Network Working Group
Internet-Draft
Updates: 4761 (if approved)
Intended status: Standards Track
Expires: February 6, 2012

B. Kothari
Cisco Systems
K. Kompella
Juniper Networks
W. Henderickx
F. Balus
Alcatel-Lucent
J. Uttaro
AT&T
July 6, 2011

BGP based Multi-homing in Virtual Private LAN Service
draft-ietf-l2vpn-vpls-multihoming-03.txt

Abstract

Virtual Private LAN Service (VPLS) is a Layer 2 Virtual Private Network (VPN) that gives its customers the appearance that their sites are connected via a Local Area Network (LAN). It is often required for the Service Provider (SP) to give the customer redundant connectivity to some sites, often called "multi-homing". This memo shows how BGP-based multi-homing can be offered in the context of LDP and BGP VPLS solutions.

Table of Contents

1. Introduction

Virtual Private LAN Service (VPLS) is a Layer 2 Virtual Private Network (VPN) that gives its customers the appearance that their sites are connected via a Local Area Network (LAN). It is often required for a Service Provider (SP) to give the customer redundant connectivity to one or more sites, often called "multi-homing". [RFC4761] explains how VPLS can be offered using BGP for auto-discovery and signaling; section 3.5 of that document describes how multi-homing can be achieved in this context. [RFC6074] explains how VPLS can be offered using BGP for auto- discovery, (BGP-AD) and [RFC4762] explains how VPLS can be offered using LDP for signaling. This document provides a BGP-based multi-homing solution applicable to both BGP and LDP VPLS technologies. Note that BGP MH can be used for LDP VPLS without the use of the BGP- AD solution.

Section 2 lays out some of the scenarios for multi-homing, other ways that this can be achieved, and some of the expectations of BGP-based multi-homing. Section 3 defines the components of BGP-based multi-homing, and the procedures required to achieve this. Section 7 may someday discuss security considerations.

1.1. General Terminology

Some general terminology is defined here; most is from [RFC4761], [RFC4762] or [RFC4364]. Terminology specific to this memo is introduced as needed in later sections.

A "Customer Edge" (CE) device, typically located on customer premises, connects to a "Provider Edge" (PE) device, which is owned and operated by the SP. A "Provider" (P) device is also owned and operated by the SP, but has no direct customer connections. A "VPLS Edge" (VE) device is a PE that offers VPLS services.

A VPLS domain represents a bridging domain per customer. A Route Target community as described in [RFC4360] is typically used to identify all the PE routers participating in a particular VPLS domain. A VPLS site is a grouping of ports on a PE that belong to the same VPLS domain. A Multi-homed (MH) site is uniquely identified by a MH site ID (MH-ID). Sites are referred to as local or remote depending on whether they are configured on the PE router in context or on one of the remote PE routers (network peers).

1.2. Conventions

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].

2. Background

This section describes various scenarios where multi-homing may be required, and the implications thereof. It also describes some of the singular properties of VPLS multi-homing, and what that means from both an operational point of view and an implementation point of view. There are other approaches for providing multi-homing such as Spanning Tree Protocol, and this document specifies use of BGP for multi-homing. Comprehensive comparison among the approaches is outside the scope of this document.

2.1. Scenarios

CE1 is a VPLS CE that is dual-homed to both PE1 and PE2 for redundant connectivity.

```
                             ...............
                            .               .    ___ CE2
                      ___ PE1                .  /
                     /    :                  PE3
                  __/    :       Service      :
              CE1 __     :       Provider    PE4
                    \     :                   : \___ CE3
                     \___ PE2                .
                            .               .
                             ...............
```

Figure 1: Scenario 1

CE1 is a VPLS CE that is dual-homed to both PE1 and PE2 for redundant connectivity. However, CE4, which is also in the same VPLS domain, is single-homed to just PE1.

```
              CE4 -------    ...............
                         \  .               .    ___ CE2
                      ___ PE1                .  /
                     /    :                  PE3
                  __/    :       Service      :
              CE1 __     :       Provider    PE4
                    \     :                   : \___ CE3
                     \___ PE2                .
                            .               .
                             ...............
```

Figure 2: Scenario 2

2.2. VPLS Multi-homing Considerations

The first (perhaps obvious) fact about a multi-homed VPLS CE, such as CE1 in Figure 1 is that if CE1 is an Ethernet switch or bridge, a loop has been created in the customer VPLS. This is a dangerous situation for an Ethernet network, and the loop must be broken. Even if CE1 is a router, it will get duplicates every time a packet is flooded, which is clearly undesirable.

The next is that (unlike the case of IP-based multi-homing) only one of PE1 and PE2 can be actively sending traffic, either towards CE1 or into the SP cloud. That is to say, load balancing techniques will not work. All other PEs MUST choose the same designated forwarder for a multi-homed site. Call the PE that is chosen to send traffic to/from CE1 the "designated forwarder".

In Figure 2, CE1 and CE4 must be dealt with independently, since CE1 is dual-homed, but CE4 is not.

3. Multi-homing Operation

This section describes procedures for electing a designated forwarder among the set of PEs that are multi-homed to a customer site. The procedures described in this section are applicable to BGP based VPLS, LDP based VPLS with BGP-AD or a VPLS that contains a mix of both BGP and LDP signaled PWs.

3.1. Provisioning Model

Figure 1 shows a customer site, CE1, multi-homed to two VPLS PEs, PE1 and PE2. In order for all VPLS PEs within the same VPLS domain to elect one of the multi-homed PEs as the designated forwarder, an indicator that the PEs are multi-homed to the same customer site is required. This is achieved by assigning the same multi-homed site ID (MH-ID) on PE1 and PE2 for CE1. When remote VPLS PEs receive NLRI advertisement from PE1 and PE2 for CE1, the two NLRI advertisements for CE1 are identified as candidates for designated forwarder selection due to the same MH-ID. Thus, same MH-ID SHOULD be assigned on all VPLS PEs that are multi-homed to the same customer site.

Note that a MH-ID=0 is invalid and a PE should discard such an advertisement.

3.2. Multi-homing NLRI

Section 3.2.2 in [RFC4761] describes the encoding of the BGP VPLS NLRI. This NLRI contains fields VE-ID, VE block offset, VE block size and label base. For multi-homing operation, the same NLRI is used for identifying the multi-homed customers sites. The VE-ID field in the NLRI is set to MH-ID; the VE block offset, VE block size and label base are set to zero. Thus, the NLRI contains 2 octets indicating the length, 8 octets for Route Distinguisher, 2 octets for MH-ID and 7 octets with value zero.

Figure 2 shows two customer sites, CE1 and CE4, connected to PE1 with CE1 multi-homed to PE1 and PE2. CE4 does not require special addressing, being associated with the base VPLS instance identified by the VSI-ID for LDP VPLS and VE-ID for BGP VPLS. However, CE1 which is multi-homed to PE1 and PE2 requires configuration of MH-ID and both PE1 and PE2 MUST be provisioned with the same MH-ID for CE1.

It is valid to have non-zero VE block offset, VE block size and label base in the VPLS NLRI for a multi-homed site. However, multi-homing operations in such a case are outside the scope of this document.

3.3. Designated Forwarder Election

BGP-based multi-homing for VPLS relies on BGP DF election and VPLS DF election. The net result of doing both BGP and VPLS DF election is that of electing a single designated forwarder (DF) among the set of PEs to which a customer site is multi-homed. All the PEs that are elected as non-designated forwarders MUST keep their attachment circuit to the multi-homed CE in blocked status (no forwarding).

These election algorithms operate on VPLS advertisements, which include both the NLRI and attached BGP attributes. In order to simplify the explanation of these algorithms, we will use a number of variables derived from fields in the VPLS advertisement. These variables are: RD, MH-ID, VBO, DOM, ACS, PREF and PE-ID. The notation ADV -> <RD, MH-ID, VBO, DOM, ACS, PREF, PE-ID> means that from a received VPLS advertisement ADV, the respective variables were derived. The following sections describe two attributes needed for DF election, then describe the variables and how they are derived from fields in VPLS advertisement ADV, and finally describe how DF election is done.

3.3.1. Attributes

The procedures below refer to two attributes: the Route Origin community (see Section 4.1) and the L2-info community (see Section 4.2). These attributes are required for inter-AS operation; for generality, the procedures below show how they are to be used. The procedures also say how to handle the case that either or both are not present.

3.3.2. Variables Used

3.3.2.1. RD

RD is simply set to the Route Distinguisher field in the NLRI part of ADV.

3.3.2.2. MH-ID

MH-ID is simply set to the VE-ID field in the NLRI part of ADV.

3.3.2.3. VBO

VBO is simply set to the VE Block Offset field in the NLRI part of ADV. This field will typically be zero.

3.3.2.4. DOM

This variable, indicating the VPLS domain to which ADV belongs, is derived by applying BGP policy to the Route Target extended communities in ADV. The details of how this is done are outside the scope of this document.

3.3.2.5. ACS

ACS is the status of the attachment circuits for a given site of a VPLS. ACS = 1 if all attachment circuits for the site are down, and 0 otherwise.

For BGP-based Multi-homing, ADV MUST contain an L2-info extended community; within this community are control flags. One of these flags is the 'D' bit, described in [I-D.kothari-l2vpn-auto-site-id]. ACS is set to the value of the 'D' bit in ADV.

3.3.2.6. PREF

PREF is derived from the Local Preference (LP) attribute in ADV as well as the VPLS Preference field (VP) in the L2-info extended community. If the Local Preference attribute is missing, LP is set to 0; if the L2-info community is missing, VP is set to 0. The following table shows how PREF is computed from LP and VP.

VP Value	LP Value	PREF Value	Comment
0	0	0	malformed advertisement, unless ACS=1
0	1 to (2^16-1)	LP	backwards compatibility
0	2^16 to (2^32-1)	(2^16-1)	backwards compatibility
>0	LP same as VP	VP	Implementation supports VP
>0	LP != VP	0	malformed advertisement

Table 1

3.3.2.7. PE-ID

If ADV contains a Route Origin (RO) community (see Section 4.1) with type 0x01, then PE-ID is set to the Global Administrator sub-field of the RO. Otherwise, if ADV has an ORIGINATOR_ID attribute, then PE-ID is set to the ORIGINATOR_ID. Otherwise, PE-ID is set to the BGP Identifier.

3.3.3. Election Procedures

The election procedures described in this section apply equally to BGP VPLS and LDP VPLS.

Election occurs in two stages. The first stage divides all received VPLS advertisements into buckets of relevant and comparable advertisements. Distinction MUST NOT be made on whether the NLRI is a multi-homing NLRI or not. In this stage, advertisements may be discarded as not being relevant to DF election. The second stage picks a single "winner" from each bucket by repeatedly applying a tie-breaking algorithm on a pair of advertisements from that bucket. The tie-breaking rules are such that the order in which advertisements are picked from the bucket does not affect the final result. Note that this is a conceptual description of the process; an implementation MAY choose to realize this differently as long as the semantics are preserved.

Note: these procedures supersede the tie breaking rules described in (Section 9.1.2.2) [RFC4271]

3.3.3.1. Bucketization for BGP DF Election

An advertisement

```
ADV -> <RD, MH-ID, VBO, ACS, PREF, PE-ID>
```

is put into the bucket for <RD, MH-ID, VBO>. In other words, the information in BGP DF election consists of <RD, MH-ID, VBO> and only advertisements with exact same <RD, MH-ID, VBO> are candidates for DF election.

3.3.3.2. Bucketization for VPLS DF Election

An advertisement

```
ADV -> <RD, MH-ID, VBO, DOM, ACS, PREF, PE-ID>
```

is discarded if DOM is not of interest to the VPLS PE. Otherwise, ADV is put into the bucket for <DOM, MH-ID>. In other words, all

advertisements for a particular VPLS domain that have the same MH-ID are candidates for VPLS DF election.

3.3.3.3. Tie-breaking Rules

This section describes the tie-breaking rules for both BGP and VPLS DF election. Tie-breaking rules for BGP DF election are applied to candidate advertisements by any BGP speaker. Since RD must be same for advertisements to be candidates for BGP DF election, use of unique RDs will result in no candidate advertisements for BGP tie-breaking rules and thus, a BGP speaker in such a case will simply not do BGP DF election. Tie-breaking rules for VPLS DF election are applied to candidate advertisements by all VPLS PEs and the actions taken by VPLS PEs based on the VPLS DF election result are described in Section 3.4.

Given two advertisements ADV1 and ADV2 from a given bucket, first compute the variables needed for DF election:

```
ADV1 -> <RD1, MH-ID1, VBO1, DOM1, ACS1, PREF1, PE-ID1>
ADV2 -> <RD2, MH-ID2, VBO2, DOM2, ACS2, PREF2, PE-ID2>
```

Note that MH-ID1 = MH-ID2 and DOM1 = DOM2, since ADV1 and ADV2 came from the same bucket. If this is for BGP DF election, RD1 = RD2 and VBO1 = VBO2 as well. Then the following tie-breaking rules MUST be applied in the given order.

```
1.  if (ACS1 != 1) AND (ACS2 == 1) ADV1 wins; stop
    if (ACS1 == 1) AND (ACS2 != 1) ADV2 wins; stop
    else continue

2.  if (PREF1 > PREF2) ADV1 wins; stop;
    else if (PREF1 < PREF2) ADV2 wins; stop;
    else continue

3.  if (PE-ID1 < PE-ID2) ADV1 wins; stop;
    else if (PE-ID1 > PE-ID2) ADV2 wins; stop;
    else ADV1 and ADV2 are from the same VPLS PE
```

For BGP DF election, if there is no winner and ADV1 and ADV2 are from the same PE, BGP DF election should simply consider this as an update.

For VPLS DF election, if there is no winner and ADV1 and ADV2 are from the same PE, a VPLS PE MUST retain both ADV1 and ADV2.

3.4. DF Election on PEs

DF election algorithm MUST be run by all multi-homed VPLS PEs. In addition, all other PEs SHOULD also run the DF election algorithm. As a result of the DF election, multi-homed PEs that lose the DF election for a MH-ID MUST put the ACs associated with the MH-ID in non-forwarding state.

DF election result on the egress PEs can be used in traffic forwarding decision. Figure 2 shows two customer sites, CE1 and CE4, connected to PE1 with CE1 multi-homed to PE1 and PE2. If PE1 is the designated forwarder for CE1, based on the DF election result, PE3 can chose to not send unknown unicast and multicast traffic to PE2 as PE2 is not the designated forwarder for any customer site and it has no other single homed sites connected to it.

4. Multi-AS VPLS

This section describes multi-homing in an inter-AS context.

4.1. Route Origin Extended Community

Due to lack of information about the PEs that originate the VPLS NLRIs in inter-AS operations, Route Origin Extended Community [RFC4360] is used to carry the source PE's IP address.

To use Route Origin Extended Community for carrying the originator VPLS PE's loopback address, the type field of the community MUST be set to 0x01 and the Global Administrator sub-field MUST be set to the PE's loopback IP address.

4.2. VPLS Preference

When multiple PEs are assigned the same site ID for multi-homing, it is often desired to be able to control the selection of a particular PE as the designated forwarder. Section 3.5 in [RFC4761] describes the use of BGP Local Preference in path selection to choose a particular NLRI, where Local Preference indicates the degree of preference for a particular VE. The use of Local Preference is inadequate when VPLS PEs are spread across multiple ASes as Local Preference is not carried across AS boundary. A new field, VPLS preference (VP), is introduced in this document that can be used to accomplish this. VPLS preference indicates a degree of preference for a particular customer site. VPLS preference is not mandatory for intra-AS operation; the algorithm explained in Section 3.3 will work with or without the presence of VPLS preference.

Section 3.2.4 in [RFC4761] describes the Layer2 Info Extended Community that carries control information about the pseudowires. The last two octets that were reserved now carries VPLS preference as shown in Figure 3.

```
+------------------------------------+
| Extended community type (2 octets) |
+------------------------------------+
|  Encaps Type (1 octet)             |
+------------------------------------+
|  Control Flags (1 octet)           |
+------------------------------------+
|  Layer-2 MTU (2 octet)             |
+------------------------------------+
|  VPLS Preference (2 octets)        |
+------------------------------------+
```

Figure 3: Layer2 Info Extended Community

A VPLS preference is a 2-octets unsigned integer. A value of zero indicates absence of a VP and is not a valid preference value. This interpretation is required for backwards compatibility. Implementations using Layer2 Info Extended Community as described in (Section 3.2.4) [RFC4761] MUST set the last two octets as zero since it was a reserved field.

For backwards compatibility, if VPLS preference is used, then BGP Local Preference MUST be set to the value of VPLS preference. Note that a Local Preference value of zero for a MH-ID is not valid unless 'D' bit in the control flags is set (see [I-D.kothari-l2vpn-auto-site-id]). In addition, Local Preference value greater than or equal to 2^16 for VPLS advertisements is not valid.

4.3. Use of BGP-MH attributes in Inter-AS Methods

Section 3.4 in [RFC4761] and section 4 in [RFC6074] describe three methods (a, b and c) to connect sites in a VPLS to PEs that are across multiple AS. Since VPLS advertisements in method (a) do not cross AS boundaries, multi-homing operations for method (a) remain exactly the same as they are within as AS. However, for method (b) and (c), VPLS advertisements do cross AS boundary. This section describes the VPLS operations for method (b) and method (c). Consider Figure 4 for inter-AS VPLS with multi-homed customer sites.

4.3.1. Inter-AS Method (b): EBGP Redistribution of VPLS Information between ASBRs

```
                            AS1                    AS2
                           ........               ........
            CE2 ________  .        .             .        .
                    ___ PE1         .           .          PE3 --- CE3
                   /    :            .         .            :
                __/    :             :         :             :
            CE1 __     :           ASBR1 --- ASBR2           :
                  \     :            :         :            :
                   \___ PE2         .           .          PE4 ---- CE4
                          .        .             .         .
                           ........                ........
```

Figure 4: Inter-AS VPLS

A customer has four sites, CE1, CE2, CE3 and CE4. CE1 is multi-homed to PE1 and PE2 in AS1. CE2 is single-homed to PE1. CE3 and CE4 are also single homed to PE3 and PE4 respectively in AS2. Assume that in addition to the base LDP/BGP VPLS addressing (VSI-IDs/VE-IDs), MH ID 1 is assigned for CE1. After running DF election algorithm, all four VPLS PEs must elect the same designated forwarder for CE1 site. Since BGP Local Preference is not carried across AS boundary, VPLS preference as described in Section 4.2 MUST be used for carrying site preference in inter-AS VPLS operations.

For Inter-AS method (b) ASBR1 will send a VPLS NLRI received from PE1 to ASBR2 with itself as the BGP nexthop. ASBR2 will send the received NLRI from ASBR1 to PE3 and PE4 with itself as the BGP nexthop. Since VPLS PEs use BGP Local Preference in DF election, for backwards compatibility, ASBR2 MUST set the Local Preference value in the VPLS advertisements it sends to PE3 and PE4 to the VPLS preference value contained in the VPLS advertisement it receives from ASBR1. ASBR1 MUST do the same for the NLRIs it sends to PE1 and PE2. If ASBR1 receives a VPLS advertisement without a valid VPLS preference from a PE within its AS, then ASBR1 MUST set the VPLS preference in the advertisements to the Local Preference value before sending it to ASBR2. Similarly, ASBR2 must do the same for advertisements without VPLS Preference it receives from PEs within its AS. Thus, in method (b), ASBRs MUST update the VPLS and Local Preference based on the advertisements they receive either from an ASBR or a PE within their AS.

In Figure 4, PE1 will send the VPLS advertisements with Route Origin Extended Community containing its loopback address. PE2 will do the same. Even though PE3 receives the VPLS advertisements for VE-ID 1 and 2 from the same BGP nexthop, ASBR2, the source PE address contained in the Route Origin Extended Community is different for the CE1 and CE2 advertisements, and thus, PE3 creates two PWs, one for CE1 (for VE-ID 1) and another one for CE2 (for VE-ID 2).

4.3.2. Inter-AS Method (c): Multi-Hop EBGP Redistribution of VPLS Information between ASes

In this method, there is a multi-hop E-BGP peering between the PEs or Route Reflectors in AS1 and the PEs or Route Reflectors in AS2. There is no VPLS state in either control or data plane on the ASBRs. The multi-homing operations on the PEs in this method are exactly the same as they are in intra-AS scenario. However, since Local Preference is not carried across AS boundary, the translation of LP to VP and vice versa MUST be done by RR, if RR is used to reflect VPLS advertisements to other ASes. This is exactly the same as what

a ASBR does in case of method (b). A RR must set the VP to the LP value in an advertisement before sending it to other ASes and must set the LP to the VP value in an advertisement that it receives from other ASes before sending to the PEs within the AS.

5. MAC Flush Operations

In a service provider VPLS network, customer MAC learning is confined to PE devices and any intermediate nodes, such as a Route Reflector, do not have any state for MAC addresses.

Topology changes either in the service provider's network or in customer's network can result in the movement of MAC addresses from one PE device to another. Such events can result into traffic being dropped due to stale state of MAC addresses on the PE devices. Age out timers that clear the stale state will resume the traffic forwarding, but age out timers are typically in minutes, and convergence of the order of minutes can severely impact customer's service. To handle such events and expedite convergence of traffic, flushing of affected MAC addresses is highly desirable.

This section describes the scenarios where VPLS flush is desirable and the specific VPLS Flush TLVs that provide capability to flush the affected MAC addresses on the PE devices. All operations described in this section are in context of a particular VPLS domain and not across multiple VPLS domains. Mechanisms for MAC flush are described in [I-D.kothari-l2vpn-vpls-flush] for BGP based VPLS and in [RFC4762] for LDP based VPLS.

5.1. MAC List FLush

If multiple customer sites are connected to the same PE, PE1 as shown in Figure 2, and redundancy per site is desired when multi-homing procedures described in this document are in effect, then it is desirable to flush just the relevant MAC addresses from a particular site when the site connectivity is lost.

To flush particular set of MAC addresses, a PE SHOULD originate a flush message with MAC list that contains a list of MAC addresses that needs to be flushed. In Figure 2, if connectivity between CE1 and PE1 goes down and if PE1 was the designated forwarder for CE1, PE1 MAY send a list of MAC addresses that belong to CE1 to all its BGP peers.

It is RECOMMENDED that in case of excessive link flap of customer attachment circuit in a short duration, a PE should have a means to throttle advertisements of flush messages so that excessive flooding of such advertisements do not occur.

5.2. Implicit MAC Flush

Implicit MAC Flush refers to the use of BGP MH advertisements by the PEs to flush the MAC addresses learned from the previous designated

forwarder.

In case of a failure, when connectivity to a customer site is lost, remote PEs learn that a particular site is no longer reachable. The local PE either withdraws the VPLS NLRI that it previously advertised for the site or it sends a BGP update message for the site's VPLS NLRI with the 'D' bit set. In such cases, the remote PEs can flush all the MACs that were learned from the PE which reported the failure.

However, in cases when a designated forwarder change occurs in absence of failures, such as when an attachment circuit comes up, the BGP MH advertisement from the PE reporting the change is not sufficient for MAC flush procedures. Consider the case in Figure 2 where PE1-CE1 link is non-operational and PE2 is the designated forwarder for CE1. Also assume that Local Preference of PE1 is higher than PE2. When PE1-CE1 link becomes operational, PE1 will send a BGP MH advertisement to all it's peers. If PE3 elects PE1 as the new designated forwarder for CE1 and as a result flushes all the MACs learned from PE1 before PE2 elects itself as the non-designated forwarder, there is a chance that PE3 might learn MAC addresses from PE2 and as a result may black-hole traffic until those MAC addresses are deleted due to age out timers.

A new flag 'F' is introduced in the Control Flags Bit Vector as a deterministic way to indicate when to flush.

Control Flags Bit Vector

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|D|A|F|Z|Z|Z|C|S| (Z = MUST Be Zero)
+-+-+-+-+-+-+-+-+
```

Figure 5

A designatd forwarder must set the F bit and a non-designated forwarder must clear the F bit when sending BGP MH advertisements. A state transition from one to zero for the F bit can be used by a remote PE to flush all the MACs learned from the PE that is transitioning from designated forwarder to non-designated forwarder.

5.3. Minimizing the effects of fast link transitions

Certain failure scenarios may result in fast transitions of the link towards the multi-homing CE which in turn will generate fast status transitions of one or multiple multi-homed sites reflected through multiple BGP MH advertisements and LDP MAC Flush messages.

It is recommended that a timer to damp the link flaps be used for the port towards the multi-homed CE to minimize the number of MAC Flush events in the remote PEs and the occurrences of BGP state compressions for F bit transitions. A timer value more than the time it takes BGP to converge in the network is recommended.

6. Backwards Compatibility

No forwarding loops are formed when PEs or Route Reflectors that do not support procedures defined in this section co exist in the network with PEs or Route Reflectors that do support.

6.1. BGP based VPLS

As explained in this section, multi-homed PEs to the same customer site MUST assign the same MH-ID and related NLRI SHOULD contain the block offset, block size and label base as zero. Remote PEs that lack support of multi-homing operations specified in this document will fail to create any PWs for the multi-homed MH-IDs due to the label value of zero and thus, the multi-homing NLRI should have no impact on the operation of Remote PEs that lack support of multi-homing operations specified in this document.

6.2. LDP VPLS with BGP Auto-discovery

The BGP-AD NLRI has a prefix length of 12 containing only a 8 bytes RD and a 4 bytes VSI-ID. If a LDP VPLS PEs running BGP AD lacks support of multi-homing operations specified in this document, it SHOULD ignore a MH NLRI with the length field of 17. As a result it will not ask LDP to create any PWs for the multi-homed Site-ID and thus, the multi-homing NLRI should have no impact on LDP VPLS operation. MH PEs may use existing LDP MAC Flush to flush the remote LDP VPLS PEs or may use the implicit MAC Flush procedure.

7. Security Considerations

No new security issues are introduced beyond those that are described in [RFC4761] and [RFC4762].

8. IANA Considerations

At this time, this memo includes no request to IANA.

9. Acknowledgments

The authors would like to thank Yakov Rekhter, Nischal Sheth, and Mitali Singh for their insightful comments and probing questions.

10. References

10.1. Normative References

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

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

[RFC6074] Rosen, E., "Provisioning, Autodiscovery, and Signaling in L2VPNs", RFC 6074, January 2011.

10.2. Informative References

[I-D.kothari-l2vpn-vpls-flush]
Kothari, B. and R. Fernando, "VPLS Flush in BGP-based Virtual Private LAN Service", draft-kothari-l2vpn-vpls-flush-00 (work in progress), October 2008.

[I-D.kothari-l2vpn-auto-site-id]
Kothari, B., Kompella, K., and T. IV, "Automatic Generation of Site IDs for Virtual Private LAN Service", draft-kothari-l2vpn-auto-site-id-01 (work in progress), October 2008.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, February 2006.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, April 2006.

[RFC4762] Lasserre, M. and V. Kompella, "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

Authors' Addresses

Bhupesh Kothari
Cisco Systems
3750 Cisco Way
San Jose, CA 95134, US
Email: bhupesh@cisco.com

Kireeti Kompella
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089 US
Email: kireeti@juniper.net

Wim Henderickx
Alcatel-Lucent
Email: wim.henderickx@alcatel-lucent.be

Florin Balus
Alcatel-Lucent
Email: florin.balus@alcatel-lucent.com

James Uttaro
AT&T
200 S. Laurel Avenue
Middletown, NJ 07748, US
Email: uttaro@att.com