Abstract

The TRILL base protocol specification, draft-ietf-trill-rbridge-protocol-13.txt, specifies minimal hooks for options. This draft specifies the format for options and an initial set of options.
Table of Contents

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

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 Bit Options and Summary Bits..............6
    2.3.2 TLV Option Format............................7
    2.3.3 Marshalling of Options.....................8

3. Specific Bit Option................................9
  3.1 ECN Bit Option..................................9

4. Specific TLV Options...............................11
  4.1 Flow ID TLV Option...............................11
  4.2 Additional Flags TLV Option....................12

5. Additions to IS-IS................................13
  5.1 Additions to Link State.........................13
  5.2 Additions to Port Capabilities.................13

6. IANA Considerations................................14
7. Security Considerations............................14

8. References..........................................15
  8.1 Normative References............................15
  8.2 Informative References..........................15

Authors’ Addresses....................................16
Copyright and IPR Provisions........................17
1. Introduction

The base TRILL protocol specification, which appears in [Protocol], provides an options feature and describes minimal hooks to safely support that feature. But it does not specify the structure of options nor the details of any particular options. This draft specifies that format and some initial options.

Section 2 below describes the general principles of operation, format, and ordering of TRILL Header Options. Such options are of two kinds: bit options, and TLV (Type, Length, Value) encoded options. Section 3 describes a specific bit option while Section 4 describes specific TLV encoded 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 [RFC2119].
2. TRILL Header Options

The TRILL Protocol includes an option capability in the TRILL Header (see [Protocol] Section 3.5). The 5-bit 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 and each option is 32-bit aligned.

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 only necessarily affect the RBridge(s) where a TRILL frame is decapsulated. Provision is also made for both "critical" and "non-critical" options. An RBridge receiving a frame with a critical option that might affect it and 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). For example, an ingress-to-egress security option could protect the value of an immutable ingress-to-egress option. But such a security option generally could not protect a mutable value as a transit RBridge could change that value but would not, in general, have the keys to recompute a signature or authentication code to take a changed value into account.

Note: Most R Bridges implementations are expected to be optimized for simple and common cases of frame forwarding and processing. Although the hard limit on options length, their 32-bit alignment, and the presence of critical option summary bits as described below, are intended assist in options processing, nevertheless the inclusion of options may cause frame processing using a "slow path" with inferior performance to "fast path" processing. Limited slow path throughput of such frames could cause them to be discarded.

2.1 RBridge Option Handling Requirements

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

All R Bridges 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 such options present, they MUST discard the frame.
Transit RBridges MUST transparently forward all immutable ingress-to-egress options in frames that 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.

In addition, a transit RBridge:

- MAY add, if space is available, or remove, hop-by-hop options as specified for that option;
- MAY change the value and/or length of a mutable option as permitted by that option’s specification (provided there is enough room if lengthening the option);
- MUST adjust the length of the options area, including changing Op-Length in the TRILL header as appropriate for any changes it has made in the options;
- MUST NOT add or remove an ingress-to-egress option.
- with regard to any non-critical hop-by-hop options that the transit RBridge does not understand, it MAY remove them if they are mutable but MUST transparently copy them when forwarding a frame if they are immutable.

2.2 No Surprises

RBridges advertise the ingress-to-egress options that they support in the core TRILL IS-IS instance and advertise the hop-by-hop options they support in Hellos they send. An RBridge is not required to support any options.

Unless an RBridge advertises support for a critical option, it would not normally receive frames with that option except due to errors or transient conditions.

An RBridge SHOULD NOT add a critical option to a frame unless,
- for a critical hop-by-hop option, it has determined that the next hop RBridge or RBridges to which the frame will be sent support that option, or
- for a critical ingress-to-egress option, it has determined that the RBridge or RBridges that will receive the frame with the option and egress it support that option.

2.3 Options Format

If any options are present in a TRILL header, as indicated by a non-zero Op-Length field, the first four octets of the options area consist of two summary bits and 30 option bits as described below. The remainder of the options area consists of TLV (Type Length Value)
encoded options aligned on 32-bit boundaries. Section 2.3.2 specifies the format of an individual TLV option. Section 2.3.3 describes the marshalling of TLV options.

### 2.3.1 Bit Options and Summary Bits

```
+------+------+--+--+--+--+--+--+-------+-------+-------+
| CHbH | CItE |                 |       |       |       |
+------+------+--+--+--+--+--+--+-------+-------+-------+
```

Figure 1: Options Area Initial Four Octets

The top two bits of the options area, bits 0 and 1 above, are called summary bits and summarize the presence of critical options. 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 R Bridges that support no options MUST drop the frame. If the CHbH bit is zero, the frame is safe, from the point of view of options processing, for a transit R Bridge to forward, even if the forwarding R Bridge doesn’t understand any options. A transit R Bridge 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 R Bridges that don’t support any 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 R Bridge to process, even if it doesn’t understand any options.

The remaining 30 bits in the initial four octets of the options area are available for bit-encoded options. Any R Bridge adding an options area to a TRILL Header must set these 30 bits to zero except when permitted to set one or more by an option that R Bridge understands. The 30 bits are categorized as follows:

<table>
<thead>
<tr>
<th>Bits</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-7</td>
<td>Critical hop-by-hop</td>
</tr>
<tr>
<td>8-15</td>
<td>Non-critical hop-by-hop</td>
</tr>
<tr>
<td>16-23</td>
<td>Critical ingress-to-egress</td>
</tr>
<tr>
<td>24-31</td>
<td>Non-critical ingress-to-egress</td>
</tr>
</tbody>
</table>

Any transit R Bridge must transparently copy bits 2-31 except as
permitted by an option implemented by that RBridge. Even if a transit RBridge removes all TLV options from a TRILL Header, it MUST NOT eliminate the options area in a forwarded frame if any of these 30 bits is non-zero.

2.3.2 TLV Option Format

TRILL Header options, other than bit options described above, are TLV 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| 8|    9-15         |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+---
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. Hop-by-hop options are potentially applicable to every RBridge that 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 as specified in the particular option.

The next to highest order bit of the first octet (NC) is zero for critical options and one for non-critical options.

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 movable value. The IE, NC, Type, and MT fields themselves are always immutable.

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. For example, if the Type indicates a Flow ID option, then it MUST be marked as a hop-by-hop, non-critical, mutable option. If the IE, NC, or MT bits have a value not permitted by the option Type specification for an option that an RBridge would otherwise act on, the RBridge MUST discard the frame.

The Length field is an unsigned quantity giving the length of the option value in octets. It gives the amount of option value data, if any, beyond the initial two Type and Length 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 120. The meaning of
"Length" values of 121 through 127 is reserved and, when such values are noticed in a frame, the frame MUST be discarded.

2.3.3 Marshalling of Options

In a TRILL Header with options, those options start immediately after the Ingress RBridge Nickname and completely fill the options area.

TLV options start immediately after the initial four octets of option and summary bits and 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. A particular option first octet value MUST NOT appear more than once in a TRILL Header. Frames that 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 32-bit aligned. Should an option not consist of a multiple of four octets, the option is padded at the end up to the next multiple of four octets with octets that MUST be zero.

If any options are present, those options, both flag and TLV, MUST be correctly summarized into the CHbH and CItE bits at the top of the initial two octets of the options area.
3. Specific Bit Option

The table below shows the state of TRILL Header bit option assignments. See Section 6 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 for critical hop-by-hop options</td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td>ECN</td>
<td>3.1</td>
</tr>
<tr>
<td>10-15</td>
<td>available for non-critical hop-by-hop options</td>
<td></td>
</tr>
<tr>
<td>16-23</td>
<td>available for critical ingress-to-egress options</td>
<td></td>
</tr>
<tr>
<td>24-31</td>
<td>available for non-critical ingress-to-egress options</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Flag Options

3.1 ECN Bit Option

RBRidges may implement an ECN (Explicit Congestion Notification) option [RFC3168]. If implemented, it SHOULD be enabled by default but can be disable on a per RBridge basis by configuration.

RBRidges that do not implement this option or on which it is disabled simply (1) set bits 8 and 9 of the bit options area zero when they add an options area to a TRILL Header and (2) transparently copy those bits, if an options area is present, when they forward a frame with a TRILL Header.

An RBridge that implements the ECN option does the following when that option is enabled:

- When ingressing an IP frame that is ECN enabled, it MUST add an options area to the TRILL Header and copy the two ECN bits from the IP header into option bits 8 and 9.
- When ingressing a frame for a non-IP protocol with a means of indicating ECN that is understood by the RBridge, it MAY add an options are to the TRILL Header with the ECN bits set from the ingressed frame.
- When forwarding a frame encountering congestion at an RBridge, if an options area is present with option bits 8 and 9 indicate ECN-capable transport, the RBridge MUST modify them to the congestion experienced value.
- When egressing an IP frame, if the TRILL Header has an options area with option bits 8 and 9 non-zero, it copies those bits into the ECN bits in the IP header.
- When egressing a non-IP protocol frame with a means of indicating ECN that is understood by the RBridge, it MAY transfer the ECN information from the ECN bits in the options area to the egressed
native frame.

The following table is modified from [RFC3168] and shows the meaning of bit values in TRILL Header option bits 8 and 9, bits 6 and 7 in the IPv4 TOS Byte, and bits 6 and 7 in the IPv6 Traffic Class Octet:

<table>
<thead>
<tr>
<th>Binary</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Not-ECT (Not ECN-Capable Transport)</td>
</tr>
<tr>
<td>01</td>
<td>ECT(1) (ECN-Capable Transport(1))</td>
</tr>
<tr>
<td>10</td>
<td>ECT(0) (ECN-Capable Transport(0))</td>
</tr>
<tr>
<td>11</td>
<td>CE (Congestion Experienced)</td>
</tr>
</tbody>
</table>

Table 2. ECN Bit Combinations

An RBridge detects congestion either by monitoring its own queue depths or from participation in a link-specific protocol. An RBridge implementing the ECN option MAY be configured to add congestion experienced marking using ECN to any frame with a TRILL Header that encounters congestion even if the frame was not previously marked as ECN-capable.
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>0x01-0x07</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x08</td>
<td>Flow ID</td>
<td>4.1</td>
</tr>
<tr>
<td>0x09-0x3B</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x3C</td>
<td>Additional Flags</td>
<td>4.2</td>
</tr>
<tr>
<td>0x39-0x3E</td>
<td>available</td>
<td></td>
</tr>
<tr>
<td>0x3F</td>
<td>reserved</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. TLV Option Types

The following subsections specify particular TRILL TLV options.

4.1 Flow ID TLV Option

In connection with multi-pathing of frames, frames that are part of the same order dependent flow need to follow the same path for correct performance. Methods to determine flows are beyond the scope of the TRILL standard; however, it may be useful, once the flow of a frame has been determined, to preserve and transmit that information for use by subsequent R Bridges.

This is a non-critical option. It is considered hop-by-hop because it can be added by a transit RBridge and can be used by transit R Bridges to make forwarding decisions. Because the ingress RBridge 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 sense in different regions.

The value length of this option is fixed for efficiency. Since 2 octets might be insufficient for some purposes, the length is set at six, the next size fitting evenly into the 32-bit alignment of options. Should there be less than 6 octets of significance to the flow ID, the value SHOULD be right justified by prefixing zero octets. If an RBridge is incapable of using all six octets for flow ID purposes, it SHOULD use a smaller number of lower order octets from the value.

The option fields and flags are as follows:
Type is 0x08.
Length is 6. The data is an unsigned integer that is the flow ID. If there are less than six value octets of significance, they SHOULD be right justified by prefixing zero octets.
IE MUST be zero. This is a hop-by-hop option.
NC and MT MUST be one. This is a non-critical mutable option.

4.2 Additional Flags TLV Option

The option provides a means of adding a variety of additional flags to the TRILL Header beyond the bit options available in the first four 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 they 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 0x3C.
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.
5. Additions to IS-IS

RBRidges use IS-IS PDUs to inform other RBRidges which options they support. The specific IS-IS TLVs or sub-TLVs used to encode this information are specified in a separate document.

5.1 Additions to Link State

RBRidges indicate in their link state which ingress-to-egress TLV and bit options 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, if any.

5.2 Additions to Port Capabilities

RBRidges indicate in their Hellos which hop-by-hop TLV and bit options they support. In addition, if they support the Additional Flags TLV option, they indicate which critical hop-by-hop Additional Flags TLV option flags they support.
6. IANA Considerations

IANA will create two subregistries within the TRILL registry. A "TRILL Header Bit Options" subregistry that is initially populated as specified in Table 1 in Section 3. And a "TRILL TLV Option Types" subregistry that is initially populated as specified in Table 3 in Section 4. References in both of those tables to sections of this document are to be replaced in the IANA subregistries by references to this document as an RFC.

New TRILL bit options and TLV option types are allocated by IETF Review [RFC5226].

IANA will create a third subregistry within the TRILL registry for flags in the four variations of the Additional Flags TLV option (the four combinations of critical and non-critical, ingress-to-egress and hop-by-hop) which is initially empty. Such flags are allocated by RFC Publication if the RFC allocates no more than three bits, which may be a mixture of the four types, or by IETF Review for any number of bits [RFC5226].

7. Security Considerations

For general TRILL protocol security considerations, see [Protocol].

RBrigdes should not trust that the options that appear in a TRILL header reflect the intent of the previous or earlier hop RBrigdes, including the ingress RBrigde, unless the option is appropriately secured. For example, through inclusion of a TRILL authentication option, which may be specified in another document.

In order to facilitate authentication, options SHOULD be specified so they do not have alternative equivalent forms. Authentication of anything with alternative equivalent forms almost always requires canonicalization which an authenticating RBrigde ignorant of the option would be unable to do.
8. References

8.1 Normative References


8.2 Informative References

None at this time.
Authors’ Addresses

Donald E. Eastlake 3rd  
Stellar Switches  
155 Beaver Street  
Milford, MA 01757  
tel: +1-508-634-2066  
email: d3e3e3@gmail.com  

Anoop Ghanwani  
Brocade Communications Systems  
1745 Technology Drive  
San Jose, CA 95110 USA  
Phone: +1-408-333-7149  
Email: anoop@brocade.com

Caitlin Bestler  
Consultant  
555 E. El Camino Real #104  
Sunnyvale, CA 94087  
tel: +1-949-528-3085  
email: cait@asomi.com
Copyright and IPR Provisions

Copyright (c) 2009 IETF Trust and the persons identified as the
document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal
Provisions Relating to IETF Documents in effect on the date of
publication of this document (http://trustee.ietf.org/license-info).
Please review these documents carefully, as they describe your rights
and restrictions with respect to this document.

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