Abstract

MAC Address Withdrawal is a mechanism described in [RFC4762] to remove or unlearn MAC addresses that have been dynamically learned, for faster convergence. Failure of mechanisms that control loop free connectivity (e.g. Split Horizon) among VPLS PE nodes may cause MAC Address Withdrawal messages looping among those nodes, leading to Denial of Service (DoS) or complete failure of control plane in the PE nodes. This document describes a mechanism to detect and prevent loops of MAC Address Withdrawal messages in a VPLS PE node on such failures.
1. Introduction

A method of Virtual Private LAN Service (VPLS) is described in [RFC4762]. A full mesh of pseudowire (PW)s is established between PE devices to construct the VPLS core. The mesh topology provides a LAN segment or broadcast domain that is fully capable of learning and forwarding on Ethernet MAC addresses at the PE devices. Since every PE is now directly connected to every other PE in the core via a PW, the topology forms a loop. A simpler loop breaking rule is used - the "split horizon" rule, whereby a PE does not forward traffic from one PW to another in the same VPLS mesh. Various mechanisms used to enforce split horizon in the VPLS full mesh is out of scope of this document.

For scalability, this VPLS full mesh core configuration can be augmented with additional non-meshed spoke or MTU-s [RFC4762] nodes to provide a Hierarchical VPLS (H-VPLS) service [RFC4762]. To protect from connection failure of the spoke PW or the PE, the MTU-s or the PE is dual-homed into two PE devices in the same VPLS instance [Figure 1]. The dual homed connectivity forms a loop that requires loop breaking mechanism. Various loop breaking mechanisms are out of scope of this document.

MAC Address Withdrawal [RFC4762] is required to remove or unlearn MAC addresses for faster convergence on topology change in H-VPLS (e.g., failure of the primary link for a dual-homed VPLS capable switch). In this document the term "MAC Flush Message" means the LDP Address Withdraw Message with MAC List TLV used for MAC address withdrawal in VPLS.
Failure of mechanisms that control loop free connectivity among VPLS PE nodes may cause MAC Flush messages looping among the nodes. This document describes a mechanism to detect and prevent such loops on MAC Flush messages.

Note that such misconfiguration may cause loops, broadcast storms in traffic forwarding as well. Data path loop detection and prevention is outside the scope of this document. Implementations may deploy methodologies available in native service forwarding to detect and prevent such loops in data path.

2. Conventions used in this document

In examples, "C:" and "S:" indicate lines sent by the client and server respectively.

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 [RFC2119].

3. Terminology

This document uses the terminology defined in [RFC5036] and [RFC4762]. Throughout this document VPLS means the emulated bridged LAN service offered to a customer. H-VPLS means the hierarchical Connectivity or layout of MTU-s and PE devices offering the VPLS [RFC4762]. The terms spoke node and MTU-s in H-VPLS are used interchangeably.
4. Problem Statement

This section describes the problem in detail.

When a MAC flush message is received by a PE device, it is processed locally and propagated to all other PE devices participating in the full mesh. The scope of MAC flush propagation in the full mesh is limited by the same split horizon rules applied for forwarding VPLS service traffic. MAC flush message received from a PE device
participating in full mesh is not forwarded to another PE device in the same full mesh. Due to failure of split horizon control mechanisms or misconfiguration in the VPLS topology, a loop may get created. Similarly, due to failure of loop breaking mechanism in dual-homed topology a loop may get created among the participating PE and/or MTU-s devices. In such a scenario, a MAC flush message may be propagated by a PE or MTU-s to cause loops of MAC flush messages through the control plane. Each such loop of a MAC flush message causes replication to (N-1) messages, where N is number of PE devices the replicating PE is connected to. Such looping of MAC flush messages may lead to denial of service (DoS) attacks or complete failure of the control plane. The control plane in PE or MTU-s devices requires a fault tolerant mechanism to detect such loops and protect against such failures. This document describes a loop detection procedure in MAC Flush in a VPLS.

5. Loop Detection in MAC Flush

MAC Flush Loop Detection is a configurable option in a VPLS capable PE or MTU-s device that provides a mechanism for finding and preventing MAC Flush messages from looping across the VPLS network.

The mechanism makes use of LDP Path Vector TLVs defined in [RFC 5036]. For loop detection in MAC Flush, Path Vector TLVs are carried in the MAC Flush messages.

The following shows the encoding of Path Vector TLV when used for MAC Flush Loop Detection. For backward compatibility the U bit and F bit MUST be set as 1.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|1|1| Path Vector (0x0104) |        Length                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|                            LSR Id 1                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|                                                               |
~                                                               ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|                            LSR Id n                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

Figure 2. Path Vector TLV in MAC Flush.
The method builds on the following basic property of the TLV:

- A Path Vector TLV contains a list of the PEs that the containing MAC flush message has traversed. A PE is identified in a Path Vector list by its unique LDP LSR Identifier, which is the first four octets of its LDP Identifier. When a PE propagates a MAC flush message containing a Path Vector, it appends its LSR Id to the Path Vector list. An LSR that receives a message with a Path Vector that contains its LSR Id, detects that the message has traversed a loop.

The following paragraphs describe the MAC Flush Loop Detection procedures. For these paragraphs, and only these paragraphs, "MUST" is redefined to mean "MUST if configured for Loop Detection". Further the term "PE device" is used specify a VPLS capable PE or a MTU-s device.

5.1 Loop Detection Procedure in MAC Flush Message

The rules that govern use of the Path Vector TLV are LDP MAC Flush messages by a PE when MAC Flush Loop Detection is enabled are the following:

- If PE device is originating the MAC Flush Message, it MUST include a Path Vector TLV of length 1 containing its own LSR Id.

- If PE device is propagating the MAC flush as a result of having received a MAC Flush from an upstream PE, then if the MAC Flush contains a Path Vector TLV:

  PE device MUST add its own LSR Id to the Path Vector, and MUST pass the resulting Path Vector to its downstream PE along with the MAC flush message propagated. If the MAC flush message contains no Path Vector TLV, then PE MUST include a Path Vector TLV of length 1 containing its own LSR Id.

- If PE device receives a MAC Flush Message from another PE device with a Path Vector TLV containing its LSR Id or which exceeds the maximum allowable length, then PE device detects that the MAC flush message has traveled in a loop. By the definition of Path Vector TLV in [RFC5036], supports the notion of a maximum allowable Path Vector Length; a PE that detects a Path Vector has reached the maximum length behaves as if containing message has traversed a loop. Such limit MAY be a locally configurable option in an implementation based on
the scope of H-VPLS topology. The limit MAY also be driven by payload size limitations of MAC flush messages.

- When PE device detects a loop, it MUST drop the MAC flush message without processing and MUST NOT propagate the message further.

6. Applicability

If MAC Flush Loop Detection is desired in VPLS Network, then it should be turned on in all PE devices within that VPLS Network, else Loop Detection will not operate properly and may result in undetected loops or in falsely detected loops.

PE devices that are configured for MAC Flush Loop Detection are not expected to store the Path Vectors as part of the VPLS service.

7. Security Considerations

- Control plane aspects
  - LDP security (authentication) methods as described in [RFC5036] is applicable here. Further this document implements security considerations as in [RFC4447] and [RFC4762].

- Data plane aspects
  - This specification does not have any impact on the VPLS forwarding plane.

8. IANA Considerations

This document does not require any IANA consideration.

9. References

9.1. Normative References


9.2. Informative References


10. Acknowledgments

The authors would like acknowledge the comments and suggestions from Wim Henderickx, Vach Kompella, Florin Balus, Mustapha Aissaoui, Mathew Bocci and Miroslav Vrana.

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

Authors’ Addresses

Paul Kwok
Alcatel-Lucent
701 E Middlefield Road,
Mountain View, CA 94043
USA.

Email: paul.kwok@alcatel-lucent.com

Pranjal Kumar Dutta
Alcatel-Lucent
701 E Middlefield Road,
Mountain View, CA 94043
USA.

Email: pranjal.dutta@alcatel-lucent.com
Copyright Notice

Copyright (c) 2010 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.