Mohammed Umair
Cisco
Kingston Smiler Selvaraj
IPInfusion
Donald Eastlake 3rd
Huawei Technologies
Lucy Yong
Self
Expires: June 7, 2018
December 4, 2017

TRILL Transparent Transport over MPLS
draft-ietf-trill-transport-over-mpls-06.txt

Abstract

This document specifies how to interconnect multiple Transparent Interconnection of Lots of links (TRILL) sites with an intervening MPLS network using existing TRILL and VPLS standards. This draft addresses two problems as follows:

1) Providing connection between more than two TRILL sites that are separated by an MPLS provider network.

2) Providing a single logical virtualized TRILL network for different tenants that are separated by an MPLS provider network.

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

Table of Contents

1. Introduction

The IETF Transparent Interconnection of Lots of Links (TRILL) protocol [RFC6325] [RFC7177] [RFC7780] provides transparent forwarding in multi-hop networks with arbitrary topology and link technologies using a header with a hop count and link-state routing. TRILL provides optimal pair-wise forwarding without configuration, safe forwarding even during periods of temporary loops, and support for multipathing of both unicast and multicast traffic. Intermediate Systems (ISs) implementing TRILL are called Routing Bridges (RBridges) or TRILL Switches

This draft, in conjunction with [RFC7173] on TRILL Transport using Pseudowires, addresses two problems:

1) Providing connection between more than two TRILL sites belongs to a single TRILL network that are separated by an MPLS provider network using [RFC7173]. (Herein also called problem statement 1.)

2) Providing a single logical virtualized TRILL network for different tenants that are separated by an MPLS provider network. In short providing connection between TRILL sites belonging to a tenant/tenants over a MPLS provider network. (Herein also called problem statement 2.)

A tenant is the administrative entity on whose behalf their associated services are managed. Here tenant refers to a TRILL campus that is segregated from other tenants for security reasons.

A key multi-tenancy requirement is traffic isolation so that one tenant's traffic is not visible to any other tenant. This draft also addresses the problem of multi-tenancy by isolating one tenant's traffic from the other.

1.1. Terminology

Acronyms used in this document include the following:

AC - Attachment Circuit [RFC4664]

Data Label - VLAN or FGL

ECMP - Equal Cost Multi Path

FGL - Fine-Grained Labeling [RFC7172]

IS-IS - Intermediate System to Intermediate

System [IS-IS]

LDP - Label Distribution Protocol

LAN - Local Area Network

MPLS - Multi-Protocol Label Switching

PBB - Provider Backbone Bridging

PE - Provider Edge Device

PSN - Packet Switched Network

PW - Pseudowire [RFC4664]

TIR - TRILL Intermediate Router (Devices that has both IP/MPLS and TRILL functionality)

TRILL - Transparent Interconnection of Lots of Links OR Tunneled Routing in the Link Layer

TRILL Site - A part of a TRILL campus that contains at least one RBridge.

VLAN - Virtual Local Area Network...........

VPLS - Virtual Private LAN Service

VPTS - Virtual Private TRILL Service

VSI - Virtual Service Instance [RFC4664]

VTSD - Virtual TRILL Switch Domain OR Virtual TRILL Service Domain A Virtual RBridge that segregates one tenant's TRILL database as well as traffic from the other.

WAN - Wide Area Network

2. TRILL Over MPLS Model

TRILL Over MPLS can be achieved in two different ways.

a) VPLS Model for TRILL
b) VPTS Model/TIR Model

Both these models can be used to solve the problem statement 1 and 2.

Herein the VPLS Model for TRILL is also called Model 1 and the VPTS Model/TIR Model is also called Model 2.

3. VPLS Model

Figure 1 shows the topological model of TRILL over MPLS using VPLS model. The PE routers in the below topology model should support all the functional Components mentioned in [RFC4664].

```
+------+                                                      +------+
| RBa +---+      ...........................      +---| RBb |
+------+   |      .                         .      |   +------+
Site 1     |    +-----+                 +-----+    |    Site 2
           +----|PE1 |                 |PE2 |----+
                +-----+    MPLS Cloud   +-----+
                  .                         .
                  .         +-----+         .
                  ..........|PE3 |...........
                            +-----+      ^
                               |         |
                               |         +-- Emulated LAN
                            +------+
                            | RBc |
                            +------+
                            Site 3
```

Figure 1: Topological Model of TRILL over MPLS connecting three TRILL Sites

Figure 2 below shows the topological model of TRILL over MPLS to connect multiple TRILL sites belonging to a tenant. (Tenant here is a TRILL campus, not a specific Data label.) VSI1 and VSI2 are two Virtual Service Instances that segregate Tenant1's traffic from other tenant traffic. VSI1 will maintain its own database for Tenant1, similarly VSI2 will maintain its own database for Tenant2.

```
       +-----+         ............................          +-----+
       |RBat1+---+     . ++++++++++++++++++++++++ .      +---|RBbt1|
       +-----+   |     . +                      + .      |   +-----+
 Tenant1 Site 1  |    +----+                  +----+     | Tenant1 Site2
                 +----|VSI1|                  |VSI1|----+
                 +----|VSI2|    MPLS  Cloud   |VSI2|----+
                 |    +----+                  +----+     |
       +-----+   |     . +                       + .     |   +-----+
       |RBat2+---+     . +++++++++ +-----+ ++++++++ .     +---|RBbt2|
       +-----+         ............|VSI1|...........         +-----+
 Tenant2 Site 2                    |VSI2|          ^      Tenant2 Site2
                                   +-----+         |
                                      |            |
                                  +------+         +-----Emulated
                                  |RBct2|                LAN
                                  +------+
                               Tenant2 Site 3

      .... VSI1 Path
      ++++ VSI2 Path
```

Figure 2: Topological Model for VPLS Model connecting 2 Tenants with 3 sites each

In this model, TRILL sites are connected to VPLS-capable PE devices that provide a logical interconnect, such that TRILL RBridges belonging to a specific tenant connected via an single bridged Ethernet. These devices are the same as PE devices specified in [RFC4026]. The Attachment Circuit ports of PE Routers are layer 2 switch ports that are connected to the RBridges at a TRILL site. Here each VPLS instance looks like an emulated LAN. This model is similar to connecting different RBridges (TRILL sites) by a layer 2 bridge domain (multi access link) as specified in [RFC6325]. This model doesn't requires any changes in PE routers to carry TRILL packets, as TRILL packets will be transferred transparently.

3.1. Entities in the VPLS Model

The PE (VPLS-PE) and CE devices are defined in [RFC4026].

The Generic L2VPN Transport Functional Components like Attachment Circuits, Pseudowires, VSI etc. are defined in [RFC4664].

The RB (RBridge) and TRILL Sites are defined in [RFC6325] as updated by [RFC7780].

3.3. TRILL Adjacency for VPLS model

As specified in section 3 of this document, the MPLS cloud looks like an emulated LAN (also called multi-access link or broadcast link). This results in RBridges at different sites looking like they are connected by a multi-access link. With such interconnection, the TRILL adjacency over the link are automatically discovered and established through TRILL IS-IS control messages [RFC7177]. These IS-IS control messages are transparently forwarded by the VPLS domain, after doing MPLS encapsulation specified in the section 3.4.

3.4. MPLS encapsulation for VPLS model

Use of VPLS [RFC4762] [RFC4761] to interconnect TRILL sites requires no changes to a VPLS implementation, in particular the use of Ethernet pseudowires between VPLS PEs. A VPLS PE receives normal Ethernet frames from an RBridge (i.e., CE) and is not aware that the CE is an RBridge device. As an example, an MPLS-encapsulated TRILL packet within the MPLS network can use the format illustrated in Appendix A of [RFC7173] for the non-PBB case. For the PBB case, additional header fields illustrated in [RFC7041] can be added by entry PE and removed by the exit PE.

3.5. Loop Free provider PSN/MPLS

No explicit handling is required to avoid loop free topology. Split Horizon technique specified in [RFC4664] will take care of avoiding loops in the provider PSN network.

3.6. Frame Processing

The PE devices transparently process the TRILL control and data frames. Procedures to forward the frames are defined in [RFC4664].

4. VPTS Model

The VPTS (Virtual Private TRILL Service) is a L2 TRILL service, that emulates TRILL service across a Wide Area Network (WAN). VPTS is similar to what VPLS does for bridge core but provides a TRILL core. VPLS provides "Virtual Private LAN Service" for different customers. VPTS provides "Virtual Private TRILL Service" for different TRILL

tenants.

Figure 3 shows the topological model of TRILL over MPLS using VPTS. In this model the PE routers are replaced with TIR (TRILL Intermediate Router) and VSI is replaced with VTSD (Virtual TRILL Switch Domain). The TIR devices must be capable of supporting both MPLS and TRILL as specified in section 4.1.1. The TIR devices are interconnected via PWs and appear as a unified emulated TRILL campus with each VTSD inside a TIR equivalent to a RBridge.

Some of the reasons for interconnecting TRILL Sites without isolating the TRILL Control plane of one TRILL site from other sites are as described below.

1) Nickname Uniqueness: One of the basic requirements of TRILL is that, RBridge Nicknames are unique within the campus [RFC6325]. If we segregate control plane of one TRILL site from other TRILL site and provide interconnection between these sites, it may result in Nickname collision.

2) Distribution Tree and its pruning: When a TRILL Data packet traverses a Distribution Tree, it will stay on it even in other TRILL site. If no end-station service is enabled for a particular Data Label in a TRILL site, the Distribution Tree may be pruned and TRILL data packets of that particular Data Label might never get to other TRILL sites. The TRILL RPF check will always be performed on the packets that are received by TIRs through pseudowires.

3) Hop Count values: When a TRILL data packet is received over a pseudowire by a TIR, the TIR does the processing of Hop Count defined in [RFC6325] and will not perform any resetting of Hop Count.

```
          +-----+                                               +-----+
          | RBa +---+      ...........................      +---| RBb |
          +-----+   |      .                         .      |   +-----+
          Site 1    |    +----+                   +----+    |    Site 2
                    +----|TIR1|                   |TIR2|----+
                         +----+    MPLS Cloud     +----+
                           .                         .
                           .         +----+          .
                           ..........|TIR3|...........
                                     +----+      ^
                                        |        |
                                        |        +-- Emulated TRILL
                                     +-----+
                                     | RBc |
                                     +-----+
                                     Site 3
```

Figure 3: Topological Model of VPTS/TIR
connecting three TRILL Sites

In the above figure (Figure 3) Site1, Site2 and Site3 (running the TRILL protocol) are connected to TIR Devices. These TIR devices, along with the MPLS cloud, look like an unified emulated TRILL network. Only the PE devices in the MPLS network should be replaced with TIRs so the intermediate Provider routers are agnostic to the TRILL protocol.

Figure 4 below extends the topological model of TRILL over MPLS to connect multiple TRILL sites belonging to a tenant (tenant here is a campus, not a Data label) using VPTS model. VTSD1 and VTSD2 are two Virtual TRILL Switch Domains (Virtual RBridges) that segregate Tenant1's traffic from Tenant2's traffic. VTSD1 will maintain its own TRILL database for Tenant1, similarly VTSD2 will maintain its own TRILL database for Tenant2.

```
    +-----+          ............................         +-----+
    |RBat1+---+      . ######################## .     +---|RBbt1|
    +-----+   |      . #                      # .     |   +-----+
Tenant1 Site 1|    +-----+                +-----+    |  Tenant1 Site 2
              +----|VTSD1|                |VTSD1|----+
              +----|VTSD2|   MPLS  Cloud   |VTSD2|----+
              |    +-----+                +-----+    |
    +-----+   |      . #                       # .    |   +-----+
    |RBat2+---+      . #########+-----+######### .    +---|RBbt2|
    +-----+          ...........|VTSD1|...........        +-----+
Tenant2 Site2                   |VTSD2|          ^       Tenant2 Site 2
                                +-----+          |
```

```
                                |             |
                             +-----+          +-----Emulated
```

```
                               |RBct2|                  TRILL
                               +-----+
                            Tenant2 Site 3

        .... VTSD1 Connectivity
        #### VTSD2 Connectivity
```

Figure 4: Topological Model of VPTS/TIR
connecting 2 tenants with three TRILL Sites

4.1. Entities in the VPTS Model

The CE devices are defined in [RFC4026].

The Generic L2VPN Transport Functional Components like Attachment Circuits, Pseudowires etc. are defined in [RFC4664].

The RB (RBridge) and TRILL Campus are defined in [RFC6325] as updated by [RFC7780].

This model introduces two new entities called TIR and VTSD.

4.1.1. TRILL Intermediate Routers (TIR)

The TIRs (TRILL Intermediate Routers) must be capable of running both VPLS and TRILL protocols. TIR devices are a superset of the VPLS-PE devices defined in [RFC4026] with the additional functionality of TRILL. The VSI instance that provides transparent bridging functionality in the PE device is replaced with VTSD in a TIR.

4.1.2. Virtual TRILL Switch/Service Domain (VTSD)

The VTSD (Virtual Trill Switch Domain) is similar to VSI (layer 2 bridge) in the VPLS model, but the VTSD acts as a TRILL RBridge. The VTSD is a superset of VSI and must support all the functionality provided by the VSI as defined in [RFC4026]. Along with VSI functionality, the VTSD must be capable of supporting TRILL protocols and forming TRILL adjacencies. The VTSD must be capable of performing all the operations that a standard TRILL Switch can do.

One VTSD instance per tenant must be maintained, when multiple tenants are connected to a TIR. The VTSD must maintain all the information maintained by the RBridge on a per tenant basis. The VTSD must also take care of segregating one tenant traffic from other.

Each VTSD should have its own nickname, If a TIR supports 10 TRILL tenants, it needs to be assigned with ten TRILL nicknames, one for the nickname space of each of its tenants, and run ten copies of TRILL protocols, one for each tenant.

4.2. TRILL Adjacency for VPTS model

The VTSD must be capable of forming TRILL adjacency with other VTSDs present in its peer VPTS neighbor, and also the neighbor RBridges present in the TRILL sites. The procedure to form TRILL Adjacency is specified in [RFC7173] and [RFC7177].

4.3. MPLS encapsulation for VPTS model

The VPTS model uses PPP or Ethernet pseudowires for MPLS encapsulation as specified in [RFC7173], and requires no changes in the packet format in that RFC.

4.4. Loop Free provider PSN/MPLS

This model isn't required to employ Split Horizon mechanism in the provider PSN network, as TRILL takes care of Loop free topology using Distribution Trees. Any multi-destination packet will traverse a distribution tree path. All distribution trees are calculated based on TRILL base protocol standard [RFC6325] as updated by [RFC7780].

4.5. Frame Processing

This section specifies multi-destination and unicast frame processing in VPTS/TIR model.

4.5.1. Multi-Destination Frame Processing

Any multi-destination (unknown unicast, multicast or broadcast, as indicated by multi-destination bit in the TRILL Header) packets inside VTSD will be processed or forwarded through the distribution tree for which they were encapsulated on TRILL ingress. If any multi-destination packet is received from the wrong pseudowire at a VTSD, the TRILL protocol running in the VTSD will perform an RPF check as specified in [RFC7780] and drop the packet.

The Pruning mechanism in Distribution Trees, as specified in [RFC6325] and [RFC7780], can also be used to avoid forwarding of multi-destination data packets on the branches where there are no potential destinations.

4.5.2. Unicast Frame Processing

Unicast packets must be forwarded in same way they get forwarded in a standard TRILL Campus as specified in [RFC6325]. If multiple equal cost paths are available over pseudowires to reach destination, then VTSD should be capable of doing ECMP for them.

5. Extensions to TRILL Over Pseudowires [RFC7173]

The [RFC7173] mentions how to interconnect a pair of Transparent Interconnection of Lots of Links (TRILL) switch ports using pseudowires. This document explains, how to connect multiple TRILL sites (not limited to only two sites) using the mechanisms and encapsulations defined in [RFC7173].

6. VPTS Model Versus VPLS Model

VPLS Model uses a simpler loop-breaking rule: the "split horizon" rule, where a PE must not forward traffic from one PW to another in the same VPLS mesh, whereas the VPTS Model uses distribution Trees for loop free topology.

7. Packet Processing Between Pseudowires

Whenever a packet gets received over a pseudowire, a VTSD will decapsulate the MPLS headers followed by checking the TRILL header. If the egress nickname in the TRILL header is for a TRILL site located beyond another pseudowire, then VTSD will encapsulate with new MPLS headers and send it across the proper pseudowire.

For example in figure 3, consider that the pseudowire between TIR1 and TIR2 fails, Then TIR1 will communicate with TIR2 via TIR3, whenever packets which are destined to TIR3 gets received from pseudowire between TIR1 and TIR3, VTSD inside TIR3 will decapsulate

the MPLS headers, then check the TRILL header's egress nickname field. If the egress nickname indicate it is destained for the RBridge in site3 then the packet will be sent to RBc, if the egress nickname is located at site2, VTSD will add MPLS headers for the pseudowire between TIR3 and TIR2 and forward the packet on that pseudowire.

8. Efficiency Considerations

Since the VPTS Model uses Distribution trees for processing of multi-destination data packets, it is always advisable to have at least one Distribution tree root to be located in every TRILL site. This will avoid data packets getting received at TRILL sites where end-station service is not enabled for that data packet.

9. Security Considerations

For general TRILL security considerations, see [RFC6325]

For transport of TRILL by Pseudowires security consideration, see [RFC7173].

For general VPLS security considerations, see [RFC4761] and [RFC4762]

10. IANA Considerations

This document requires no IANA actions. RFC Editor: Please delete this section before publication

Normative References

[IS-IS] "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002, 2002".

[RFC4761] Kompella, K., Ed., and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <https://www.rfc-editor.org/info/rfc4761>.

[RFC4762] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <https://www.rfc-editor.org/info/rfc4762>.

[RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, DOI 10.17487/RFC6325, July 2011, <https://www.rfc-editor.org/info/rfc6325>.

[RFC7041] Balus, F., Ed., Sajassi, A., Ed., and N. Bitar, Ed., "Extensions to the Virtual Private LAN Service (VPLS) Provider Edge (PE) Model for Provider Backbone Bridging", RFC 7041, DOI 10.17487/RFC7041, November 2013, <https://www.rfc-editor.org/info/rfc7041>.

[RFC7173] Yong, L., Eastlake 3rd, D., Aldrin, S., and J. Hudson, "Transparent Interconnection of Lots of Links (TRILL) Transport Using Pseudowires", RFC 7173, DOI 10.17487/RFC7173, May 2014, <https://www.rfc-editor.org/info/rfc7173>.

[RFC7177] Eastlake 3rd, D., Perlman, R., Ghanwani, A., Yang, H., and V. Manral, "Transparent Interconnection of Lots of Links (TRILL): Adjacency", RFC 7177, DOI 10.17487/RFC7177, May 2014, <https://www.rfc-editor.org/info/rfc7177>.

[RFC7780] Eastlake 3rd, D., Zhang, M., Perlman, R., Banerjee, A., Ghanwani, A., and S. Gupta, "Transparent Interconnection of Lots of Links (TRILL): Clarifications, Corrections, and Updates", RFC 7780, DOI 10.17487/RFC7780, February 2016, <https://www.rfc-editor.org/info/rfc7780>.

Informative References

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, DOI 10.17487/RFC4026, March 2005, <https://www.rfc-editor.org/info/rfc4026>.

[RFC4664] Andersson, L., Ed., and E. Rosen, Ed., "Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, DOI 10.17487/RFC4664, September 2006, <https://www.rfc-editor.org/info/rfc4664>.

[RFC7172] Eastlake 3rd, D., Zhang, M., Agarwal, P., Perlman, R., and D. Dutt, "Transparent Interconnection of Lots of Links (TRILL): Fine-Grained Labeling", RFC 7172, DOI 10.17487/RFC7172, May 2014, <https://www.rfc-editor.org/info/rfc7172>.

Acknowledgements

The contributions of Andrew G. Malis is gratefully acknowledged in improving the quality of this document.

Authors' Addresses

Mohammed Umair
Cisco Systems
SEZ, Cessna Business Park
Sarjapur - Marathahalli Outer Ring road
Bengaluru - 560103, India

EMail: mohammed.umair2@gmail.com

Kingston Smiler Selvaraj
IPInfusion
RMZ Centennial
Mahadevapura Post
Bangalore - 560048 India

EMail: kingstonsmiler@gmail.com

Donald E. Eastlake 3rd
Huawei Technologies
155 Beaver Street
Milford, MA 01757
USA

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

Lucy Yong
Self

Phone: +1-469-227-5837
EMail: lucyyong@gmail.com