egress and hop-by-hop). Such flags are allocated by TRILL Expert Approval.

6. Security Considerations

TBD

7. Acknowledgement

The Port ID option was initially suggested as part of the TRILL Header by Silvano Gai.
8. Normative References


9. Informative References

Disclaimer

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

Additional IPR Provisions

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

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

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

Copyright (C) The IETF Trust 2008 This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

Author’s Address

Donald E. Eastlake 3rd
Stellar Switches
155 Beaver Street
Milford, MA 01757

email: Donald.Eastlake@stellarswitches.com