Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 10, 2013

M. Wasserman
Painless Security
D. Eastlake
Huawei R&D USA
D. Zhang
Huawei Technologies
September 6, 2012

Transparent Interconnection of Lots of Links (TRILL) over IP
draft-mrw-trill-over-ip-02.txt

Abstract

The Transparent Interconnection of Lots of Links (TRILL) protocol is implemented by devices called Routing Bridges (RBridges). TRILL supports both point-to-point and multi-access links and is designed so that a variety of link protocols can be used between RBridge ports. This document standardizes methods for encapsulating TRILL in UDP/IP(v4 or v6) to provide a unified TRILL campus.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on March 10, 2013.

Copyright Notice

Table of Contents

1. Requirements Terminology

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 [RFC2119].

2. Introduction

RBridges are devices that implement the IETF TRILL protocol [RFC6325] [RFC6326] [RFC6327].

RBridges provide transparent forwarding of frames within an arbitrary network topology, using least cost paths for unicast traffic. They support VLANs and multipathing of unicast and multi-destination traffic. They use IS-IS link state routing and encapsulation with a hop count. The are compatible with IEEE 802.1 customer bridges, and can incrementally replace them.

Two or more RBridges can communicate over a variety of different link types, such as Ethernet [RFC6325] or PPP [RFC6361].

This document defines a method for RBridges to communicate over UPD/IP(v4 or v6). TRILL over IP will allow remote, Internet-connected RBridges to form a single RBridge campus, or multiple TRILL over IP networks within a campus to be connected as a single TRILL campus via a TRILL over IP backbone.

TRILL over IP connects RBridge ports using IPv4 or IPv6 as a transport in such a way that the ports appear to TRILL to be connected by a single link. The link will be a multi-access link if more than two RBridge ports are connected via a single TRILL over IP link, so that any pair of ports can communicate.

To support cases where RBridges are connected via links (such as the public Internet) that are not under the same administrative control as the TRILL campus, this document specifies the use of Datagram Transport Layer Security (DTLS) [RFC4327] to secure communication between RBridges running TRILL over IP.

3. Use Cases for TRILL over IP

In this document, we consider two use cases that are typical of situations where network administrators may choose to use TRILL over an IP network: a remote office scenario, and an IP backbone scenario.

3.1. Remote Office Scenario

In the Remote Office Scenario, a remote TRILL network is connected to a TRILL campus across a multihop non-TRILL IP network, such as the public Internet. The TRILL network in the remote office becomes a logical part of TRILL campus, and nodes in the remote office can be attached to the same VLANs as local campus nodes. In many cases, a remote office may be attached to the TRILL campus by a single pair of RBridges, one on the campus end, and the other in the remote office. In this use case, the TRILL over IP link will often cross logical and physical IP networks that do not support TRILL, and are not under the same administrative control as the TRILL campus.

3.2. IP Backbone Scenario

In the IP Backbone Scenario, TRILL over IP is used to connect a number of TRILL networks to form a single TRILL campus. For example, a TRILL over IP backbone could be used to connect multiple TRILL networks on different floors of a large building, or to connect TRILL networks in separate buildings of a multi-building site. In this use case, there may often be several TRILL RBridges on a single TRILL over IP link, and the IP link(s) used by TRILL over IP are typically under the same administrative control as the rest of the TRILL campus.

3.3. Important Properties of the Scenarios

There are a number of differences between the two scenarios listed above, some of which drive features of this specification. These differences are especially pertinent to the security requirements of the solution, how multicast data frames are handled, and how the RBridges discover each other.

3.3.1. Security Requirements

In the IP Backbone Scenario, TRILL over IP is used between a number of RBridges, on a network link that is in the same administrative control as the remainder of the TRILL campus. While it is desirable in this scenario to prevent the association of rogue RBridges, this can be accomplished using existing IS-IS security mechanisms. There may be no need to protect the data traffic, beyond any protections that are already in place on the local network.

In the Remote Office Scenario, TRILL over IP may run over a network that is not under the same administrative control as the TRILL network. Nodes on the network may think that they are sending traffic locally, while that traffic is actually being sent, in a UDP/IP tunnel, over the public Internet. It is necessary in this

scenario to protect user privacy, as well as ensuring that no unauthorized RBridges can gain access to the RBridge campus. The data privacy requirement is addressed by the use of DTLS for both IS-IS frames and data frames between RBridges in this scenario.

3.3.2. Multicast Handling

In the IP Backbone scenario, native mutlicast may be supported on the TRILL over IP link. If so, it will be used to send TRILL IS-IS and multicast data frames, as discussed later in this document.

In the Remote Office Scenario, there will often be only one pair of RBridges connecting a given site, and even when multiple RBridges are used to connect a Remote Office to the TRILL campus, the intervening network may not provide reliable (or any) multicast connectivity. Also, it is difficult to provide strong data privacy for multicast traffic. For all of these reasons, the connections between local and remote RBridges will be treated like point-to-point links, and all TRILL IS-IS control messages and multicast data frames that are transmitted between the Remote Office and the TRILL campus will be serialized, as discussed later in this document.

3.3.3. RBridge Discovery

In the IP Backbone Scenario, RBridges that use TRILL over IP will use the normal TRILL IS-IS Hello mechanisms to discover the existence of other RBridges on the link [RFC6327], and to establish authenticated communication with those RBridges.

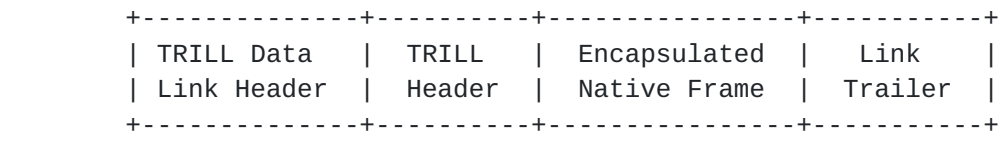
In the Remote Office Scenario, a DTLS session will need to be established between RBridges before TRILL IS-IS traffic can be exchanged, as discussed below. In this case, one of the RBRidges will need to be configured to establish a DTLS session with the other RBridge. This will typically be accomplished by configuring the RBridge at a Remote Office to initiate a DTLS session, and subsequent TRILL exchanges, with a TRILL over IP-enabled RBridge attached to the TRILL campus.

4. TRILL Frame Formats

To support the TRILL base protocol standard [RFC6325]. , two types of frames will be transmitted between RBridges: TRILL Data frames and TRILL IS-IS frames.

4.1. TRILL Data Frame

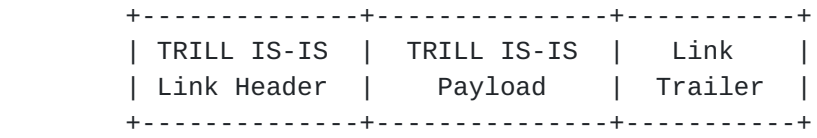
The on-the-wire form of a TRILL Data frame in transit between two neighboring RBridges is as shown below:



Where the Encapsulated Native Frame is in Ethernet frame format with a VLAN tag but with no trailing Frame Check Sequence (FCS).

4.2. TRILL IS-IS Frame

TRILL IS-IS frames are formatted on-the-wire as follows:



The Link Header and Link Trailer in these formats depend on the specific link technology. The Link Header usually contains one or more fields that distinguish TRILL Data from TRILL IS-IS. For example, over Ethernet, the TRILL Data Link Header ends with the TRILL Ethertype while the TRILL IS-IS Link Header ends with the L2-IS-IS Ethertype; on the other hand, over PPP, there are no Ethertypes but PPP protocol code points are included that distinguish TRILL Data from TRILL IS-IS.

In TRILL over IP, we will use UDP/IP (v4 or v6) as the link header, and the TRILL frame type will be determined based on the UDP port number. In TRILL over IP, no Link Trailer is specified, although one may be added when the resulting IP packets are encapsulated for transmission on a network (e.g. Ethernet).

5. Link Protocol Specifics

TRILL Data packets can be unicast to a specific RBridge or multicast to all RBridges on the link. TRILL IS-IS packets are always multicast to all other RBridge on the link (except for TRILL IS-IS MTU PDUs, which may be unicast). On Ethernet links, the Ethernet

multicast address All-RBridges is used for TRILL Data and All-IS-IS-RBridges for TRILL IS-IS.

To properly handle TRILL base protocol frames on a TRILL over IP link, either native multicast mode must be enabled on that link, or multicast must be simulated using serial unicast, as discussed below.

In TRILL Hello PDUs used on TRILL IP links, the IP addresses of the connected IP ports are their SNPA addresses. Thus, all TRILL Neighbor TLVs in such Hellos MUST specify that the size of the SNPA is 4-bytes for an IPv4 link or 16-bytes for an IPv6 link [rfc6326bis]. Note that SNPA addresses and their size are independent of TRILL System IDs which are 6-bytes.

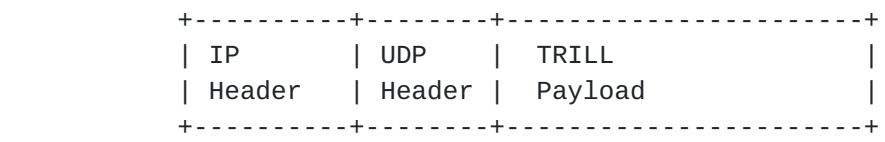
6. Port Configuration

Each RBridge port used for a TRILL over IP link MUST have at least one IP (v4 or v6) address. Implementations MAY allow a single physical port to operate as multiple IPv4 and/or IPv6 logical ports.

TBD: MUST be able to configure list of IP addresses for serial unicast. MUST be able to configure non-standard IP multi-cast addresses.

7. TRILL over UDP/IP Format

The general format of a TRILL over UDP/IP packet is shown below.



Where the UDP Header is as follows:

TBD

8. Handling Multicast

8.1. Multicast of TRILL IS-IS Packets

By default, TRILL IS-IS packets are sent to an IPv4 multicast address.

8.2. Multicast Data Frames

TBD

9. Use of DTLS

All RBridges that support TRILL over IP MUST implement DTLS and support the use of DTLS to secure both TRILL IS-IS and data traffic. When DTLS is used to secure a TRILL over IP link, the DTLS session MUST be fully established before any TRILL IS-IS or data frames are exchanged.

RBridges that implement TRILL over IP MUST support the use of certificates for DTLS, and MUST support the following algorithm:

o TLS_RSA_WITH_AES_128_CBC_SHA [RFC5246]

RBridges that support TRILL over IP MAY support the use of pre-shared keys for DTLS. If the communicating RBridges have IS-IS authentication enabled with a pre-shared key, then that key may be used for DTLS unless some other pre-shared key is configured. If pre-shared keys are supported, the following cryptographic algorithms MUST be supported for use with pre-shared keys:

o TLS_PSK_WITH_AES_128_CBC_SHA [RFC5246]

10. Transport Considerations

10.1. Recursive Encapsulation

TRILL is designed to transport end station Ethernet traffic and IP is frequently transported over Ethernet. Thus, an end station Ethernet frame EF might get TRILL encapsulated to TRILL(EF) which was then send on a TRILL over IP over Ethernet link resulting in an Ethernet frame of the form Ethernet(IP(TRILL(EF))). There is a risk of such an Ethernet frame being re-ingressed by the same TRILL campus, due to physical or logical misconfiguration, looping round, being further encapsulated and re-ingressed, etc. The frame might get discarded if it got too large but if fragmentation is enabled, it would just keep getting split into fragment that would continue to loop and grow and re-fragment until the path was saturated with junk and frames were

being discarded due to queue overflow. TTL would provide no protection because each TRILL encapsulation adds a new TTL.

To protect against this scenario, TRILL over IP output ports MUST be able to test whether a TRILL frame they are above to send is, in fact a TRILL encapsulation of a TRILL over IP over Ethernet frame. That is, is it of the form TRILL(Ethernet(IP(TRILL(...))). If so, the default action of the TRILL over IP output port is to discard the frame. However, there are cases where some level of multiple encapsulations is desired so it MUST be possible to configure the port to allow such frames.

11. MTU Considerations

In TRILL each RBridge advertises the largest LSP frame it can accept (but not less than 1,470 bytes) on any of its interfaces (at least those interfaces with adjacencies to other RBridges in the campus) in its LSP number zero through the originatingLSPBufferSize TLV [RFC6325] [rfc6326bis]. The campus minimum MTU, denoted Sz, is then established by taking the minimum of this advertised MTU for all RBridges in the campus. Links that do not meet the Sz MTU are not included in the routing topology. This protects the operation of IS-IS from links that would be unable to accommodate some LSPs.

Exact methods of determining originatingLSPBufferSize for an RBridge with one or more TRILL over IP ports are beyond the scope of this document. However, if an IP link either can accommodate jumbo frames or is a link on which IP fragmentation is enabled and acceptable, then it is unlikely that the IP link will be a constraint on the RBridge's originatingLSPBufferSize. On the other hand, if the IP link can only handle smaller frames and fragmentation is to be avoided when possible, a TRILL over IP port might constrain the RBridge's originatingLSPBufferSize. Because TRILL sets the minimum values of Sz at 1,470 bytes, there may be links that meet the minimum MTU for the IP protocol (1,280 bytes for IPv6, theoretically 68 bytes for IPv4) on which it would be necessary to enable fragmentation for TRILL use.

The optional use of TRILL IS-IS MTU PDUs, as specified in [RFC6325] and [RFC6327] can provide added assurance of the actual MTU of a link.

12. Middlebox Considerations

TBD

13. Security Considerations

TRILL over IP is subject to all of the security considerations for the base TRILL protocol. In addition, there are specific security requirements for different TRILL deployment scenarios, as discussed in the "Use Cases for TRILL over IP" section above.

This document specifies that all RBridges that support TRILL over IP MUST implement DTLS, and makes it clear that it is both wise and good to use DTLS in all cases where a TRILL over IP link will traverse a network that is not under the same administrative control as the rest of the TRILL campus. DTLS is necessary, in these cases to protect the privacy and integrity of data traffic.

TRILL over IP is completely compatible with the use of IS-IS security, which can be used to authenticte RBridges before allowing them to join a TRILL campus. This is sufficient to protect against rogue RBridges, but is not sufficient to protect data frames that may be sent, in UDP/IP tunnels, outside of the local network, or even across the public Internet. To protect the privacy and integrity of that traffic, use DTLS.

In cases were DTLS is used, the use of IS-IS security may not be necessary, but there is nothing about this specification that would prevent using both DTLS and IS-IS security together. In cases where both types of security are enabled, implementations MAY allow users to configure a single shared key that will be used for both mechanisms.

14. IANA Considerations

IANA has allocated the following UDP Ports for the TRILL IS-IS and Data channels:

UDP Port	Protocol
(TBD)	TRILL IS-IS Channel
(TBD)	TRILL Data Channel

IANA has allocated one IPv4 and one IPv6 multicast address, as shown below, which correspond to the All-RBridges and All-IS-IS-RBridges multicast MAC addresses that the IEEE Registration Authority has assigned for TRILL. Because the low level hardware MAC address dispatch considerations for TRILL over Ethernet do not apply to TRILL over IP, one IP multicast address for each version of IP is

sufficient.

[Values recommended to IANA:]

Name	IPv4	IPv6
All-RBridges	233.252.14.0	FF0X:0:0:0:0:0:0:205

Note: when these IPv4 and IPv6 multicast addresses are used and the resulting IP frame is sent over Ethernet, the usual IP derived MAC address is used.

[Need to discuss scopes for IPv6 multicast (the "X" in the addresses) somewhere. Default to "site" scope but MUST be configurable?]

15. Acknowledgements

This document was written using the xml2rfc tool described in RFC 2629 [RFC2629].

The following people have provided useful feedback on the contents of this document: Sam Hartman.

16. References

16.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4327] Dubuc, M., Nadeau, T., Lang, J., and E. McGinnis, "Link Management Protocol (LMP) Management Information Base (MIB)", RFC 4327, January 2006.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

[RFC6325] Perlman, R., Eastlake, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.

[RFC6326] Eastlake, D., Banerjee, A., Dutt, D., Perlman, R., and A. Ghanwani, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 6326, July 2011.

[RFC6327] Eastlake, D., Perlman, R., Ghanwani, A., Dutt, D., and V. Manral, "Routing Bridges (RBridges): Adjacency", RFC 6327, July 2011.

16.2. Informative References

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

[RFC6361] Carlson, J. and D. Eastlake, "PPP Transparent Interconnection of Lots of Links (TRILL) Protocol Control Protocol", RFC 6361, August 2011.

Authors' Addresses

Margaret Wasserman
Painless Security
356 Abbott Street
North Andover, MA 01845
USA

Phone: +1 781 405-7464
Email: mrw@painless-security.com
URI: http://www.painless-security.com

Donald Eastlake
Huawei R&D USA
155 Beaver Street
Milford, MA 01757
USA

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

Dacheng Zhang
Huawei Technologies
Q14, Huawei Campus
No.156 Beiqing Rd.
Beijing, Hai-Dian District 100095
P.R. China

Phone:
Email: zhangdacheng@huawei.com
URI: