Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 10, 2012

X. Wan
X. Yang
HangZhou H3C Co. Limited
V. Manral
A. Retana
Hewlett-Packard Co.

June 8, 2012

Extending TRILL over WAN
draft-xl-trill-over-wan-01.txt

Abstract

TRILL is a key technology for large-scale layer 2 networking within a data center, TRILL over WAN_(ToW) provides a scalable and simple solution that interconnect multiple TRILL networks to form a single TRILL domain using the currently deployed enterprise or service provider networks. This document provides an overview of this solution.

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 December 10, 2012.

Table of Contents

1. Overview

TRILL over WAN is a technology for supporting L2 VPNs over an L2/L3 infrastructure. It provides an "over-the-top" method of doing virtualization among several sites where the routing and forwarding state is maintained at the network edge, but not within the site or in the core.

TRILL over WAN can reside in a small number of device at the edge between TRILL sites and the core, we call these devices "Joint Devices" which perform typical transit Rbridges functions(nickname-based forwarding) and perform overlay functions on their core facing interfaces.

TRILL traffic which requires traversing the WAN to reach its destination, is prepended with an IP header. As shown in figure1, if a destination RBridge(Nickname) is reachable via Joint Device S2(with a core facing IP address of IPB), other Joint Devices forwarding traffic to such Rbridge will add on IP header with a destination IP address of IPB and forward the traffic into the core. The core will forward traffic based on IP address IPB, once the traffic makes it to Joint Device S2 it will be stripped of the overlay IP header and it will be forwarded into the site in the same way a regular RBridge would forward a packet. Broadcast or multicast traffic is encapsulated with a multicast header and follows a similar process.

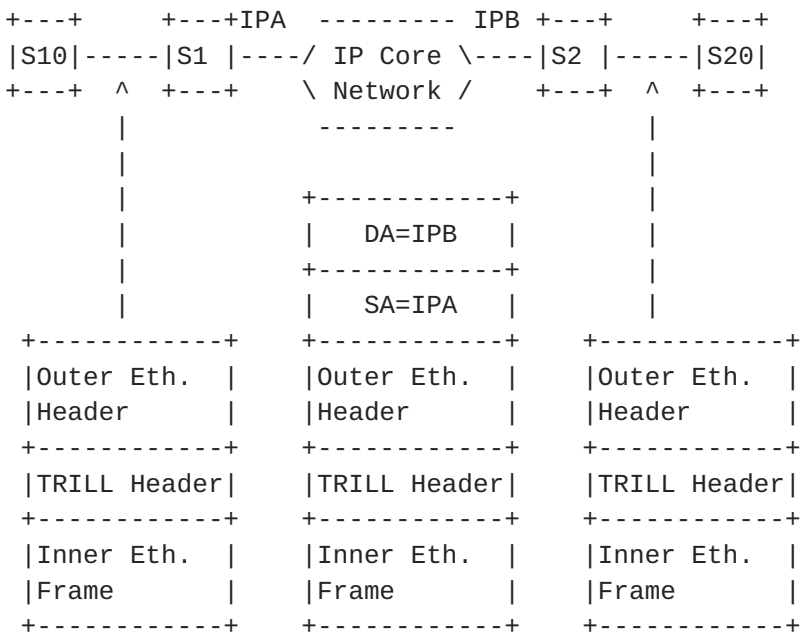
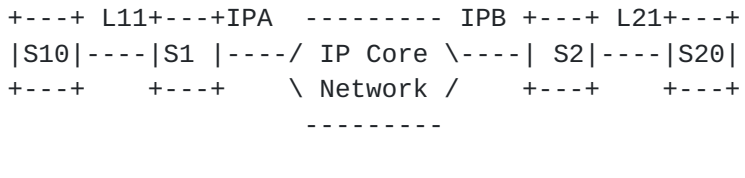


Figure 1: Traffic encapsulation

The key piece that TRILL over WAN adds is the state to map a given egress Nickname to an IP address of the Joint Device behind which that egress Nickname is located. TRILL over WAN forwarding is a function of mapping a destination Nickname to a Joint Device IP address in the overlay network.

To achieve this, a control plane is required to exchange the reachability information among different Joint Devices. After Joint Device discovery and adjacency setup, the virtual links to neighbor Joint Devices will be treated as TRILL interface, and the TRILL's control plane runs over these virtual links. Through these steps, all the Rbridges in multiple site forms a single TRILL domain. Each Rbridge (including Joint Device) calculates its TRILL forwarding table independently. Figure 2 shows what the resulting forwarding tables look like in a simple example.



S1

Nickname	Interface
S10	L11
S20	IPB
S1	self
S2	IPB

S2

Nickname	Interface
S10	IPA
S20	L21
S1	IPA
S2	self

Figure 2: Forwarding Tables

TRILL over WAN supports multi-homing for sites where one or more of the Joint Devices connected to core network. It can support active-active multi-homing capability and loop elimination by nature of TRILL. No need to add other extra mechanism.

2. 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].

[[anchor2: NOTE TO RFC EDITOR: Please remove the following text before publication.]]
Some ideas of this specification is being discussed on the EAI mailing list. See https://www1.ietf.org/mailman/listinfo/ima for information about subscribing. The list's archive is at http://www1.ietf.org/mail-archive/web/ima/index.html.

3. Control Plane

3.1. Provider Control Plane

The provider control plane enables unicast reachability among the Joint Devices and also provides the multicast group than makes Joint Devices adjacent from the overlay control plane perspective. It also provides the multicast trees in the core that will be used for optimal forwarding of the TRILL data traffic.

3.2. Overlay Control Plane

The overlay control plane provides auto-discovery of the Joint Devices that are members of an overlay VPN.

The TRILL control plane traffic between Joint Devices of different sites will be carried over the virtual link.

Detailed to be added.

3.2.1. Edge Device Discovery and Adjacency setup

See draft-hasmit-otv-03 2.2.1.

3.2.2. Control Plane Packet Encapsulation and Relay

Any Jonit Device of the site is virtual link with the other Jonit Devices of the other sites. Any Jonit Device should encapsulate the ISIS routing information of its site, and then relay it to the other Joint Devices of the other sites.

There is no different between the virtual links and the physical links to the TRILL protocol. The TRILL protocol calculates the

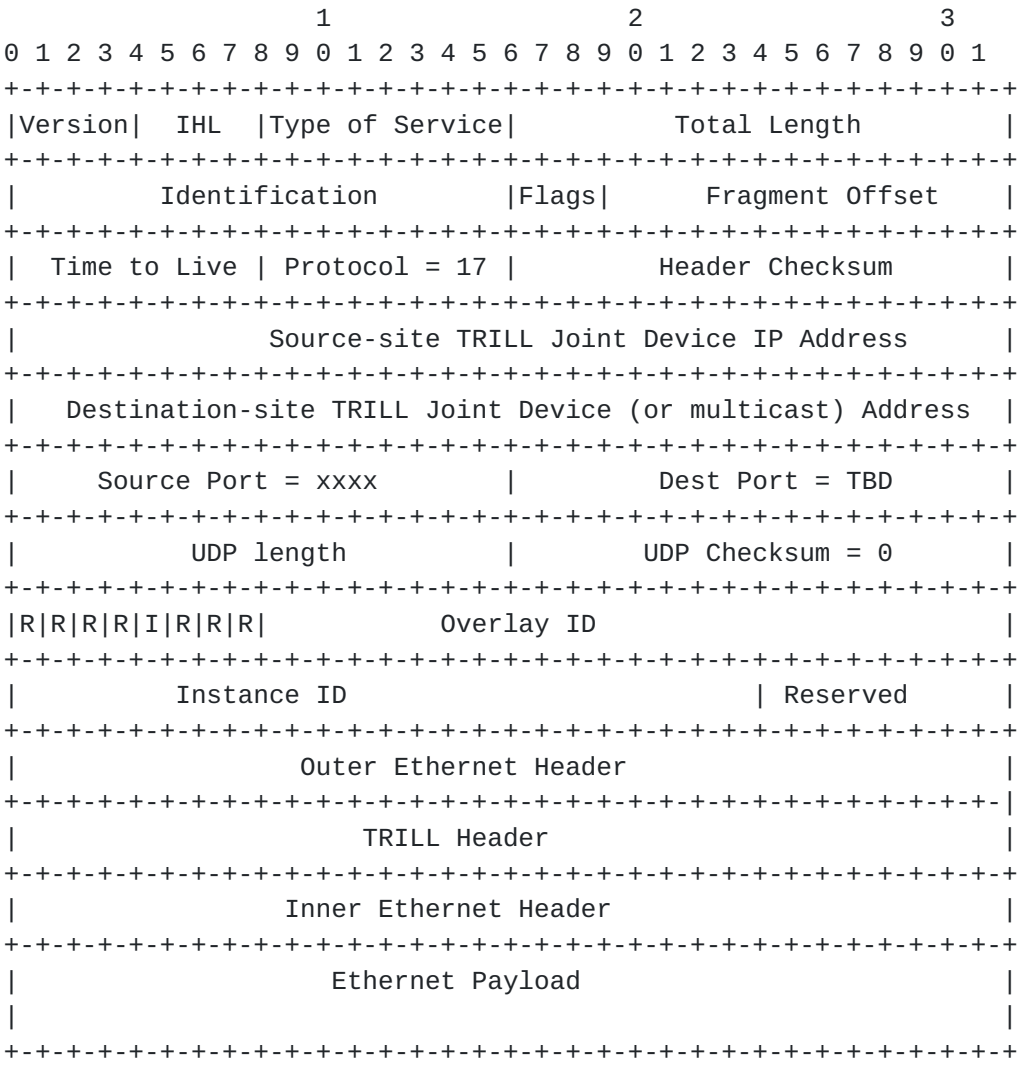
routing information among the whole TRILL domain.

4. Data Plane

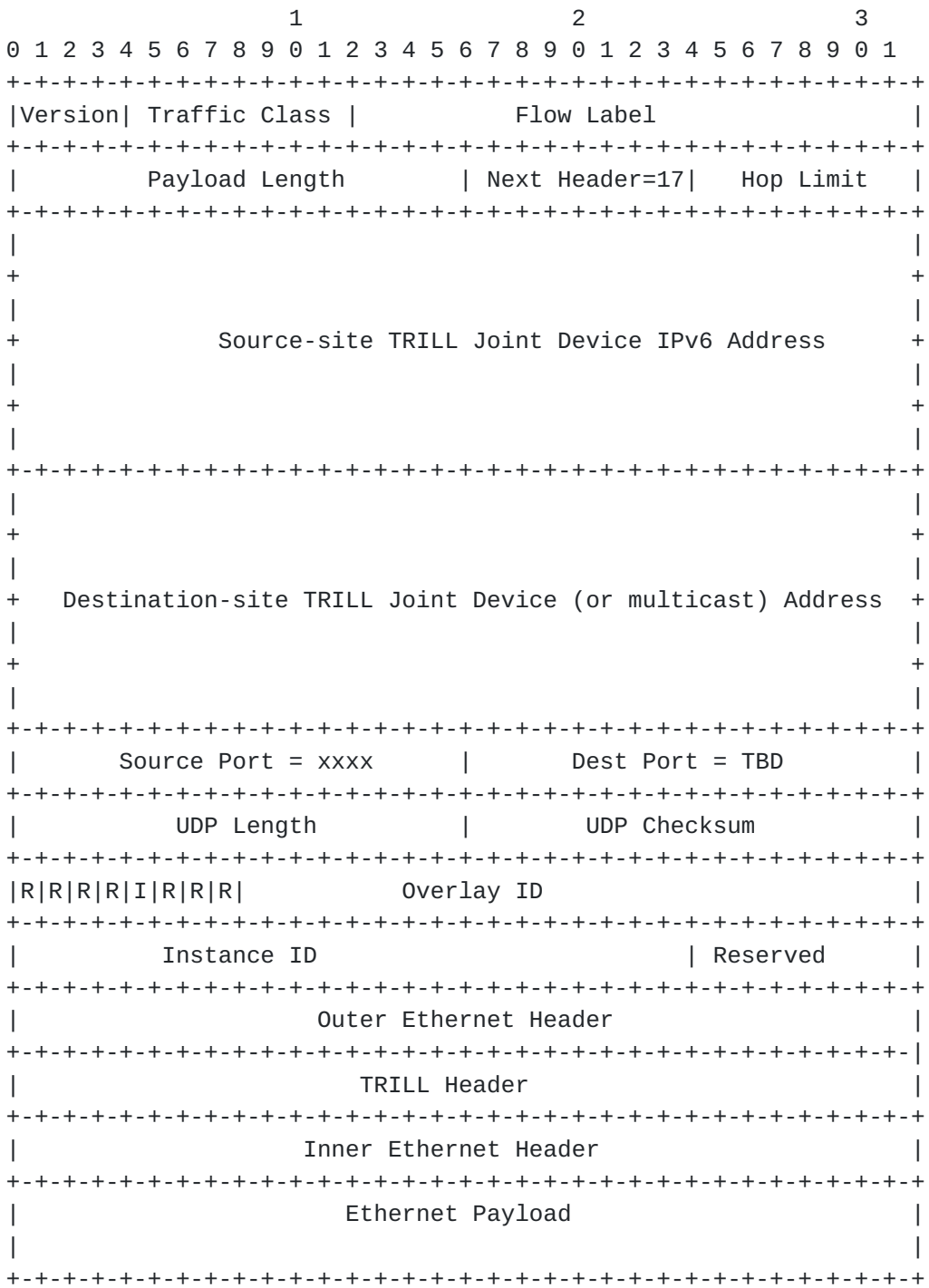
4.1. Encapsulation

The encapsulation format is TRILL frame encapsulated in UDP inside of IPv4 or IPv6.

The format of the UDP IPv4 encapsulation is as follows:



The format of the UDP IPv6 encapsulation is as follows:



4.2. Forwarding Process

4.2.1. Forwarding from an Internal Link to the Overlay

The forwarding within a site is normal TRILL forwarding, so here only describes the forwarding from an Internal link to the Overlay Link, or vice versa.

A Joint Device is a transit Rbridge from TRILL point of view. When a TRILL packet is received from the internal interface, egress Nickname is used to lookup the Nickname table which will yield a next-hop IP address entry pointing to a remote Joint Device. Then the packet is encapsulated with UDP/IP header and sent over the overlay interface to destination Joint Device at Layer-3 as a regular IP packet.

4.2.2. Forwarding from the Overlay Link to an Internal Link

When a packet is received on the overlay interface, it will be IP decapsulated to reveal the inner TRILL(including the outer MAC) header for forwarding. The egress Nickname will used for forwarding, the forwarding action is same as a transit RBridge.

4.2.3. Multicast Packet Flows

To be added.

4.2.4. Broadcast Packet Flows

To be added.

4.2.5. Mac Address Learning

The TRILL edge devices learn remote MAC addresses(including the MAC addresses in other data centers) in data plane by hardware. In most cases, the Joint device is like a transit RBridge, and doesn't learn end host's MAC addresses. From DCI(Data Center Interconnect) perspective, the Joint Device is DCI device at the same time, so TRILL over WAN can relieve the pressure of MAC addresses table capability in DCI device.

4.2.6. Multi-homing

In the situation of multi-homing shown as Figure 3, all the Joint Devices can be active by the nature of TRILL.

Figure 4 shows what the resulting forwarding tables would look like in the multi-homing example.

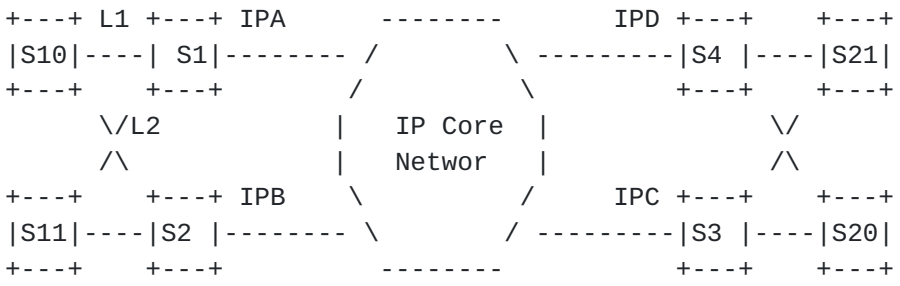


Figure 3: Multi-homing Scenario

S1

Nickname	Interface
S1	self
S2	-
S3	IPC
S4	IPD
S10	L1
S11	L2
S20	IPC/IPD
S21	IPC/IPD

Figure 4: Forwarding Table of S1

In S1 device, the traffic destinated to S10 and S21 have two next hops, IPC and IPD. In forwarding process, hashing of TRILL packet inner information will be used to determine which next hop IP address to use. Thus, the ingress traffic will be load balanced between multiple Joint Devices within a site.

5. IANA Considerations

IANA need to allocate the following UDP Ports for the TRILL IS-IS and Data channels:

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

6. Security Considerations

6.1. Failure Separation

To be added.

7. Acknowledgements

Much of the wording herein was adapted from draft-hasmit-otv-03.

8. Normative References

[I-D.hasmit-otv] Grover, H., Rao, D., Farinacci, D., and V. Moreno, "Overlay Transport Virtualization", draft-hasmit-otv-03 (work in progress), July 2011.

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

[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.

Authors' Addresses

Xiaolan Wan
HangZhou H3C Co. Limited
No. 2 ChuangYe Road, HaiDian District
Beijing
P.R. China

Phone: +86 10 82774971
Email: wxlan@h3c.com

Xiaopeng Yang
HangZhou H3C Co. Limited
No. 2 ChuangYe Road, HaiDian District
Beijing
P.R. China

Phone: +86 10 82774963
Email: yxp@h3c.com

Vishwas Manral
Hewlett-Packard Co.
19111 Pruneridge Ave.
Cupertino, CA
USA

Email: vishwas.manral@hp.com

Alvaro Retana
Hewlett-Packard Co.
2610 Wycliff Road
Raleigh, NC
USA

Email: alvaro.retana@hp.com