Extension to BGP-VPLS for E-Tree
draft-cao-l2vpn-bgp-vpls-etree-00.txt

Status of this Memo

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

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 This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.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 be created, except to publish it as an RFC and to translate it into languages other than English.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

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 material or to cite them other than as "work in progress."
Abstract

This document proposes an approach to support Metro Ethernet Forum (MEF) Ethernet Tree (E-Tree) in Virtual Private LAN Service using BGP for auto-discovery and signaling [RFC4761]. The proposed solution is characterized by breaking communication channels between Leafs to fulfill the specific E-Tree requirement: Leaf cannot communicate with Leaf. Backward compatibility is also considered.

Table of Contents

1. Introduction................................................3
2. Conventions used in this document..........................4
3. Terminology................................................4
4. Reference Model............................................4
5. Extension to VPLS for E-Tree................................6
   5.1. Assumptions............................................6
1. Introduction

A specific Rooted-multipoint service, Ethernet Tree (E-Tree), has been defined by Metro Ethernet Forum [MEF6.1]. Compared with MEF Ethernet LAN (E-LAN) service where there is no communication restriction between its attachment circuits, each attachment circuit of E-tree is designated as either a Root or a Leaf. A Root-AC can communicate with all other attachment circuit in a E-Tree, however a Leaf-AC can only communicate with Root-ACs but not Leafs.

[Draft VPLS ETREE Req] provides the functional requirements for MEF E-Tree support in VPLS.

This document presents a minimal extension to the current VPLS standard [RFC4761] to break the "communication channel" between Leaf attachment circuits.

Figure 1 below describes scenario for Leaf-to-Leaf communication restriction.

\[\text{Figure 1 Scenario for Leaf-to-Leaf Communication Restriction}\]
If PE2 receives one frame from PE1 over Ethernet PW, PE2 does NOT know whether the frame comes from Root AC or Leaf AC, so it can not decide to forward the frame to AC4 (Leaf AC) or not with the current VPLS standards [RFC4761] [RFC4762].

2. Conventions used in this document

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. Terminology

There are two solutions to restrict Leaf-to-Leaf communication,

1. Each frame carries additional information to indicate that it is originated from a Leaf endpoint or Root endpoint on the ingress PE, then egress PE can know forward behavior of the frame, if it comes from Leaf, it will NOT be forwarded to the local Leaf ACs. [Draft Ext VPLS for ETree] proposes this solution.

2. If a frame from local Leaf-ACs, one PE in a VPLS will NOT forward it to its other local Leaf-ACs; if there is no PW between local Leafs and remote Leaf-ACs which are connected to remote PE, a frame from local Leafs also cannot be forwarded to remote Leafs. Then we can restrict the communication between Leafs.

The purposed solution in this document prefers to the second and two terms are introduced,

o Root-endpoint. One endpoint which connects only Root-ACs, one or more Root-ACs.

o Leaf-endpoint. One endpoint which connects only Leaf-ACs, one or more Leaf-ACs.

There is no endpoint which connects both Root-ACs and Leaf-ACs in the second solution.

4. Reference Model

Figure 2 below describes a typical reference model where Root ACs (AC1, AC5, AC6) can communicate with all other Ethernet ACs in the
VSI, but Leaf ACs (AC2, AC3, AC4) only can communicate with Root ACs (AC1, AC5, AC6), and can not communicate with each other.

Figure 2  E-Tree typical Reference Model

In most use cases, an E-Tree architecture has only a few Root ACs but many Leaf-ACs. On any PE in E-Tree, there are 3 cases,

- **Root-only ACs.** All ACs connected to VSI are Root ACs, say AC5 and AC6 of PE 3 in Figure 2 and at least one VE ID which stands for one Root endpoint is assigned.

- **Leaf-only ACs.** All ACs connected to VSI are Leaf ACs, say AC3 and AC4 of PE 2 in Figure 2 and at least one VE ID which stands for one Leaf endpoint is assigned.

- **Root-Leaf-Mixed ACs.** Some ACs connected to VSI are Root ACs and some are Leaf ACs, say AC1 and AC2 of PE 1 in Figure 2. Here network administrator should at least assign two VE IDs, one for Root ACs and called as Root-endpoint, one for Leaf ACs and called as Leaf-endpoint.

Within an E-Tree,
o All Root-ACs of a Root endpoint can receive frame from and transmit frame to any other endpoints in an E-Tree.

o A Root-AC of a Root endpoint can receive frame from and transmit frame to its other Root-ACs;

o A Leaf endpoint can receive frame from and transmit frame to any Root endpoints in an E-Tree.

o A Leaf endpoint cannot receive frame from and transmit frame to any other Leaf endpoints in the E-Tree.

o A Leaf-AC of a Leaf endpoint cannot receive frame from and transmit frame to its other Leaf-ACs;

For one VSI, PE 1 has both Root and Leaf ACs, so on PE 1, PE 1 establish one PW (PW12) for AC 1 (Root AC, also belongs to a Root endpoint) with PE 2 where remote ACs are Leaf-only, two PWs with PE 3 where PE 1 receives frames from or transmits frames to remote Root ACs for local Root ACs over PW13-1 and receives frames from or transmits frames to remote Root ACs for local Leaf ACs; PE 2 has Leaf-only ACs, so PE 2 establish one PW (PW12) with remote Root ACs on PE 1 and one PW (PW23) with remote Root ACs on PE3; PW3 which has Root-only ACs can establish PW with remote Leaf ACs and Root ACs, so PE 3 establish two PWs with PE 1 and one PW with PE2. However the ACs on PE 2 are Leaf, so any Ethernet frame which is received from AC 3 cannot be transmitted to other local Leaf ACs, say AC4. PE 2 also can not transmit Ethernet frame to remote Leaf ACs since there is no PW for it.

This applies to all traffic, including Unicast Known, Unicast Unknown, Broadcast or Multicast.

5. Extension to VPLS for E-Tree

5.1. Assumptions

The PEs are assumed to be (logically) fully meshed with tunnels over which packets that belong to a service (such as VPLS or E-Tree) are encapsulated and forwarded.

Any E-Tree endpoint comprises only one AC type. If a PE in a VPLS has both Root ACs and Leaf ACs, it SHOULD be configured with at least two endpoints, one is composed of Root ACs, and another is composed of Leaf ACs. It is illegal for any endpoint to have both at same time.
5.2. AC type

Each AC connected to an E-Tree on PE MUST have an AC attribute, Root AC or Leaf AC. For backward compatibility, the default AC type MUST be Root for current VPLS standard [RFC4761] [RFC4762].

- Root AC. It can communicate with all ACs in a VPLS or E-Tree and SHOULD belong to a Root endpoint.
- Leaf AC. It only can communicate with all Root ACs in a VPLS or E-Tree, and SHOULD belong to a Leaf endpoint.

5.3. PW setup Matrix

Just as mentioned in Section 3, there is no PW between Leaf-endpoints and Table 1 describes PW setup matrix,

| +---------------+-------+------+ |
| | endpoint type + Root + Leaf + |
| | +------------------------+ |
| | Root + Setup + Setup+ |
| | +------------------------+ |
| | Leaf + Setup + n/a + |
| +------------------------+ |

Table 1 PW setup Matrix

In the following cases PW may be established,

- Local Root-Remote Root
- Local Root-Remote Leaf or Local Leaf-Remote Root

Between PE 1 and PE 2 in Figure 1, we have to setup 3 PWs,

- PW 1: Communication between Root ACs, i.e., AC 1 and AC 3 in Figure 1.
- PW 2: Communication between Root AC and Leaf AC, i.e., AC 1 and AC 4 in Figure 1.
- PW 3: Communication between Root AC and Leaf AC, i.e., AC 2 and AC 3 in Figure 1.

5.4. VPLS BGP NLRI

Section 3.2.2 in [RFC 4761] defines VPLS BGP NLRI with a new AFI and SAFI to exchange VPLS membership and demultiplexers.
A PE participating in a E-Tree must have at least one VE ID, but for a VSI on a PE which has both Root-ACs and Leaf-ACs, it must have at least two VE IDs, one is called as Root endpoint and one is called as Leaf endpoint.

Here whole VE ID set is divided into two parts, one is Root VE ID set, and one is Leaf VE id set.

L = \{VBO, VBO+1, \ldots, VBO+LVBS-1\},

R = \{VBO+LVBS, VBO+LVBS +1,\ldots, VBO+VBS-1\},

Here VE ID, Leaf VE block Size (LVBS) and VE Block Size (VBO) are typically assigned by the network administrator. All Root VE IDs are in R set, and all Leaf VE IDs are in L set. If there are Root-only ACs on a PE, LVBS SHOULD be set as zero; if there are Leaf-only ACs, LVBS SHOULD be equal to VBS.

The endpoint which is identified by VE ID in L set only can establish PW with the endpoint identified by VE ID in R set, but the endpoint identified by the VE ID in R set can establish PW with all VPLS endpoint identified by VE ID in RUL.

5.5. PW setup and teardown

Suppose PE-a is part of E-Tree foo and has both Root-ACs and Leaf-ACs. For Root ACs, it is assigned with VE ID r which is in Root VE ID set, VE Block Offset VBO, VE Block Size VBS, and label base rLB; For Leaf ACs, it is assigned with VE ID l which is in Leaf VE ID set, VE block offset VBO+LVBS, VE Block Size VBS-LVBS, and label base lLB. If PE-b
is also part of E-Tree foo with VE ID w (Root or Leaf) and gets NLRI advertisement from PE-a, it will do the following,

5.5.1. Root endpoint

1. Checks if w is part of PE-a’s ‘remote VE set’: if VBO <= w < VBO+ VBS, then w is part of PE-a’s remote VE set. If not, PE-b ignores this message, and skips the rest of this procedure.

2. Sets up a PW to PE-a: the demultiplexor label to send traffic from PE-b to PE-a is computed as (rLB + W - VBO).

3. Checks if r is part of any ‘remote VE set’ that PE-b announced, i.e., PE-b checks if r belongs to some remote VE set that PE-b announced, say with VE Block Offset VBO’, VE Block Size VBS’, and label base LB’. If not, PE-b MUST make a new announcement as described.

4. Sets up a PW from PE-a: the demultiplexor label over which PE-b should expect traffic from PE-a is computed as: (LB’ + r - VBO’).

If PE-a withdraws an NLRI for r that PE-b was using, then PE-b MUST tear down its ends of the pseudowire between PE-a and PE-b.

5.5.2. Leaf endpoint

1. Checks if w is part of PE-a’s ‘remote VE set’: if VBO+LVBS <= w < VBO+ VBS, then w is part of PE-a’s remote Root VE set. If not, PE-b ignores this message, and skips the rest of this procedure.

2. Sets up a PW to PE-a: the demultiplexor label to send traffic from PE-b to PE-a is computed as (LB + w - VBO).

3. Checks if l is part of any ‘remote VE set’ that PE-b announced, i.e., PE-b checks if l belongs to some remote VE set that PE-b announced, say with VE Block Offset VBO’, VE Block Size VBS’, and label base LB’. If not, PE-b MUST make a new announcement as described.

4. Sets up a PW from PE-a: the demultiplexor label over which PE-b should expect traffic from PE-a is computed as: (LB’ + l - VBO’).

If PE-a withdraws an NLRI for l that PE-b was using, then PE-b MUST tear down its ends of the pseudowire between PE-a and PE-b.
5.6. Signaling PE Capabilities

The extended attribute in Section [RFC4761] 3.2.4, the "Layer2 Info Extended Community", is used to signal control information about the pseudowires to be setup for a VPLS. It also can carry endpoint information. It will be extended in later version.

5.7. Backward Compatibility

Root-ACs and Leaf-ACs are used only in cases where PEs support E-Tree and have no impact on VPLS PEs already in operation.

In a case where a common VPLS is composed of both PEs supporting the solution and PEs not supporting it, ACs attached to PEs which don't support E-tree are taken as Root-ACs. The Leaf-to-Leaf communication restriction will be implemented within the scope of the compliant PEs.

6. Compliance with Requirements

The proposed solution in this document meets the requirements mentioned in [Draft VPLS ETree Req] Section 5.

The solution prohibits communication between any two Leaf ACs in a VPLS.

The solution allows multiple Root ACs in a VPLS instance.

The solution allows Root AC and Leaf AC of a VPLS instance co-exist on any PE.

The solution is applicable to BGP-VPLS [RFC4761].

The solution is applicable to Case 1: Single technology "VPLS-only".

7. Security Considerations

This will be added in later version.

8. References

8.1. Normative References


8.2. Informative References

[Draft VPLS ETreereq] Key, et al., Requirements for MEF E-Tree Support in VPLS, draft-key-l2vpn-vpls-etree-reqt-02.txt, October 2010

[Draft Ext VPLS for ETtree] Key, et al., Extension to VPLS for E-Tree, draft-key-l2vpn-vpls-etree-04.txt, October 2010

9. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

Authors’ Addresses

Yuqun Cao
Ruijie Networks
618 Jinshan Road, Fuzhou 350002, China

Email: yuqun.cao@gmail.com