

TRILL Working Group
INTERNET-DRAFT
Intended Status: Proposed Standard

Sam Aldrin
Donald Eastlake
Huawei Technologies
Tissa Senevirathne
Ayan Banerjee
Santiago Alvarez
CISCO
Xiaolan Wan
XiaoPeng Yang
Hangzhou H3C Tech. Co. Ltd

Expires: January 10, 2013

July 9, 2012

TRILL Campus Extension
draft-aldrin-trill-campus-extension-00

Abstract

This document presents a solution suite for TRILL campuses to be connected over WAN networks. TRILL protocol is primarily designed to work within intra-data centers. Connecting different sites over WAN using overlay tunnel protocols is the primary method employed at present. Though this presents a simple mechanisms to extend TRILL campuses over WAN networks, it also brings in the problem of scalability for TRILL nicknames exchanged between sites, latency, duplication of traffic etc. This draft proposes a way to extend the TRILL sites without having the WAN network aware of TRILL campus network and its topology. This document details on how to extend TRILL campus sites and to establish connections between various sites without having to exchange entire TRILL campus information, thus reducing the information flow to the required sites only. This document do not add or define datacenter interconnect protocol.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

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

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2012 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 (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	4
1.1	Terminology	5
1.2	Contributors	5
2	Solution overview	5
2.1	Site inter-connection	6
2.2	Requirements overview	6
2.2	TRILL campus extension	7
2.3	TRILL nickname exhaustion	7
3	Solution analysis and comparison	7
3.1	TRILL campus extension	8
3.2	TRILL campus extension over IP	9
3.3	TRILL campus interconnection with TRILL-EVPN	10
3.4	TRILL campus interconnection over VPN's	11
4	Operational Overview	12
4.1	Campus and Backbone Areas	12

4.2	Unicast forwarding	12
4.3	Multicast Forwarding	13
4.4	MAC learning	14
4.5	TRILL nickname aggregation	15
5	TRILL campus inter-connectivity	15

5.1	Border RBridges interconnection	16
5.2	Inter-site nickname exchange	16
5.3	Border RBridge capability exchange	16
5.4	TRILL adjacency resolution	17
5.5	Forwarding of data frames	17
6	Proposed additions and extensions	17
6.1	Border RBridge capability TLV	18
6.2	TRILL nickname aggregation sub-TLV	18
7	TRILL campus extension over IP	19
7.1	Underlying Network	19
7.2	Overlay Control Plane	19
7.2.1	BRbridge discovery and adjacency setup	19
7.2.2	Control plane packet encapsulation	19
7.3	Overlay Data Plane	20
7.3.1	Encapsulation	20
7.4	Forwarding Process	22
7.4.1	Forwarding from an Internal Link to the Overlay	22
7.4.2	Forwarding from the Overlay Link to an Internal Link	22
7.4.5	Mac Address Learning	22
7.4.6	Multi-homing	23
7.5	Control plane termination	23
7.6	Control plane termination and Data plane de-capsulation	24
8	Security Considerations	24
9	IANA Considerations	24
10	References	24
10.1	Normative References	24
9.2	Informative References	25
	Authors' Addresses	25

1 Introduction

TRILL protocol is primarily designed as an intra-datacenter protocol by leveraging the routing functionality to interconnect bridges. Traditional Ethernet networks provided a single path for forwarding the traffic, which is usually derived using protocols like Spanning Tree. TRILL provided a way to utilize multiple links for forwarding, thus utilizing the resources effectively. Even though TRILL is new protocol, it seamlessly integrates with legacy bridging networks without having to forklift upgrade of all the bridges to support TRILL. By not having to learn the MAC addresses of end stations by intermediate devices, provided a powerful way to interconnect bridges within a datacenter and maximizing the resource usage and providing multipath usage option.

TRILL enabled network creates efficiency by having reduced forwarding table size. By doing TRILL nickname based forwarding created a layer of abstraction and much easier to implement the protocol. This enabled to address the scalability of a L2 domain, where thousands of Rbridges could exist to meet the needs of a datacenter. By leveraging IS-IS protocol, the information exchange and leveraging the path computation technology brought forth a new paradigm into bridging technology. TRILL Base Protocol Specification [[RFC6325](#)] specifies a tree based paradigm to forward broadcast and multicast traffic as well as unknown unicast traffic.

Even though the TRILL is enabled within a campus network or a datacenter, it is not primarily designed to work over WAN. There is a need to interconnect various TRILL campuses to extend beyond a single

LAN domain. Some enterprise or campus networks could be having multiple TRILL sites and these TRILL enabled sites could run independently or could share resources in order to cater to the needs of customers or users. As such, there exist few proposals based on overlay technologies which interconnect these sites but those solutions require MAC learning at the edge R Bridges and stripping of TRILL nickname on the frame. Another option is to interconnect these TRILL sites transparently over Pseudowires and making a huge TRILL campus site. This is useful option but the downside of this is when provider would like to maintain independent sites and exchange only the required data to be shared across sites, it becomes complicated to maintain the networks.

[TRILL-EVPN] draft details the data center protocol in interconnect different LAN domains, including TRILL sites, over VPN's. This requires establishing VPN's over WAN using BGP protocol, thus requiring WAN service provider involvement in establishing interconnection between different sites.

Aldrin, et.al.

Expires January 10, 2013

[Page 4]

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

This draft solves a specific problem where a TRILL campus needs to be extended, be it a different geographical location or network administrative domain. Though the primary goal is effective extension of TRILL campuses, this draft is not a replacement for datacenter interconnect protocols like PBB_EVPN. All the extensions are limited to TRILL protocol, without any extensions to WAN protocols for interconnect, such that, effective unicast and multicast traffic could be sent over WAN links and keeping core network agnostic of the TRILL network.

1.1 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](#)].

1.2 Contributors

In addition to the authors listed above the following individuals also contributed to this document:

Xiaopeng Yan

Hangzhou H3C Tech. Co., Ltd.

Vishwas Manral
Alvaro Retana
Dave Tucker
Hewlett-Packard

[2.](#) Solution overview

This section provides the high-level overview of the solution to the problems in various scenarios. More detailed representation of the solution is covered in the later sections of the draft.

TRILL site or TRILL campus uses IS-IS to setup the RBridge interconnection. A RBridge knows how to reach another RBridge within the campus. When two TRILL campuses are interconnected, one could visualize it in two different perspectives. First one is a merger of two TRILL campuses into one. This requires each RBridge to know about the other and the IS-IS should be able to compute the shortest paths from one to another. The downside of this model is that information exchange explosion and the size of IS-IS db and number of PDU's being exchanged could increase exponentially.

The second perspective is to have these two TRILL campuses being interconnected over a WAN, but their functioning nature is independent to each other. The two campuses exchange only the

Aldrin, et.al.

Expires January 10, 2013

[Page 5]

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

required information between border RBridges of the campuses. This will be more optimal and leads to interconnecting multiple campuses without having to redesign the whole network to ensure uniqueness and identity of RBridges.

Solution being proposed is in line with second option, which maintains site independency with a simplified solution and modeled around the existing and proven technologies. Some of the enhancements were already proposed in other drafts like multilevel TRILL [TISSA-MLEVEL] and the solution leverages by extending those definitions as necessary. The solution addresses the following areas described in the following sections

[2.1](#) Site inter-connection

Each TRILL campus is considered as an independent site or an L1 IS-IS domain. These TRILL campus sites are interconnected over WAN. Each area will have an appointed border R Bridges. These R Bridges exchange the information of other border R Bridges of different TRILL campus sites to establish connection with each other. When a TRILL campus is extended by interconnecting two or more campuses, the interconnection could be over L2 or L3 WAN connection. A frame from an R Bridge in campus, to reach other R Bridge in other campus, should be aware of the path it should take. When a frame is sent from the border R Bridge to another Border R Bridge in the other campus, it traverses the WAN network. This is transparent to the source R Bridge in order to reach the destination R Bridge. The WAN connection established between the campuses is not aware of the traffic it is carrying and is not established based on the TRILL campuses.

[2.2](#) Requirements overview

There are various requirements necessary to be met in order to provide a seamless TRILL campus extension. Some of the important requirements are as follows

- o Extend TRILL technology over interconnect
- o Ability to provide the same fast convergence for mobility as it does in intra-TRILL campus
- o Ability to work transparent with various WAN technologies
- o Option for dynamic establishment of connectivity across sites
- o Minimal changes to TRILL protocol and definitions

- o Backward compatible to existing networks and their functions.

In some scenarios it may be required for TRILL to be extended over an IP network, for example a case where the network administrator does not have control over the wide area network configuration. The TRILL over IP solution should meet the requirements outlined above and additionally:

- o Require only unicast reachability between R Bridges and multicast support from the WAN provider
- o Provide Auto-discovery of Border R bridges

[2.2](#) TRILL campus extension

By interconnecting TRILL campus sites over WAN, one could extend the L1 area, but that would cause other issues as detailed in the earlier section. When two campuses merge, there will be a possibility for nickname collision. The ideal situation is to keep the ability for each campus nicknames overlapped with the other campus. The other option is to provide a mechanism to do the nickname resolution. This version of the document details the solution of the later option. In the subsequent version, the former option for retaining nicknames will be addressed.

[2.3](#) TRILL nickname exhaustion

Though this draft is not meant to provide solution for TRILL nickname exhaustion, it enables provider to deal with the problem effectively and not having to re-design the network, every time a new campus is interconnected. The proposed solution has R Bridges which are not required to be exposed outside of the campus and there are other R Bridges which are also known as border R Bridges. This version the document assumes global uniqueness of TRILL nicknames across various campuses. When a frame has to be forwarded to an R Bridge which resides in another campus, the originating R Bridge knows how to get to the border R Bridge. This border R Bridge should have the list of R Bridges of other campus sites and thus could select the appropriate link connecting to the destination campus and encapsulate the TRILL frame and forward over that. More details are covered in the unicast and multicast sections of the detailed solution.

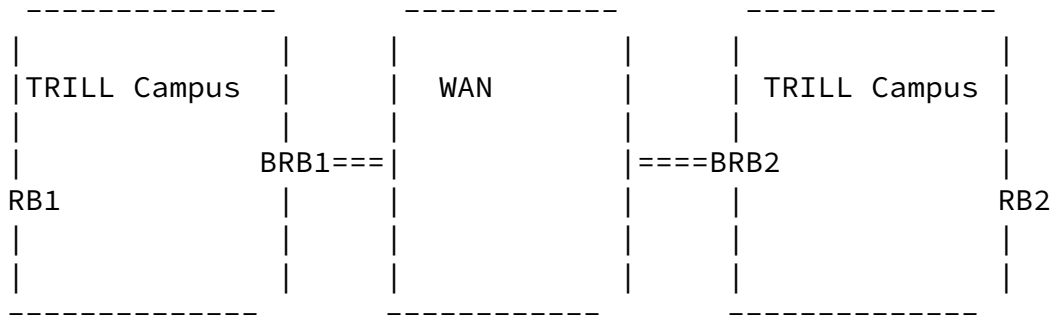
[3.](#) Solution analysis and comparison

As eluded to in the earlier sections, there are various methods on interconnecting different TRILL campus sites. Before going into the details of proposed solution a close examination of some of the

proposed solutions, provides better perspective of this solution.

3.1 TRILL campus extension

In this model TRILL campuses are connected over WAN using technologies like PW. This is the most simple way of interconnecting the sites. When campuses are interconnected, the TRILL campus will get expanded and each RBridge could reach each other. The main criteria for this will be to maintain unique nickname for RBridges.

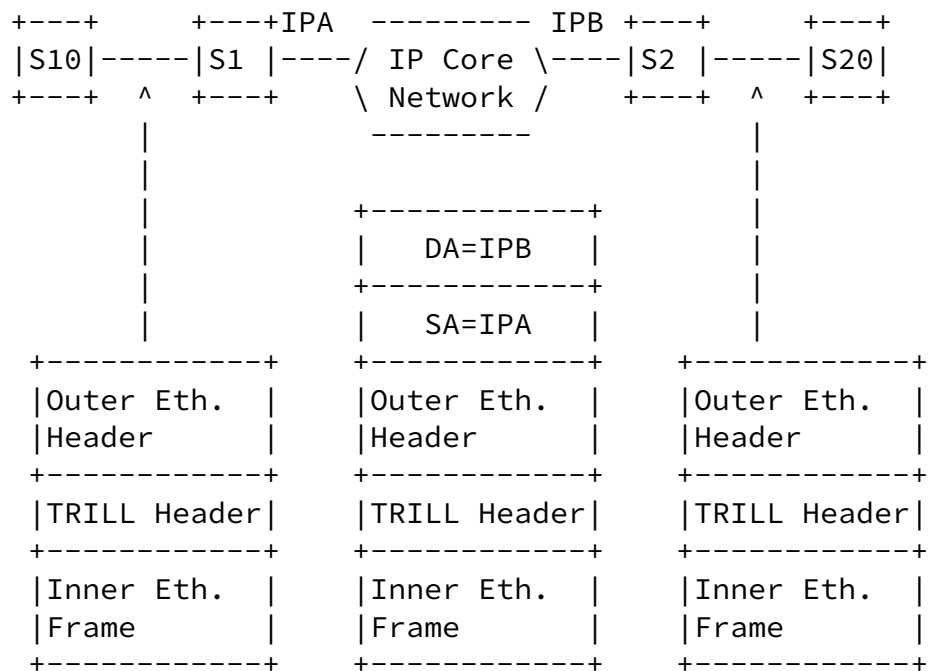


As shown in the figure above, two TRILL campuses are interconnected over WAN. Border RBridges establish connection over WAN using PW or other WAN technologies. All the nicknames within each campus sites have to be unique. The WAN in this case is transparent to the TRILL campuses and the path computation doesn't involve WAN component, instead it will be like one TRILL campus. When RB1 originates a TRILL frame destined to RB2, it traverses over BRB1 and BRB2 and reaches RB2.

This solution is workable when the campuses are small and do NOT need to change or requires interconnecting more TRILL campuses. The other downside for this model is, when two campuses are interconnected and there is overlap of nicknames, the network has to be re-designed to eliminate the duplicate nicknames and make each RBridge to have a unique nickname.

[3.2](#) TRILL campus extension over IP

The TRILL campus could be extended by encapsulating TRILL inside an IP Tunnel. By using this method the WAN becomes completely transparent and is only required to provide unicast reachability between BRbridges and establish the multicast trees.



INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

Each BRbridge maintains a mapping table between the egress nickname and an IP address. When a BRbridge receives a frame destined for a remote BRBridge it looks up the egress nickname in the mapping table and applies an outer IP header where the destination address is equal to the Remote BRBridge IP.

```

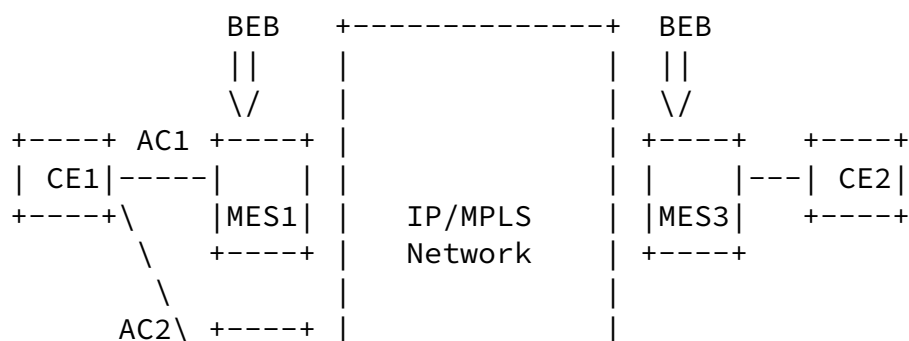
+---+ L11+---+IPA  ----- IPB +---+ L21+---+
|S10|----|S1 |----/ IP Core \----| S2|----|S20|
+---+      +---+      \ Network /      +---+      +---+
                        -----

```

S1			S2		
Nickname	Interface		Nickname	Interface	
S10	L11		S10	IPA	
S20	IPB		S20	L21	
S1	self		S1	IPA	
S2	IPB		S2	self	

[3.3](#) TRILL campus interconnection with TRILL-EVPN

TRILL campuses could be extended over WAN using TRILL-EVPN.



These VPN's could be established statically or dynamically. In order to establish dynamically, the border R Bridges need to exchange information of the nicknames and connect different sites. The hierarchical model like H-VPLS could be established as well. One other option which is much similar to the first model where campuses exchange TRILL nicknames with other campuses over VPN's. Though this model groups different sites according to the way VPN's are configured, avoiding flooding of TRILL nicknames or site independency cannot be achieved.

[4. Operational Overview](#)

[4.1 Campus and Backbone Areas](#)

Each TRILL campus will be assigned a border R Bridge. This is identified using the 'Attached' bit in the IS-IS PDU. The border R Bridge has list of the R Bridges of each campus site. These list of bridges are exchanged using the TRILL nickname aggregation sub-TLV. Details of the sub-TLV are detailed in the below section.

Every TRILL campus need not exchange all the R Bridge nicknames with other campuses. Let us take the scenario of campus A to be interconnected with campus B. In campus A, there are R Bridges RB1...RB10, which are interconnected in L1. These nicknames are not tunneled or exchanged with other L1 campus sites. Similarly campus B has RB11...RB20 and need not be distinct from campus A R Bridge nicknames. So, if a new campus is connected to the domain, there is a conflict of nicknames. As described in the earlier section, when nicknames have to be uniquely maintained, the draft describes a solution. Solution with keeping same nickname across different TRILL campuses will be addressed in the later versions of the draft.

When campuses are interconnected over WAN links, there are two possible terminations of the WAN links, the border R Bridge and an edge PE device. If the R Bridge is connected to PE device, the TRILL

frames could be sent over the link connecting to the PE device to be transported across WAN. This process is transparent to the TRILL network and the RBridge doesn't remove the TRILL encapsulation, rather tunnel the frames over the WAN to the far end RBridge.

There are many ways the WAN connections could be provided from RBridge. GRE tunnels, IP tunnels, MPLS LSP's etc. Details of these is outside the scope of this draft as it is transparent to RBridge.

The border RBridges will have the complete information of its campus RBridges. Not all of the RBridges nicknames need to be advertised globally. So, the globally exchanged nicknames of RBridges should be unique across campuses. Depending on the policy established, these Border RBridges will exchange the TRILL information with other campus border Bridges, between different campuses. In IS-IS domain the equivalent of this is the L2 backbone area, which in this case, is established over WAN.

[4.2](#) Unicast forwarding

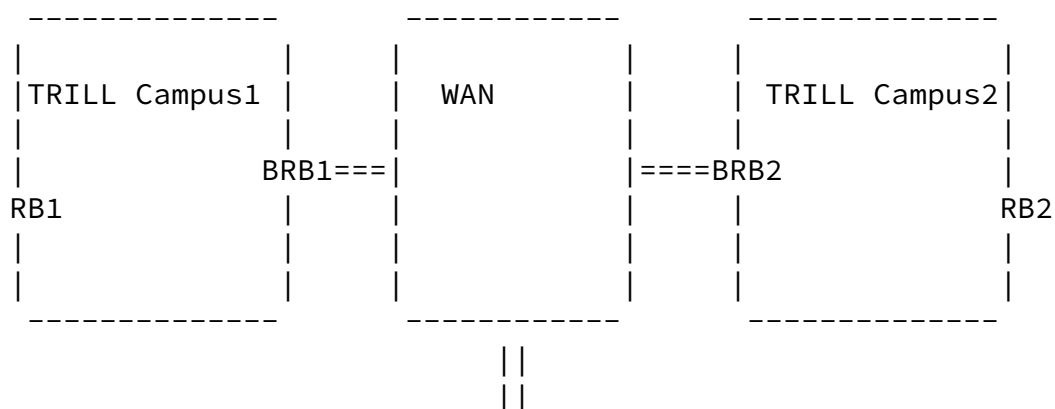
If the destination TRILL nickname is not known, the originating or transit RBridges forwards it to border RBridge. As the border RBridge has all the nicknames of each campus, it forwards the frame to the

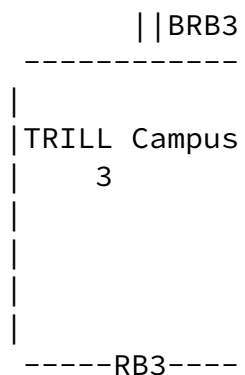
right campus border RBridge, which in turn could forward within its campus as per base protocol specification [[RFC6325](#)]. In the case where the destination is unknown, the frame is flooded to each campus. Using the MAC learning procedures, the associated RBridge will be learnt for the subsequent frames to be forwarded as unicast, instead of flooding. Flooding into various campuses or TRILL data sites happen only if the the frame is of global ID based on VLAN identification.

[4.3](#) Multicast Forwarding

Whether the traffic scope is local or global across campuses is identified by VLAN or port or fine-grain label. If the traffic is to be forwarded within campus, it is done using the local tree. But if the forwarding has to be done globally, it needs to use the global tree. When the frame has global context, it will be flooded into other TRILL sites as well.

Multicast networks are established within the core networks. In the case of RBridges, which are part of the customer networks and do not participate in the core networks and their topologies, the multicast tree could be built using IP multicast or leverage MVPN services offered by the core network service provider. The global trees are established between border RBridges with the help of information exchanges between border RBridges. As the IS-IS is limited to individual campus sites, the information for the backbone tree over WAN has to be exchanged between border RBridges either as a IP multicast PIM message or specific TLV indented for other campus border RBridges. More details on this will be added in the later versions of the draft.





In the above figure when the multicast frame has to be sent from campus 1 to campus 2 and 3, the frame arrives at border RBridge BRB1. With the default global tree between border RBridges of different campuses, the forwarding is setup to egress the frame or replicated over multicast trees to all other campus sites. If the frame is destined for non-default global tree, the frame is forwarded according to the forwarding information established for the tree.

Once the frame is reached on the border router of the campuses, the frame is locally multicast forwarded. The same technique as employed in the multilevel draft [[TISSA-MLEVEL](#)] is used here as well.

If mVPN services are deployed interconnecting campus sites, the multicast tree is built over these services based on the customer VLANs.

[4.4](#) MAC learning

When a frame is to be forwarded from customer MAC A to customer MAC B, the frame is set as unknown unicast frame over TRILL networks. If the MAC A and MAC B are connected over WAN, the frame is transmitted over WAN to the other campus. When the frame is reached at the RBridge connecting to MAC B, it will learn about the originator RBridge for MAC A. While responding, the egress RBridge know the originating RBridge, itwill unicast the frame to the originator.

[4.5](#) TRILL nickname aggregation

Nicknames are allocated or assigned to RBridges in a given campuses using various methods. It could be OSS, CLI or could be a dynamic

control protocol which configures the nicknames. As the nicknames are confined to each L1 area, the nickname management, when sites are connected over WAN, it is essential to optimize the name allocation in order to use the name space effectively. Name allocation is not in the scope of this draft. If there is a necessity, the topic could be considered in the later revisions of the draft. For this version we do recommend some of the optimization techniques for nickname allocation defined in the multilevel draft [[TISSA-MLEVEL](#)].

Each border RBridges needs to exchange the nicknames of each campuses with other border RBridges. As the border RBridges are connected over various types of WAN networks, mandating enhancement to a specific protocol is deemed not the right approach. As the information exchange has to be done, certain characteristics for the data exchange have to be met.

- o The amount of data exchange has to be minimal and optimized.
- o the information exchange has to be quick.
- o Conflicts and duplicate information flow has to be avoided

This draft proposes a generic TLV, which is to be exchanged between border RBridges. If the nickname allocation is done in terms of ranges, the same could be exchanged between border RBridges, seamlessly. As the TLV has to be terminated at the border RBridges, it is better suited to be sent as a unicast message to the neighboring border RBridge. This could be sent as an IP message with TRILL header containing the target border RBridge. More details on how to encapsulate and process the frame should be in the later versions of the draft.

[5.](#) TRILL campus inter-connectivity

The primary reason to interconnect TRILL campuses is to maintain geographically distant, segmented sites and customer specific segregation possible by interconnecting and not having to redo the network and campus redesign for every change and need. With customers being mobile or services offered by service providers could be re-located depending on the time-zones and resource availability, restricting to a specific site is a thing of the past.

These constraints brought forth the need to have different sites interconnected over the WAN, be it MPLS or VPLS or IP and to provide

the services on demand to meet the needs of customers and their data center needs. As TRILL has proven to be an effective protocol by bringing routing technologies into bridges or L2 forwarding, the short coming of TRILL interconnect is to be overcome. As eluded to in the earlier sections on different kinds of solutions, meeting all the needs of the TRILL campus extension as laid out in the requirements section, is the primary goal of this draft.

[5.1](#) Border RBridges interconnection

Border RBridges only participate in interconnecting various TRILL campuses. These border RBridges are elected or identified as described in the earlier section i.e. using IS-IS protocol advertisement. These border RBridges, when required to interconnect with other campuses, over WAN, depending on configuration of choice, an overlay tunnels like GRE could be established between the border RBridges. advertise the route to other site border RBridges using the BGP enhancement. The connectivity between different campuses over WAN is already established. When two RBridges needs to be connected over, a GRE or IP tunnel is established between those two. Detailed mechanisms on this establishments will be addressed in the later versions of the draft.

If L2 connectivity is to be used with protocols like VPLS, a similar method could be employed, where the PWE3 could be established between border RBridges. More details to be added in the later versions of the draft.

[5.2](#) Inter-site nickname exchange

There are three types of nicknames which are exchanged between border RBridges.

- o Nickname of border RBridges
- o Nicknames of RBridges for each campus
- o Nicknames of RBridges which are part of a specific customer VLAN or VPN

The nickname aggregation TLV is used as payload to be exchanged between border RBridges. This information is used to establish inter-connectivity between TRILL sites per customer VLAN or default global tree. There exchange of information could be done with existing protocols and is not restricted to any specific protocol.

[5.3](#) Border RBridge capability exchange

An additional capability TLV is defined to exchange info on what each of the border RBridge is capable of. This is very essential for forward and backward capability . Capability information not only indicates the capability version but could also force the interconnection to be restricted as per the policy set by the customer. Some of the capability advertisements are as follows.

- o Version.
- o default nick name resolution
- o connect more campuses
- o active-active link support
- o Ability to support multicast forwarding

[5.4](#) TRILL adjacency resolution

When a frame is to be forwarded from one campus to another, the adjacency resolution has to be done on the border RBridge. When TRILL nicknames are advertised from one border RBridge to another, the border RBridge keeps the database of all the nicknames.

Once the frame is received on the border RBridge, it will look in the forwarding table to identify the next hop. The adjacency information could indicate the outgoing WAN encapsulation. For VPN, there could be additional encapsulation depending on the network configuration. The TRILL frame is encapsulated, without removing the TRILL header and is forwarded over the WAN link.

[5.5](#) Forwarding of data frames

The TRILL frames are forwarded as per the base protocol [[RFC6325](#)] within a campus site. The forwarding of the frames from TRILL campus to campus over WAN connection is pre-established between border R Bridges. The encapsulation of the Frame with WAN header is based on the adjacency resolution made in the forwarding on the border RBridge.

[6.](#) Proposed additions and extensions

There are certain extensions being proposed in this draft to interconnect TRILL campuses or datacenters. These include new additions to routing and also new TLV's to exchange information between border RBridges. There are few references to the extensions proposed in other drafts which are used in this draft as well.

Aldrin, et.al.

Expires January 10, 2013

[Page 17]

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

[6.1](#) Border RBridge capability TLV

This TLV as defined in earlier section, defines the capability of a border RBridge, to be exchanged with other border RBridges for seamless inter-working across campus sites.

```

  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
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = <TBD>          |          Length = 8          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Version          |          Flags          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Flags                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
```

Definition of flag bits will be identified and defined later.

[6.2](#) TRILL nickname aggregation sub-TLV

The nickname aggregation TLV defined in multilevel draft [TISSA-MLEVEL] is used in advertising the nicknames into other border Routers. Some new additions or changes will be proposed in later versions of the draft.

[7.](#) TRILL campus extension over IP

[7.1](#) Underlying Network

The underlying network in the TRILL campus extension over IP solution could be Wide Area Network. As changes to this network are not in the administrative control of the TRILL campus, TRILL over IP extends the campus network by having unicast reach ability between BRbridges and the ability to establish IP multicast trees over it. There are three ways the TRILL campus could be extended over IP.

1. TRILL frames, both control and data frames are transparently tunneled over IP. This is also known as overlay model.
2. TRILL and IS-IS control frames are terminated at the campus edge and data frames are tunneled transparently.
3. TRILL and IS-IS control frames are terminated at the campus edge and data frames are stripped off the TRILL header

[7.2](#) Overlay Control Plane The TRILL over IP overlay control plane is responsible for the auto-discovery of the BRbridges within the same domain. The TRILL control plane traffic between BRbridges will be carried over the virtual link. There is no termination of either TRILL control plane or data plane frames at the edge of the TRILL

campus networks.

[7.2.1](#) BRbridge discovery and adjacency setup BRbridges become part of the domain when they join the multicast group on the underlying network associated with the domain. The auto-discovery mechanism happens when BRbridges peer with each other as if they were directly connected at layer-2. This peering is possible as all the traffic for the BRbridge is encapsulated with the underlying network multicast group address and sent into the core.

[7.2.2](#) Control plane packet encapsulation

Any BRbridge in a TRILL over IP domain should encapsulate the ISIS routing information of its campus and then relay it to the other BRbridges in the domain. The control frames are tunneled through the IP connection established between BRbridges.

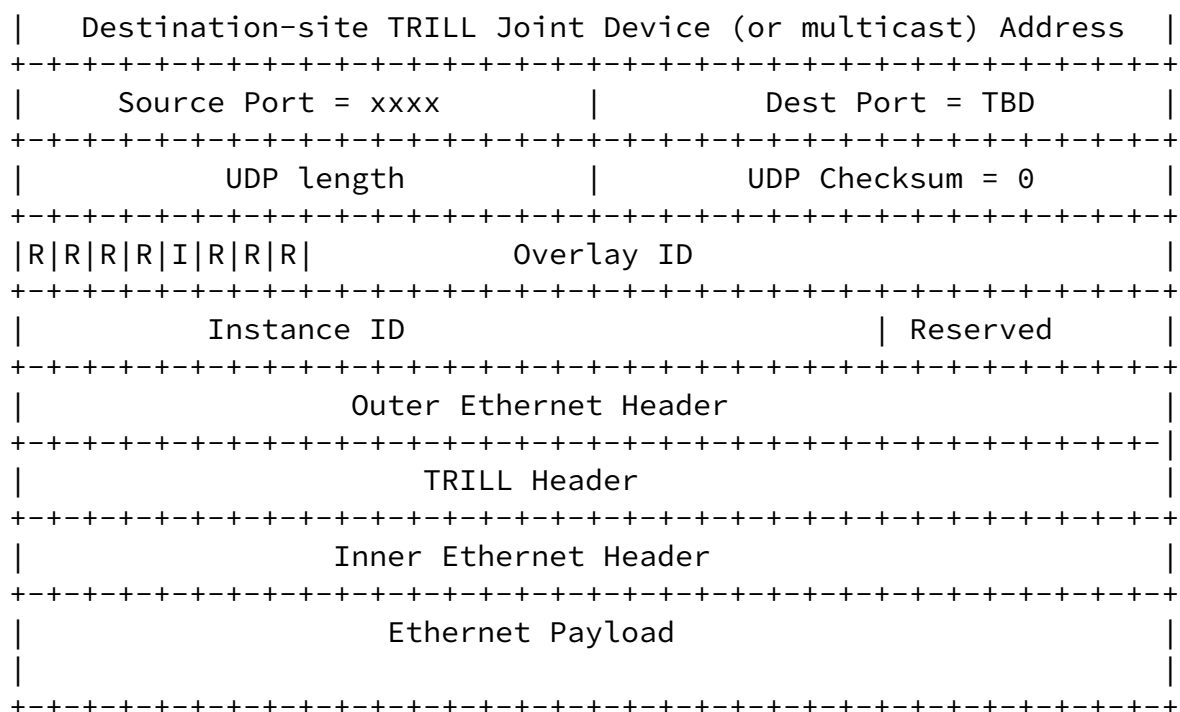
[7.3](#) Overlay Data Plane

[7.3.1](#) Encapsulation

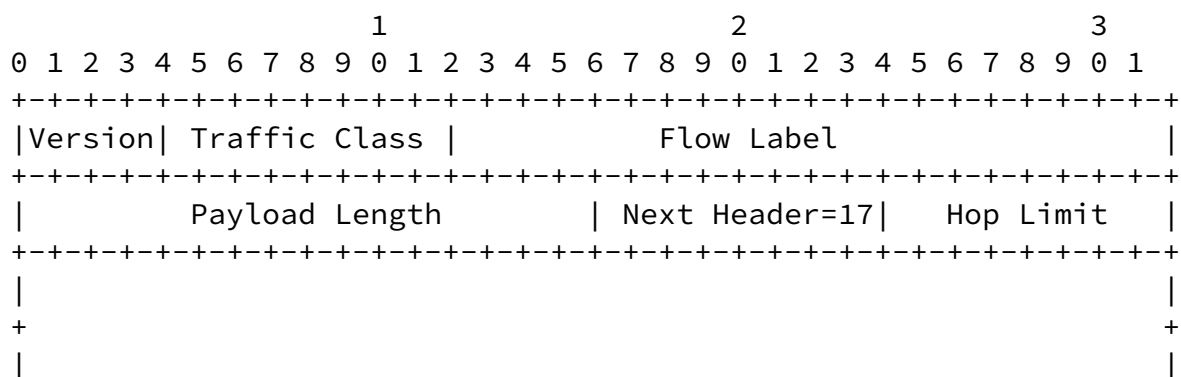
The encapsulation format is TRILL frame encapsulated in UDP inside of IPv4 or IPv6.

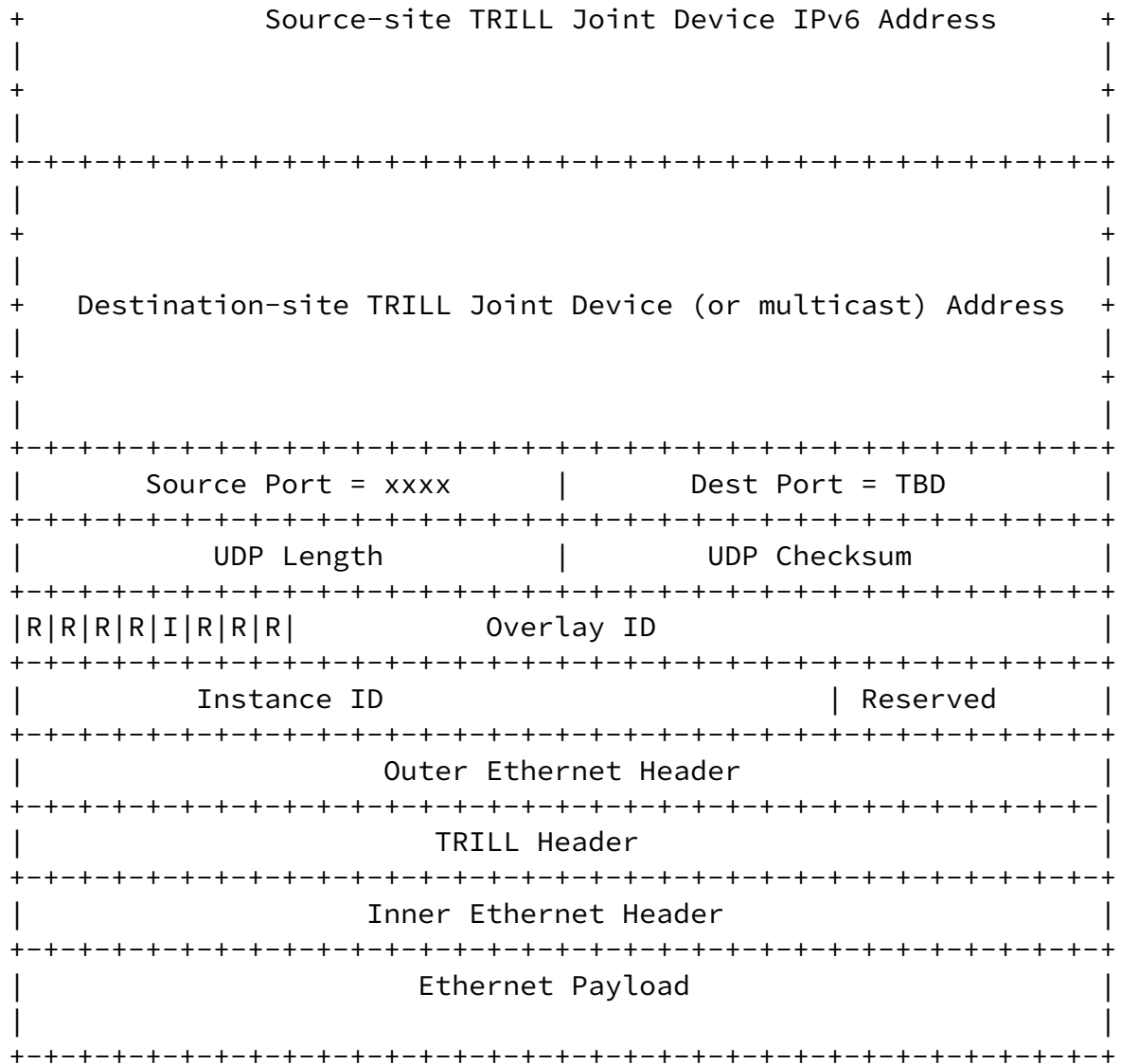
The format of the UDP IPv4 encapsulation is as follows:

```
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
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|Version|  IHL  |Type of Service|                Total Length        |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|                Identification                |Flags|  Fragment Offset  |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| Time to Live | Protocol = 17 |                Header Checksum        |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|                Source-site TRILL Joint Device IP Address            |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```



The format of the UDP IPv6 encapsulation is as follows:





[7.4](#) Forwarding Process

[7.4.1](#). Forwarding from an Internal Link to the Overlay

The forwarding within a campus is as defined in the base protocol of

TRILL. This section here 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.

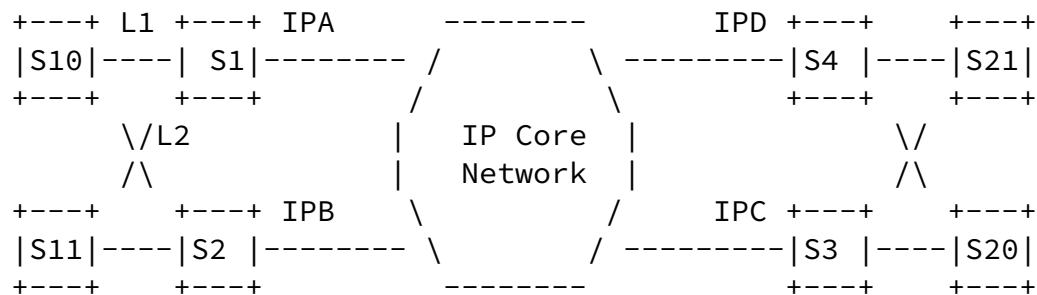
[7.4.2.](#) Forwarding from the Overlay Link to an Internal Link When a packet is received on the overlay interface, it will be IP de-capsulated to reveal the inner TRILL(including the outer MAC) header for forwarding. The egress Nickname will used for forwarding, the forwarding action is same as a transit RBridge.

[7.4.5.](#) 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 campus extension perspective, the border device is DCI device at the same time, so TRILL over WAN can relieve the pressure of MAC addresses table capability in DCI device.

7.4.6. Multi-homing

In the situation of multi-homing shown as Figure 3, all the BRbridges can be active by the nature of TRILL. Figure 4 shows what the resulting forwarding tables would look like in the multi-homing example.



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

In S1 device, the traffic destined 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.

7.5 Control plane termination

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

When TRILL campuses are extended by interconnecting them, the administrative domain for control plane could be independent, where as the data plane could be transparently interconnected. This requires the campus border RBridges to terminate the control plane. The IS-IS control frames do not get encapsulated and sent over to the the IP connection. Border RBridges exchange the campus information over IP in order to program the forwarding table and adjacency resolution.

Data frames when arrived at the border RBridges, destined for the other campus, will be looked for adjacency and encapsulated with IP and sent over the IP. In the case of multicast, the TRILL frame is encapsulated and sent over IP multicast network.

[7.6](#) Control plane termination and Data plane de-capsulation

When two campuses are interconnected but no control plane or data plane is extended over the WAN, the TRILL frames are stripped off the TRILL header at the border RBridge of the campus. The payload is then encapsulated as a regular IP packet and sent over the IP network to the other campus. This requires the border RBridges to learn the host MAC address to properly encapsulate as IP packet and routed across the WAN link.

[8](#) Security Considerations

<TBD>

[9](#) IANA Considerations

It is requested that the IANA 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

[10](#) References

10.1 Normative References

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

Aldrin, et.al.

Expires January 10, 2013

[Page 24]

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

[RFC2234] Crocker, D., Ed., and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", [RFC 2234](#), November 1997.

[RFC4971] Vasseur, JP., Ed., Shen, N., Ed., and R. Aggarwal, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", [RFC 4971](#), July 2007.

[RFC6325] Perlman, R., et.al, "Routing Bridges (RBridges): Base Protocol Specification", [RFC 6325](#), July 2011.

[trillcmt] Senevirathne, T., et.al, "Coordinated Multicast Trees (CMT) for TRILL", Work in Progress, January 2012.

9.2 Informative References

[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.

[RFC6326] Eastlake, D, et.al, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", [RFC 6326](#), July 2011.

[rfc6326bis] Eastlake, D, et.al, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", Work in Progress, [draft-eastlake-isis-rfc6326bis-04.txt](#), January 2012.

[TISSA-MLEVEL] Senevirathne, "RBridges: Multilevel TRILL", Work in Progress, [draft-tissa-trill-multilevel-00.txt](#), February 2012.

[PBB-EVPN] Sajassi, et.al, "PBB-EVPN", Work in Progress, [draft-ietf-l2vpn-pbb-evpn-03](#), June 2012.

[TRILL-EVPN] Sajassi, et.al, "TRILL-EVPN", Work in Progress, [draft-](#)

Authors' Addresses

Sam Aldrin
Huawei Technologies
2330 Central Express Way
Santa Clara, CA 95951

Email: aldrin.ietf@gmail.com

Tissa Senevirathne

Aldrin, et.al.

Expires January 10, 2013

[Page 25]

INTERNET DRAFT

TRILL Campus Extension

July 9, 2012

CISCO Systems
375 East Tasman Drive
San Jose CA 95134

Phone: 408-853-2291
Email: tsenevir@cisco.com

Ayan Banerjee
CISCO Systems
425 East Tasman Drive
San Jose CA 95134

Phone: 408-527-0539
Email: ayabaner@cisco.com

Donald Eastlake
Huawei Technologies
155 Beaver Street
Milford, MA 01757 USA

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

Santiago Alvarez
CISCO systems

Email: saalvare@cisco.com

Xiaolan Wan
HangZhou H3C Tech. 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