TRILL Weiguo Hao
Yizhou Li
Donald Eastlake
Liang Xia
Internet Draft Huawei
Andrew Qu
MediaTec
Muhammad Durrani
Brocade
Ponkarthick S.
IP Infusion
Intended status: Standards Track October 15, 2014
Expires: March 2015

TRILL Distributed Layer 3 Gateway
draft-hao-trill-irb-05.txt

Abstract

Currently TRILL protocol provides optimal pair-wise data frame forwarding for
layer 2 intra-subnet traffic but not for layer 3 inter-subnet traffic. A
centralized gateway solution is typically used for layer 3 inter-subnet traffic
forwarding but has following issues:

1. Sub-optimum forwarding path for inter-subnet traffic.

2. Huge number of gateway interfaces, 16 million in extreme case, need to be
supported on the centralized gateway.

3. Traffic bottleneck at the gateway.

An optional TRILL distributed gateway solution that resolves these centralized
gateway issues is specified in this document.

Status of this Memo

be created, and it may not be published except as an Internet-Draft.

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not

Copyright Notice

Table of Contents

1. Introduction

The IETF has standardized the TRILL (Transparent Interconnection of Lots of Links) protocol [RFC6325] that provides a solution for least cost transparent routing in multi-hop networks with arbitrary topologies and link technologies, using [IS-IS] [RFC7176] link-state routing and a hop count. TRILL switches are sometimes called RBridges (Routing Bridges).

Currently, TRILL provides optimal unicast forwarding for Layer 2 intra-subnet traffic but not for Layer 3 inter-subnet traffic. In this document, an optional TRILL-based distributed Layer 3 gateway solution is specified to provide optimal unicast forwarding for Layer 3 inter-subnet traffic. With distributed gateway support an edge RBridge provides both routing based on Layer 2 identity (address and virtual network (VN)) among end stations (ESs) that belong to same subnet and routing based on Layer 3 identity among ESs that belong to different subnets of the same routing domain. An edge RBridge needs to provide routing instances and Layer 3 gateway interfaces for local connected ESs. The routing instances are for IP address isolation between tenants. In the TRILL distributed Layer 3 gateway solution, inter-subnet traffic can be fully dispersed among edge RBridges, so there is no single bottleneck.

This document is organized as follows: Section 3 describes why a distributed gateway solution is beneficial. Section 4 gives the Layer 3 traffic

forwarding
model. Section 5 provides a distributed gateway solution overview. Section 6
gives
a distributed gateway example. And Section 7 describes the TRILL protocol
extensions needed to support this distributed gateway solution.

2. Conventions Used in This Document

The terms and acronyms in [RFC6325] are used with the following additions:

Data Label: VLAN or FGL [RFC7172].

DCN: Data Center Network.

ES: End Station. VM (Virtual Machine) or physical server, whose
address is
either the destination or source of a data frame.

GW: Gateway.

Gateway interface: Layer 3 virtual interface on gateway aka gateway interface) terminates layer 2 forwarding and forwards IP traffic to the destination as per IP forwarding rules. Incoming traffic from a physical port on a gateway will be distributed to its virtual gateway interface based on Data Label (VLAN or FGL).

L2:   Layer 2.

L3:   IP Layer 3.

ND: IPv6's Neighbor Discovery [RFC4861].

RD: Routing Domain.

ToR: Top of Rack.

VN: Virtual Network. In a TRILL campus, each virtual network is identified by a unique 12-bit VLAN ID or 24-bit Fine Grained Label [RFC7172].

VRF: Virtual Routing and Forwarding. In IP-based computer networks, Virtual Routing and Forwarding (VRF) is a technology that allows multiple instances of a routing table to co-exist within the same router at the same time.

3. Problem Statement

```
         -------                -------
         | GW1 |                | GW2 |
         -------                -------
            |                      |
         -------                -------
         |AGG1 |                |AGG2 |
         -------                -------
            |                      |
      -----------------------------------------------------
      |  -------------|------------------|----------------|
      |  |            |  |                |  |          |  |
    -------          -------           -------        -------
    |TOR1 |          |TOR2 |           |TOR3 |        |TOR4 |
    -------          -------           -------        -------
     |    |           |    |            |    |         |    |
   ----  ----       ----  ----        ----  ----     ----  ----
   |E |  |E |       |E |  |E |        |E |  |E |     |E |  |E |
```

```
      |S1|  |S2|       |S3|  |S4|        |S5|  |S6|     |S7|  |S8|
      ----  ----       ----  ----        ----  ----     ----  ----
```

Figure 1 A typical DC network

Figure 1 depicts a Data Center Network (DCN) using TRILL where edge RBridges are Top of Rack (ToR) switches. Centralized gateway GW1 and GW2 in figure 1 provide the layer 3 packet forwarding for both north-south traffic and east-west inter-subnet traffic between ESs.

End stations in one IP subnet expect to send IP traffic for a different subnet to an IP router. In addition, there is normally a Data Label (VLAN or FGL) associated with each IP subnet but there is no facility in TRILL to change the Data Label for traffic between subnets. If two end stations of the same tenant are on two different subnets and need to communicate with each other, their packets are typically forwarded all the way to a centralized IP Layer 3 gateway to perform L3 forwarding and, if necessary, change the Data Label. This is generally sub-optimal because the two end stations may be connected to the same ToR where L3 switching could have been performed locally. For example, in above Figure 1, assuming ES1 (10.1.1.2) and ES2 (20.1.1.2) belong to different subnets of same tenant, the unicast IP traffic between them has to go through a centralized gateway. It can't be locally forwarded on TOR1. If an edge RBridge has distributed gateway capabilities, then it can perform optimum L2 forwarding for intra-subnet traffic and optimum L3 forwarding for inter-subnet traffic, delivering optimum forwarding for unicast packets in all important cases.

When Fine Grained Label [RFC7172] is introduced, up to 16 million Layer 2 VN can be supported in a TRILL campus. To support inter-subnet traffic, up to 16 million Layer 3 gateway interfaces should be created on a centralized gateway if each VN corresponds to a subnet. It is a huge burden for the centralized gateway to support so many interfaces. In addition all inter-subnet traffic will go through the centralized gateway that may become the traffic bottleneck.

In summary, the centralized gateway has the following issues:

1. Sub-optimum forwarding paths for inter-subnet traffic due to the requirements to perform IP routing and possibly change Data Labels.

2. Huge number of gateway interfaces, 16 million in the extreme case, need to be supported on the centralized gateway.

3. Traffic bottleneck at the gateway.

A distributed gateway on edge RBridges addresses these issues.

4. Layer 3 Traffic Forwarding Model

```
                  +----------------------------------------------+
                  |                                              |
                  |      +------------+         +------------+   |
                  |      | Tenant n  |---------|  VRF n     |   |
                  |   +-------------+ |      +-------------+  |   |
                  |   |  +-----+    | |      |            |  |   |
                  |   |  | VN1 |    | |      |            |  |   |
                  |   |  +-----+    | |      |    VRF 1   |  |   |
                  |   |     ..     +--------+            |  |   |
                  |   |  +-----+    | |      |            |  |   |
                  |   |  | VNm |    | |      |            |  |   |
                  |   |  +-----+    | |      |            |  |   |
                  |   |  Tenant 1  |-+      |            |  |   |
                  |   +-------------+        |            |  |   |
                  |   +-------------+       +-------------+      |
                  |                                              |
                  |               Edge RBridge                   |
                  +----------------------------------------------+
```

Figure 2 Edge RBridge Model as distributed GW

In a data center network (DCN), each tenant may include one or more Layer 2 virtual networks and, in normal cases, each tenant corresponds to one routing domain (RD). Normally each Layer 2 virtual network uses a different Data Label and corresponds to one or more subnets.

Each Layer 2 virtual network in a TRILL campus is identified by a unique 12-bit VLAN ID or 24-bit Fine Grained Label [RFC7172]. Different routing domains may have overlapping address space but need distinct and separate routes. The end stations that belong to the same subnet communicate through L2 forwarding, end systems of the same tenant that belong to different subnets communicate through L3 forwarding.

The above figure 2 depicts the model where there are N VRFs corresponding to N tenants with each tenant having up to M segments/subnets (virtual network).

5. Distributed Gateway Solution Overview

In the TRILL distributed gateway scenario, an edge RBridge must perform Layer 2 routing for the ESs that are on the same subnet and IP routing for the ESs that are on the different subnets of the same tenant.

As the IP address space in different routing domains can overlap, VRF instances need to be created on each edge RBridge to isolate the IP forwarding process among different routing domains present on the edge RBridge. A globally unique tenant ID identifies each routing domain. The network operator should ensure the consistency of the tenant ID on each edge RBridge for each routing domain. If a routing domain spreads over multiple edge RBridges, routing information for the routing domain should be synchronized among these edge RBridges to ensure the reachability to all ESs in that routing domain. The Tenant ID should be carried with the routing information to differentiate the routing domains.

From the data plane perspective, all edge RBridges are connected to each other via one or multiple TRILL hops, however they are always a single IP hop away. When an ingress RBridge receives inter-subnet traffic from a local ES whose destination MAC is the edge RBridge's gateway MAC, that RBridge will perform Ethernet header termination and look up in its IP forwarding table to forward the traffic to the IP next hop. If the destination ES is connected to a remote edge RBridge, the remote RBridge will be the IP next hop for traffic forwarding. The ingress RBridge will perform TRILL encapsulation for such inter-subnet traffic and route it to the remote RBridge through the TRILL campus.

When that remote RBridge receives the traffic, it will decapsulate the packet and then lookup in the RBridge's IP forwarding table to route it to the destination ES. Through this method, TRILL with distributed gateways provides pair-wise data routing for inter-subnet traffic.

5.1. Local routing information

An ES can be locally connected to an edge RBridges through a layer 2 network or externally connected through a layer 3 IP network.

If the ES is connected to an edge RBridge through a Layer 2 network, then the edge RBridge must act as a Layer 3 GW for the ES. The gateway interface should be established on the edge RBridge for the connecting ES. Because the ESs in the same subnet may be spread over multiple edge RBridges, each of these edge RBridges should establish its gateway interface for the subnet and these gateway interfaces on different edge RBridges share the same gateway MAC and gateway IP address.

Before an ES starts to send inter-subnet traffic, it should acquire its gateway's MAC through the ARP/ND process. Local connecting edge RBridges that are supporting this distributed gateway feature always respond with the gateway MAC address when receiving ARP/ND requests for the gateway IP. Through the ARP/ND process, the edge RBridge can learn the IP and MAC correspondence of local ES connected to the edge RBridge by Layer 2 and then generate local IP routing entries for the ES in the corresponding routing domain.

If an ES is located in an external IP network, the ES also can be connected to the TRILL campus through a TRILL edge RBridge. The TRILL edge RBridge runs a unified routing protocol with the external IP network for each routing domain. The edge RBridge learns the IP prefix corresponding to the ES through the IP routing protocol, then the RBridge generates local IP routing entries in the corresponding routing domain.

5.2. Local routing information synchronization

Each edge RBridge should announce its own tenant gateway MAC to the TRILL campus. The tenant gateway MAC is to differentiate inter-subnet Layer 3 traffic or intra-subnet Layer 2 traffic on an egress RBridge; the ingress RBridge will use the tenant gateway MAC announced by the egress RBridge as the Inner.MacDA for inter-subnet traffic TRILL encapsulation. All tenants on a RBridge can share the same tenant gateway MAC for inter-subnet traffic purposes.

When a routing instance is created on an edge RBridge, the tenant ID, tenant Label (VLAN or FGL), and their correspondence should be set and globally advertised. The ingress RBridge uses the Label advertised by the egress RBridge as the inner VLAN or FGL when it performs inter-subnet traffic TRILL encapsulation. The egress RBridge relies on tenant Label to find the local VRF instance for the IP forwarding process when receiving inter-subnet traffic from the TRILL campus. (The role of tenant Label is akin to an MPLS VPN Label in an MPLS IP/MPLS VPN network.) Tenant Labels are independently allocated on each edge RBridge for each routing domain, an edge RBridge can pick up an access Label in a routing domain to act as inter-subnet Label, or the edge RBridge can use a different Label from any access Labels to act as tenant Label. It's implementation dependant and there is no restriction on this.

When a local IP prefix is learned in a routing instance on an edge RBridge, the edge RBridge should advertise the IP prefix information for the routing instance to other edge RBridges to generate IP routing entries. A globally unique tenant ID also should be carried to differentiate IP prefixes between different tenants, because the IP address space of different tenants can overlap.

TRILL FS-LSP [rfc7180bis] extensions can be used for IP routing information synchronization in each routing domain among edge RBridges. Based on the synchronized information from other edge RBridges, each edge RBridge generates remote IP routing entries in each routing domain.

Through this solution, intra-subnet forwarding function and inter-subnet IP routing functions are integrated and network management and deployment will be simplified.

5.3. Data traffic forwarding process

After a Layer 2 connected ES1 in VLAN-x acquires its gateway's MAC, it can start inter-subnet data traffic process to ES2 in VLAN-y. When the local connecting edge RBridge receives inter-subnet traffic from ES1, the RBridge performs Layer 2 header termination, then, using the local VRF corresponding to VLAN-x, it performs the IP forwarding process in that VRF.

If destination ES2 is also attached to the ingress RBridge, the traffic will be locally forwarded to ES2 on the ingress RBridge. Compared to the centralized gateway solution, the forwarding path is optimal and a traffic detour is avoided.

If ES2 is attached to a remote edge RBridge, the remote edge RBridge is IP next hop and the inter-subnet traffic is forwarded to the IP next hop through TRILL encapsulation. If there are multiple equal cost shortest path between ingress RBridge and egress RBridge, all these path can be used for inter-subnet traffic forwarding, so pair-wise load spreading can be achieved for inter-subnet traffic.

When the remote RBridge receives the inter-subnet TRILL encapsulated traffic, the RBridge decapsulates the TRILL encapsulation and checks the Inner.MacDA, if that MAC address is the local gateway MAC corresponding to the inner Label (VLAN or FGL), the inner Label will be used to find the corresponding local VRF, then the IP forwarding process in that VRF will be performed, and the traffic will be locally forwarded to the destination ES2.

In summary, through this solution, traffic detours to a central gateway are avoided, both inter-subnet and intra-subnet traffic can be forwarded along pair-wise shortest paths, and network bandwidth is conserved.

6. Distributed Layer 3 Gateway Process Example

```
                        ---------             ---------
                        |  RB3  |             |  RB4  |
                        ---------             ---------
                         #   *                        *
                         #   **************************
                         ###########################  *
                         #   *                     #  *
                         #   *                     #  *
                         #   *                     #  *
                        ---------              ---------
                        |  RB1  |              |  RB2  |
                        ---------              ---------
                           |                      |
```

```
                                    ----                   ----
                                     |E |                   |E |
                                     |S1|                   |S2|
                                    ----                   ----
```

Figure 3 Distributed gateway scenario

In figure 3 above, RB1 and RB2 support the distribution gateway function, ES1 connects to RB1, ES2 connects to RB2. ES1 and ES2 belong to Tenant1, but are in different subnets.

The IP address, VLAN, and subnet information of ES1 and ES2 are as follows:

ES	Tenant	IP Address	Subnet	VLAN
ES1	Tenant1	10.1.1.2	10.1.1.1/32	10
ES2	Tenant1	20.1.1.2	20.1.1.1/32	20

Figure 4 ES information

The nickname, VRF, tenant VLAN, tenant gateway MAC for Tenant1 on RB1 and RB2 are as follows:

RB	Nickname	Tenant	VRF	Tenant VLAN	Gateway MAC
RB1	nick1	Tenant1	VRF1	100	MAC1
RB2	nick2	Tenant1	VRF2	100	MAC2

Figure 5 RBridge information

6.1. Control plane process

RB1 announces the following local routing information to the TRILL campus:

Tenant ID: 1

Tenant gateway MAC: MAC1.

Tenant VLAN for Tenant1: VLAN 100.

IP prefix in Tenant1: 10.1.1.2/32.

RB2 announces the following local routing information to TRILL campus:

Tenant ID: 1

Tenant gateway MAC: MAC2.

Tenant VLAN for Tenant1: VLAN 100.

IP prefix in Tenant1: 20.1.1.2/32.

Relying on the routing information from RB2, remote routing entries on RB1 are generated as follows:

Prefix/Mask	Inner.MacDA	inner VLAN	egress nickname
20.1.1.2/32	MAC2	100	nick2

Figure 6 Tenant 1 remote routing table on RB1

Similarly, relying on the routing information from RB1, remote routing entries on RB2 are generated as follows:

Prefix/Mask	Inner.MacDA	inner VLAN	egress nickname
10.1.1.2/32	MAC1	100	nick1

Figure 7 Tenant 1 remote routing table on RB1

6.2. Data plane process

Assuming ES1 sends unicast inter-subnet traffic to ES2, the traffic forwarding process is as follows:

1. ES1 sends unicast inter-subnet traffic to RB1 with RB1's gateway's MAC as the destination MAC.

2. Ingress RBridge (RB1) forwarding process:

RB1 checks the destination MAC, if the destination MAC equals the local gateway MAC, the gateway function will terminate the Layer 2 header and perform L3 forwarding process.

RB1 looks up IP routing table information by destination IP and Tenant ID to get IP next hop information, which includes the egress RBridge's gateway MAC (MAC2), tenant VLAN (VLAN 100) and egress nickname (nick2). Using this information, RB1 will perform inner Ethernet header encapsulation and TRILL encapsulation. RB1 will use MAC2 as the Inner.MacDA, MAC1 (RB1's own gateway MAC) as the Inner.MacSA, VLAN 100 as the Inner.VLAN, nick2 as the egress nickname and nick1 as the ingress nickname.

RB1 looks up TRILL forwarding table by egress nickname and sends the traffic to the TRILL next hop as per [RFC6325]. The traffic will be sent to RB3 or RB4 as result of load balancing.

Assuming the traffic is forwarded to RB3, the following occurs:

3. Transit RBridge (RB3) forwarding process:

RB3 looks up TRILL forwarding information by egress nickname and forwards the traffic to RB2 as per [RFC6325].

4. Egress RBridge forwarding process:

As the egress nickname is RB2's own nickname, RB2 performs TRILL decapsulation. Then it checks the Inner.MacDA and, because that MAC is equal to the local gateway MAC, performs inner Ethernet header termination. Relying on inner VLAN, RB2 finds the local corresponding VRF and looks up the packets destination IP address in the VRF's IP routing table. The traffic is then be locally forwarded to ES2.

7. TRILL Protocol Extensions

If an edge RBridge RB1 participates in the distributed gateway function, it should announce its tenant gateway MAC, tenant Label and IPv4/IPv6 prefix to the TRILL campus through the tenant gateway MAC APPsub-TLV, tenant Label APPsub-TLV and IPv4/IPv6 prefix APPsub-TLV. Other edge RBridges belonging to the same routing domain use this information to generate IP routing entries in that routing domain. The ingress RBridge uses the tenant gateway MAC and tenant Label of the egress RBridge to perform inter-subnet traffic TRILL encapsulation when it receives inter-subnet traffic from a local ES. The tenant gateway MAC is used as the Inner.MacDA and the tenant Label is used as the Inner.Label.

The following APPsub-TLVs MUST be included in a TRILL GENINFO TLV in FS-LSPs [rfc7180bis].

7.1. The tenant gateway MAC APPsub-TLV

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type                        | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Length                      | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+
|   Tenant gateway MAC          | (6 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+
```

o Type: Set to TENANT-GWMAC sub-TLV (TBD1). Two bytes, because this APPsub-TLV appears in an extended TLV [RFC7356].

o Length: 6.

o Tenant gateway MAC: The local tenant gateway MAC for inter-subnet traffic forwarding.

7.2. The tenant Label APPsub-TLV

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type                     | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Length                   | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Tenant ID (4 bytes)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|Resv|     Label1     | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Resv2|     Label2     | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o Type: Set to TENANT-LABEL sub-TLV (TBD2). Two bytes, because this APPsub-TLV appears in an extended TLV [RFC7356].

o Length: If Label1 field is used to represent a VLAN, the value of the length field is 12. If Label1 and Label2 field are used to represent an FGL, the value of the length field is 14.

o Tenant ID: This identifies a global tenant ID.

o L: 1 bit. When Label1 and Label2 field are used to identify an FGL, this bit is set to 1. When Label1 field is used to identify a VLAN, it is set to 0.

o Resv: 3 bits that MUST be sent as zero and ignored on receipt.

o Label1: If the value of length field is 12, the field is to identify tenant VLAN ID. If the value of length field is 14, the field is to identify higher 12 bits of tenant FGL.

o Resv2: 4 bits that MUST be sent as zero and ignored on receipt. Only present if the length field is 14.

o Label2: This field has the lower 12 bits of tenant FGL. Only present if the length field is 14.

7.3. The IPv4 Prefix APPsub-TLV

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type                         | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Total Length                 | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
|                    Tenant ID                       |(4 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
| Prefix Length(1)|(1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
|                    Prefix (1)                      |(variable)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
|     .....       |(1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
|                   .....                            |(variable)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
| Prefix Length(N)|(1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
|                    Prefix (N)                      |(variable)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
```

o Type: Set to IPV4-PREFIX sub-TLV (TBD3). Two bytes, because this APPsub-TLV appears in an extended TLV [RFC7356].

o Total Length: This 2-byte unsigned integer indicates the total length of Tenant ID, Prefix Length and Prefix fields in octets. A value of 0 indicates that no IPv4 prefix is being advertised.

o Tenant ID: This identifies a global tenant ID.

o Prefix Length: The Prefix Length field indicates the length in bits of the IPv4 address prefix. A length of zero indicates a prefix that

matches all IPv4 addresses (with prefix, itself, of zero octets).

o Prefix: The Prefix field contains an IPv4 address prefix, followed by enough trailing bits to make the end of the field fall on an octet boundary. Note that the value of the trailing bits is irrelevant.

7.4. The IPv6 Prefix APPsub-TLV

```
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   Type                        | (2 bytes)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   Total Length                | (2 bytes)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   |                     Tenant ID                  |(4 bytes)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   | Prefix Length(1)|(1 byte)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   |                     Prefix (1)                 |(variable)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   |     .....       |(1 byte)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   |                    .....                       |(variable)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   | Prefix Length(N)|(1 byte)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
   |                     Prefix (N)                 |(variable)
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...-+-+-+-+-+-+-+-+-+
```

o Type: Set to IPV6-PREFIX sub-TLV (TBD4). Two bytes, because this APPsub-TLV appears in an extended TLV [RFC7356].

o Total Length: This 2-byte unsigned integer indicates the total length of Tenant ID, Prefix Length and Prefix fields in octets. A value of 0 indicates that no IPv6 prefix is being advertised.

o Tenant ID: This identifies a global tenant ID.

o Prefix Length: The Prefix Length field indicates the length in bits of the IPv6 address prefix. A length of zero indicates a prefix that matches all IPv6 addresses (with prefix, itself, of zero octets).

o Prefix: The Prefix field contains an IPv6 address prefix, followed by enough trailing bits to make the end of the field fall on an octet boundary. Note that the value of the trailing bits is irrelevant.

8. Security Considerations

Correct configuration of the edge RBridges participating is important to assure
that data is not delivered to the wrong tenant, which would violate security constrains. IS-IS security [RFC5310] can be used to secure the information advertised by the edge RBridges.

Particularly sensitive data should be encrypted end-to-end, that is, from the source end station to the destination end station.

For general TRILL Security Considerations, see [RFC6325].

9. IANA Considerations

IANA is requested to assign four APPsub-TLV type numbers less than 255 under the TRILL GENINFO TLV [RFC7357] as follows:

Type	Name	References
TBD1	TENANT-GWMAC	[this document]
TBD2	TENANT-LABEL	[this document]
TBD3	IPV4-PREFIX	[this document]
TBD4	IPV6-PREFIX	[this document]

10. Normative References

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

[2]   [RFC6325] Perlman, R., Eastlake 3rd, D., Dutt, D., Gai, S.,and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.

[3]   [RFC7172] Eastlake, D., M. Zhang, P. Agarwal, R. Perlman, D. Dutt, "TRILL (Transparent Interconnection of Lots of Links): Fine-Grained Labeling", RFC7172, May 2014.

[4]   [RFC7176] - Eastlake, D., T. Senevirathne, A. Ghanwani, D. Dutt and A. Banerjee" Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC7176, May 2014.

[5]   [RFC7357] - Zhai, H., Hu, F., Perlman, R., Eastlake 3rd, D., and O. Stokes, "Transparent Interconnection of Lots of Links (TRILL): End Station Address Distribution Information (ESADI) Protocol", RFC 7357, September 2014.

11. Informative References

[1] [RFC5310] - Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, February 2009.

12. Acknowledgments

The authors wish to acknowledge the important contributions of

Guangrui Wu, Zhenbin Li, Zhibo Hu.

Authors' Addresses

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Phone: +86-25-56623144
Email: haoweiguo@huawei.com

Yizhou Li
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Phone: +86-25-56625375
Email: liyizhou@huawei.com

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

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

Liang Xia(Frank)
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Phone: +86-25-56624539
Email: frank.xialiang@huawei.com

Andrew Qu
MediaTec
Email: laodulaodu@gmail.com

Muhammad Durrani
Brocade
130 Holger Way
San Jose, CA 95134
EMail: mdurrani@brocade.com

Ponkarthick Sivamurugan
Address: IP Infusion,
RMZ Centennial
Mahadevapura Post
Bangalore - 560048
EMail: Ponkarthick.sivamurugan@ipinfusion.com