L3VPN Working Group   Eric C. Rosen (Editor)
Internet Draft   IJsbrand Wijnands
Intended Status: Proposed Standard   Cisco Systems, Inc.
Expires: December 28, 2012

                                                        Yiqun Cai
                                                        Microsoft

                                                       Arjen Boers

                                                    June 28, 2012

                   MVPN: Using Bidirectional P-Tunnels

                   draft-ietf-l3vpn-mvpn-bidir-02.txt

Abstract

   The documents specifying multicast support for BGP/MPLS IP VPNs allow
   customer multicast data to be transported through a service
   provider's network through a set multicast tunnels. Such tunnels are
   advertised by BGP in a BGP attribute known as the "Provider Multicast
   Service Interface (PMSI) Tunnel Attribute". The base specifications
   allow the PMSI Tunnel Attribute to advertise bidirectional multicast
   distribution trees as "PMSI Tunnels"; however, those documents do not
   provide all the necessary details for using those tunnels. These
   details are provided in this document. This document also specifies
   the procedures for assigning customer multicast flows to specific
   bidirectional PMSI tunnels.

Status of this Memo

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt.

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

Copyright and License Notice

Table of Contents

1. Introduction

   The documents specifying multicast support for BGP/MPLS IP VPNs allow
   customer multicast data to be transported through a service
   provider's network through a set multicast tunnels. Such tunnels are
   advertised by BGP in a BGP attribute known as the "Provider Multicast
   Service Interface (PMSI) Tunnel Attribute". The base specifications
   allow the PMSI Tunnel Attribute to advertise bidirectional multicast
   distribution trees as "PMSI Tunnels"; however, those documents do not
   provide all the necessary details for using those tunnels. These
   details are provided in this document.

1.1. Terminology

This document uses terminology from [MVPN] and, in particular, uses
the prefixes "C-" and "P-", as specified in Section 3.1 of [MVPN], to
distinguish addresses in the "customer address space" from addresses
in the "provider address space". The following terminology and
acronyms are particularly important in this document:

  - MVPN

    Multicast Virtual Private Network -- a VPN [L3VPN] in which
    multicast service is offered.

  - VRF

    VPN Routing and Forwarding table [L3VPN].

  - PE

    A Provider Edge router, as defined in [L3VPN].

  - LSP

    An MPLS Label Switched Path.

  - MP2MP

    Multipoint-to-multipoint.

  - P-tunnel

    A tunnel through the network of one or more Service Providers
    (SPs).

  - C-S

    Multicast Source. A multicast source address, in the address
    space of a customer network.

  - C-G

    Multicast Group. A multicast group address (destination address)
    in the address space of a customer network.

  - C-multicast flow or C-flow

    A customer multicast flow. Each C-flow is identified by the
    ordered pair (source address, group address), where each address
    is in the customer's address space. The identifier of a
    particular C-flow is usually written as (C-S,C-G).

  - RP

    A "Rendezvous Point", as defined in [PIM].

  - C-RP

    A Rendezvous Point whose address is in the customer's address
    space.

  - RPA

    A "Rendezvous Point Address", as defined in [BIDIR-PIM].

  - C-RPA

    An RPA in the customer's address space.

  - P-RPA

    An RPD in the Service Provider's address space

  - Selective P-tunnel

    A P-tunnel that is joined only by Provider Edge (PE) routers that
    need to receive one or more of the C-flows that are traveling
    through that P-tunnel.

  - Inclusive P-tunnel

    A P-tunnel that is joined by all PE routers that attach to sites
    of a given MVPN.

  - Intra-AS I-PMSI A-D route

    Intra Autonomous System Inclusive Provider Multicast Service
    Interface Auto-Discovery route. Carried in BGP Update messages,
    these routes can be used to advertise the use of Inclusive
    P-tunnels.

   - S-PMSI A-D route

     Selective Provider Multicast Service Interface Auto-Discovery
     route. Carried in BGP Update messages, these routes are used to
     advertise the fact that particular C-flows are bound to (i.e.,
     are traveling through) particular P-tunnels.

   - PE Distinguisher Labels

     These are upstream-assigned MPLS labels that can be used, on a
     MP2MP LSP, to enable the receiver of a data packet to infer the
     identity of the PE router that transmitted the packet onto the
     LSP.

   - PE Distinguisher Labels Attribute

     A BGP path attribute, defined in [MVPN-BGP], that is used for
     advertising the use of PE Distinguisher Labels.

   We say that the NLRI ("Network Layer Reachability Information") of a
   BGP S-PMSI A-D route or Source Active A-D route contains (C-S,C-G) if
   its "Multicast Source" field contains C-S and its "Multicast Group"
   field contains C-G. If either or both of these fields is encoded as
   a wildcard, we will say that the NLRI contains (C-*,C-*) (both fields
   encoded as wildcard), (C-*,C-G) (multicast source field encoded as
   wildcard) or (C-S,C-*) (multicast group field encoded as wildcard).

   Familiarity with multicast concepts and terminology [PIM] is also
   presupposed.

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document, when appearing in all caps, are to be interpreted as
   described in [RFC2119].

1.2. Overview

   The base documents for MVPN, [MVPN] and [MVPN-BGP], define a "PMSI
   Tunnel Attribute" (PTA) that may be carried in the BGP "I-PMSI A-D
   routes" and BGP "S-PMSI A-D routes" that are defined therein. The
   base documents define the way that bidirectional P-tunnels are
   identified in the PTA, and the way in which the identifier of a
   bidirectional P-tunnel is encoded in the PTA.

   However, those documents do not contain the full set of
   specifications governing the use of the PTA to advertise
   bidirectional P-tunnels; rather, those documents declare these

specifications to be "out of scope." Similarly, the use of
bidirectional P-tunnels advertised in S-PMSI A-D routes with
wildcards is declared by [MVPN-WILDCARDS] to be "out of scope." This
document provides the necessary specifications to allow the use of
bidirectional P-tunnels.

This document also specifies the procedures for assigning customer
multicast flows to specific bidirectional PMSI tunnels.

Two kinds of bidirectional P-tunnel are discussed in this document:

  - Multicast distribution trees that are created through the use of
    BIDIR-PIM [BIDIR-PIM].

  - Multipoint-to-multipoint Label Switched Paths (MP2MP LSPs),
    created by Label Distribution Protocol (LDP)
    Multipoint-to-Multipoint extensions [mLDP].

Other possible kinds of bidirectional P-tunnels are outside the scope
of this document.

This document also specifies three methods of using bidirectional
P-tunnels:

  - Partitioned Method without LSP Hierarchy.

    In this method, when a PE advertises a bidirectional P-tunnel in
    the PTA of an S-PMSI A-D route, the PE must be the "root node" of
    the tunnel. There are a specific set of rules for using tunnels
    of this sort, specified in section 4.2 of this document. This
    method is discussed in [MVPN] Section 11.2.3, where it is called
    "Partial Mesh of MP2MP P-tunnels". This method can be used with
    MP2MP LSPs or with BIDIR-PIM P-tunnels. It does not require the
    use of upstream-assigned labels, and does not use the PE
    Distinguisher Labels attribute.

    When a packet is received from a P-tunnel, the PE that receives
    it can infer the identity of the P-tunnel from the MPLS label
    that has risen to the top of the packet's label stack. Once the
    P-tunnel is known, the root node of the P-tunnel is also known.
    In the "Partitioned Method without LSP Hierarchy", the root node
    of the P-tunnel on which the packet arrived is treated as the
    "distinguished PE" for that packet.

    If the received packet is part of a unidirectional C-flow, its
    "distinguished PE" is the PE that transmitted the packet onto the
    P-tunnel. If the packet is part of a bidirectional C-flow, its
    "distinguished PE" is not necessarily the PE that transmitted it,

      but rather the transmitter's "upstream PE" for the C-RPA of the
      bidirectional C-group. See section 4.2.

   - Partitioned Method with LSP Hierarchy.

      This method is only applicable when MP2MP LSPs are being used as
      the P-tunnels. In this method, a PE advertising a bidirectional
      P-tunnel in the PTA of an S-PMSI A-D route does not need to be
      the root of the P-tunnel. However, each P-tunnel MUST be
      advertised by its root, and the root MUST include a PE
      Distinguisher Labels attribute.

      This method is discussed in [MVPN], section 11.2.2. The detailed
      specification is provided in Section 4.3 of this document. This
      method provides the same functionality as the "Partitioned Method
      without LSP Hierarchy", but requires the use of upstream-assigned
      MPLS labels, which are not necessarily supported by all
      platforms. The upstream-assigned labels are used to provide an
      LSP hierarchy, in which an "outer" MP2MP LSP carries multiple
      "inner" MP2MP LSPs. P routers only maintain state for the outer
      MP2MP LSP.

      As in the "Partitioned Method without LSP Hierarchy", when a
      packet is received from a P-tunnel, the PE that receives it can
      infer the identity of the P-tunnel from the MPLS label that has
      risen to the top of the packet's label stack. However, the
      packet's "distinguished PE" is not necessarily the root node of
      the P-tunnel. Rather, the identity of the packet's distinguished
      PE is inferred from the PE Distinguisher Label further down in
      the label stack. (See [MVPN] Section 12.3.)

   - Unpartitioned Method.

      This method can be used with MP2MP LSPs or with BIDIR-PIM
      P-tunnels. If used with MP2MP LSPs, it can be used along with
      the PE Distinguisher Labels attribute. However, in this case the
      PE Distinguisher Label carried by a packet always corresponds to
      the PE that transmitted the packet onto the tunnel.

   This document does not specify any new data encapsulations for
   bidirectional P-tunnels. Section 12 of [MVPN] applies unchanged.

2. **Advertising and Creating Bidirectional P-Tunnels**

   A bidirectional P-tunnel may be advertised in the PTA of an Intra-AS
   I-PMSI A-D route or in the PTA of an S-PMSI A-D route. The
   advertisement of a bidirectional P-tunnel in the PTA of an Inter-AS
   I-PMSI A-D route is outside the scope of this document.

2.1. **BIDIR-PIM P-Tunnels**

   Each BIDIR-PIM P-Tunnel is identified by a unique P-group address
   [MVPN, section 3.1]. (The P-group address is called a "P-Multicast
   Group" in [MVPN-BGP]). Section 5 of [MVPN-BGP] specifies the way to
   identify a particular BIDIR-PIM P-tunnel in the PTA of an I-PMSI or
   S-PMSI A-D route.

   Ordinary BIDIR-PIM procedures are used to set up the BIDIR-PIM P-
   tunnels. A BIDIR-PIM P-group address is always associated with a
   unique "Rendezvous Point Address" (RPA) in the SP's address space.
   We will refer to this as the "P-RPA". Every PE needing to join a
   particular BIDIR-PIM P-tunnel must be able to determine the P-RPA
   that corresponds to the P-tunnel's P-group address. To construct the
   P-tunnel, PIM Join/Prune messages are sent along the path from the PE
   to the P-RPA. Any P routers along that path must also be able to
   determine the P-RPA, so that they too can send PIM Join/Prune
   messages towards it. The method of mapping a P-group address to an
   RPA may be static configuration, or some automated means of RPA
   discovery that is outside the scope of this specification.

   If a BIDIR-PIM P-tunnel is used to instantiate an I-PMSI or an
   S-PMSI, it is RECOMMENDED that the path from each PE in the tunnel to
   the RPA consist entirely of point-to-point links. On a
   point-to-point link, there is no ambiguity in determining which
   router is upstream towards a particular RPA, so the BIDIR-PIM
   "Designated Forwarder Election" is very quick and simple. Use of a
   BIDIR-PIM P-tunnel containing multiaccess links is possible, but
   considerably more complex.

   When the PTA of an Intra-AS I-PMSI A-D route or an S-PMSI A-D route
   identifies a BIDIR-PIM tunnel, the route SHOULD NOT have a PE
   Distinguisher Labels attribute. If it does, that attribute MUST be
   ignored.

   For a given BIDIR-PIM P-tunnel, the PE router or P router that is
   closest to the P-RPA is considered to be the "root node" of the
   tunnel.

2.2. **MP2MP LSPs**

   Each MP2MP LSP is identified by a unique "MP2MP FEC (Forwarding
   Equivalence Class) element" [mLDP]. The FEC element contains the IP
   address of the "root node", followed by an "opaque value" that
   identifies the MP2MP LSP uniquely in the context of the root node's
   IP address. This opaque value may be configured or autogenerated,
   and within an MVPN, there is no need for different root nodes to use
   the same opaque value. The mLDP specification supports the use of
   several different ways of constructing the tunnel identifiers. The
   current specification does not place any restriction on the type of
   tunnel identifier that might be used. However, a given
   implementation might not support every possible type of tunnel
   identifier.

   Section 5 of [MVPN-BGP] specifies the way to identify a particular
   MP2MP P-tunnel in the PTA of an I-PMSI or S-PMSI A-D route.

   Ordinary mLDP procedures for MP2MP LSPs are used to set up the MP2MP
   LSP.

3. **The All BIDIR-PIM Wild Card**

   When an MVPN customer is using BIDIR-PIM, it is useful to be able to
   advertise an S-PMSI A-D route whose semantics are: "by default, all
   BIDIR-PIM C-multicast traffic (within a given VPN) that has not been
   bound to any other P-tunnel is bound to the bidirectional P-tunnel
   identified by the PTA of this route". This can be especially useful
   if one is using a bidirectional P-tunnel to carry the C-BIDIR flows,
   while using unidirectional P-tunnels to carry other flows. To do
   this we, need to have a way to express a (C-*,C-*) wildcard that is
   restricted to BIDIR-PIM C-groups.

   We therefore define a special value of the group wildcard, whose
   meaning is "all BIDIR-PIM groups". The "BIDIR-PIM groups wildcard"
   is encoded as a group field whose length is 8 bits and whose value is
   zero. That is, the "multicast group length" field contains the value
   0x08, and the "multicast group" field is a single octet containing
   the value 0x00. We will use the notation (C-*,C-BIDIR) to refer to
   the "all BIDIR-PIM groups" wildcard.

4. Methods of Using Bidirectional P-Tunnels

   There are two different methods of using BIDIR-PIM P-tunnels, the
   "Partitioned Method" and the "Unpartitioned Method".

   If a bidirectional P-tunnel is used to instantiate an I-PMSI, the
   Unpartitioned Method MUST be used.

   If a bidirectional P-tunnel is used to instantiate an S-PMSI
   (including the case of a (C-*,C-*) S-PMSI), either the Partitioned
   Method or the Unpartitioned Method may be used. The method by a
   given VRF used is determined by provisioning. It SHOULD be possible
   to provision this on a per-MVPN basis, but all the VRFs of a single
   MVPN MUST be provisioned to use the same method.

4.1. Unpartitioned Method

   This section applies when and only when a bidirectional P-tunnel is
   used to instantiate a PMSI using the Unpartitioned Method.

   When instantiating an I-PMSI with a bidirectional P-tunnel, any VRF
   in a given MVPN that originates an Intra-AS I-PMSI A-D route must
   include a PTA with that route. All such PTAs MUST identify the same
   P-tunnel. (Any scenario in which they do not advertise the same
   P-tunnel in their Intra-AS I-PMSI A-D routes is outside the scope of
   this document.) The identity of this P-tunnel is known by
   provisioning.

   When instantiating a (C-*,C-*) S-PMSI with a bidirectional P-tunnel,
   any VRF in a given MVPN that originates an S-PMSI A-D route
   containing (C-*,C-*) must include a PTA with that route. All such
   PTAs MUST identify the same P-tunnel. (Any scenario in which they do
   not advertise the same MP2MP LSP in their (C-*,C-*) S-PMSI A-D routes
   is outside the scope of this document.) The identity of this
   P-tunnel is known by provisioning.

   When instantiating S-PMSIs with bidirectional P-tunnels, different
   S-PMSI A-D routes that do not contain (C-*,C-*), originated by the
   same or by different PEs, MAY have PTAs that identify the same
   bidirectional tunnel, and they MAY have PTAs that do not identify the
   same bidirectional tunnel.

   An I-PMSI or S-PMSI A-D route whose PTA identifies a bidirectional
   P-tunnel does not need to be originated by the root node of the
   tunnel. In fact, the root node does not even need to be a PE router.

   The Unpartitioned Method SHOULD NOT be used for instantiating an

S-PMSI to which one or more C-BIDIR flows are bound, as it cannot be used to support the "Partitioned Set of PEs" method discussed in [MVPN] section 11.2 and [RFC6517] section 3.6.

An I-PMSI or S-PMSI A-D route whose PTA identifies an MP2MP LSP MAY include the PE Distinguisher Labels attribute. Note that the procedures of [MVPN] Section 9.1.1 are not applicable if the PE Distinguisher Labels attribute is not included.

4.1.1. When an S-PMSI is a 'Match for Transmission'

Given the need for a PE to transmit multicast data packets of a particular customer C-flow, [MVPN-WILDCARDS] Section 3.1 gives a four-step algorithm for determining the S-PMSI A-D route, if any, that "matches" that C-flow for transmission. When referring to that section, please recall that BIDIR groups are also "Any Source Multicast" (ASM) groups.

When bidirectional P-tunnels are used in the Unpartitioned Method, the same algorithm applies, with one modification, when the PTA of an S-PMSI A-D route identifies a bidirectional P-tunnel. One additional step is added to the algorithm. This new step occurs before the fourth step of the algorithm, and is as follows:

  - Otherwise, if there is an S-PMSI A-D route currently originated by PE1, whose NLRI contains (C-*,C-BIDIR), and if C-G is a BIDIR group, the (C-S,C-G) C-flow matches that route.

4.1.2. When an S-PMSI is a 'Match for Reception'

Given the need for a PE to receive multicast data packets of a particular customer C-flow, [MVPN-WILDCARDS] Section 3.2 specifies the procedures for determining the S-PMSI A-D route, if any, that advertised the P-tunnel on which the PE should expect to receive that C-flow.

When bidirectional P-tunnels are used in the Unpartitioned Method, the same procedures apply, with one modification.

The last paragraph of Section 3.2.2 of [MVPN-WILDCARDS] begins:

    "If (C-*,C-G) does not match a (C-*,C-G) S-PMSI A-D route from PE2, but PE1 has an installed (C-*,C-*) S-PMSI A-D route from PE2, then (C-*,C-G) matches the (C-*,C-*) route if one of the following conditions holds:"

This is changed to:

> "If (C-*,C-G) does not match a (C-*,C-G) S-PMSI A-D route from
> PE2, but C-G is a BIDIR group and PE1 has an installed
> (C-*,C-BIDIR) S-PMSI A-D route, then (C-*,C-G) matches that
> route. Otherwise, if PE1 has an installed (C-*,C-*) S-PMSI A-D
> route from PE2, then (C-*,C-G) matches the (C-*,C-*) route if one
> of the following conditions holds:"

4.2. Partitioned Method without LSP Hierarchy

This section applies when and only when the Partitioned Method
without LSP Hierarchy is used to instantiate a PMSI. Whether a
particular VPN uses this method is known by provisioning. Whether a
particular VPN uses MP2MP LSPs or whether it uses BIDIR-PIM trees for
its P-tunnels is also known by provisioning.

The Partitioned Method without LSP Hierarchy MUST NOT be used to
instantiate an I-PMSI; it is only used to instantiate S-PMSIs. It
may however be used to instantiate a (C-*,C-*) S-PMSI or a
(C-*,C-BIDIR) S-PMSI.

When the Partitioned Method without LSP Hierarchy is used to
instantiate a (C-*,C-*) S-PMSI, a (C-*,C-BIDIR) S-PMSI, or a
(C-*,C-G) S-PMSI where C-G is a BIDIR group, each of a "selected set"
of PEs in a given MVPN MUST originate an S-PMSI A-D route with a PTA
identifying a bidirectional P-tunnel. The PE originating the route
MUST be the root node of the identified bidirectional P-tunnel. It
follows that two different PEs may not advertise the same
bidirectional P-tunnel.

If BIDIR-PIM P-tunnels are used, each P-tunnel MUST have a distinct
P-group address. If MP2MP LSPs are used, each P-tunnel MUST have
have a distinct MP2MP FEC (i.e., distinct combination of "root node"
and "opaque value").

A PE is considered to be in the "selected set" if at least one of the
following conditions hold:

  - The "Partitioned Sets of PEs" method of supporting C-BIDIR
    traffic is being used, and the PE's route to the Customer's
    Rendezvous Point Address (C-RPA) for one or more C-BIDIR groups
    is via a VRF interface.

   - The "Partitioned Sets of PEs" method of supporting C-BIDIR
     traffic is being used, it is desired to transmit some or all of
     the customer's unidirectional multicast traffic (for the given
     MVPN) on the same LSPs used for carrying C-BIDIR traffic, and the
     PE has customer multicast traffic to transmit to other PEs.

   There may be other conditions under which a PE is considered to be in
   the "selected set"; these are outside the scope of this document.

   If any VRF of a given MVPN uses this method when instantiating an
   S-PMSI with a bidirectional P-tunnel, then all VRFs of that MVPN MUST
   use this method.

   The PE Distinguisher Label attribute SHOULD NOT be included in a BGP
   S-PMSI A-D route when this method is being used; if included it must
   be ignored.

   When the Partitioned Method without LSP Hierarchy is used to
   instantiate an S-PMSI, it may be used to implement the "Partitioned
   Sets of PEs" method of supporting C-BIDIR, as discussed in section
   11.2 of [MVPN] and section 3.6 of [RFC6517]. A C-BIDIR flow MUST be
   carried only on a (C-*,C-G), (C-*,C-BIDIR), or (C-*,C-*) S-PMSI. A
   PE MUST NOT originate a (C-S,C-G) S-PMSI A-D route for any C-G that
   is a C-BIDIR group.

   When a BGP A-D route's PTA specifies a BIDIR-PIM P-tunnel, the PE
   Distinguisher Labels attribute SHOULD NOT be included; if it is
   included, it MUST be ignored.

4.2.1. When an S-PMSI is a 'Match for Transmission'

   Given the need for a PE, say PE1, to transmit multicast data packets
   of a particular C-flow, [MVPN-WILDCARDS] Section 3.1 gives a four-
   step algorithm for determining the S-PMSI A-D route, if any, that
   "matches" that C-flow for transmission.

   If the C-flow is not a BIDIR-PIM C-flow, these rules apply unchanged.
   If the C-flow is a BIDIR-PIM C-flow, the rules as applied by a
   particular PE, say PE1, are given below:

     - If the C-RPA for C-G is a C-address of PE1, or if PE1's route to
       the C-RPA is via a VRF interface, then:

         * if there is an S-PMSI A-D route, currently originated by PE1,
           whose NLRI contains (C-*,C-G) and whose PTA identifies a
           bidirectional P-tunnel, then the C-flow matches that route

   * otherwise, if there is an S-PMSI A-D route, currently
     originated by PE1, whose NLRI contains (C-*,C-BIDIR) and
     whose PTA identifies a bidirectional P-tunnel, then the
     C-flow matches that route

   * otherwise, if there is an S-PMSI A-D route, currently
     originated by PE1, whose NLRI contains (C-*,C-*) and whose
     PTA identifies a bidirectional P-tunnel, then the C-flow
     matches that route

  - If PE1 determines the upstream PE for C-G's C-RPA to be some
    other PE, say PE2, then the following rules apply:

   * if there is an installed S-PMSI A-D route, originated by PE2,
     whose NLRI contains (C-*,C-G) and whose PTA identifies a
     bidirectional P-tunnel, then the C-flow matches that route

   * otherwise, if there is an installed S-PMSI A-D route,
     originated by PE2, whose NLRI contains (C-*,C-BIDIR) and
     whose PTA identifies a bidirectional P-tunnel, then the
     C-flow matches that route

   * otherwise, if there is an S-PMSI A-D route, currently
     originated by PE2, whose NLRI contains (C-*,C-*) and whose
     PTA identifies a bidirectional P-tunnel, then the C-flow
     matches that route

   PE1 MUST transmit the C-flow on the P-tunnel identified in the PTA of
   the matching S-PMSI A-D route.

[4.2.2](4.2.2). When an S-PMSI is a 'Match for Reception'

   Given the need for a PE to receive multicast data packets of a
   particular C-flow, [MVPN-WILDCARDS] Section 3.2 specifies procedures
   for determining the S-PMSI A-D route, if any, that "matches" that
   C-flow for reception. Those rules apply unchanged for C-flows that
   are not BIDIR-PIM C-flows.

   For BIDIR-PIM C-flows, the rules of [MVPN-WILDCARDS] Section 3.2.1 do
   not apply.

   The rules of [MVPN-WILDCARDS] Section 3.2.2 are replaced by the
   following rules.

   Suppose that a PE router (call it PE1) needs to receive (C-*,C-G)
   traffic, where C-G is a C-BIDIR group. Suppose also that PE1 has
   determined that PE2 is the "upstream PE" [MVPN] for the C-RPA of C-G.

Then:

   - if PE1 has an installed S-PMSI A-D route originated by PE2, whose
     NLRI contains (C-*,C-G), then (C-*,C-G) matches this route.

   - otherwise, if PE1 has an installed (C-*,C-BIDIR) route from PE2,
     then (C-*,C-G) matches this route.

   - otherwise, if PE1 has an installed (C-*,C-*) S-PMSI A-D route
     from PE2, then (C-*,C-G) matches this route.

If a customer multicast data packet addressed to C-G is received on a
P-tunnel that was not advertised in an S-PMSI A-D route matching
(C-*,C-G), the packet MUST be discarded.

4.3. Partitioned Method with LSP Hierarchy

This section applies when and only when the Partitioned Method with
LSP Hierarchy is used to instantiate a PMSI. Whether a particular
VPN uses this method is known by provisioning. The Partitioned Method
with LSP Hierarchy is only used with MP2MP LSPs, and is not defined
for BIDIR-PIM P-tunnels.

The Partitioned Method with LSP Hierarchy MUST NOT be used to
instantiate an I-PMSI; it is only used to instantiate S-PMSIs. It
may however be used to instantiate a (C-*,C-*) S-PMSI or a
(C-*,C-BIDIR) S-PMSI.

When the Partitioned Method with hierarchy is used to instantiate a
(C-*,C-*) S-PMSI, a (C-*,C-BIDIR) S-PMSI, or a (C-*,C-G) S-PMSI where
C-G is a BIDIR group, each of a "selected set" of PEs in a given MVPN
MUST originate an S-PMSI A-D route with a PTA identifying a
bidirectional P-tunnel. The PE originating the route is not
necessarily the root node of the identified bidirectional P-tunnel;
multiple PEs may advertise the same bidirectional P-tunnel. However,
the root node of the P-tunnel MUST be a PE and MUST advertise that
P-tunnel in an S-PMSI A-D route. Further, whenever the root node of
the P-tunnel advertises the P-tunnel in an S-PMSI A-D route, the root
mode must include a PE Distinguishers Label attribute, created as
specified in [MVPN] Section 11.2.2

A PE is considered to be in the "selected set" if the "Partitioned
Sets of PEs" method of supporting C-BIDIR traffic is being used, and
the PE is provisioned to originate a (C-*,C-*) or (C-*,C-BIDIR)
S-PMSI A-D route, and to use an MP2MP LSP to instantiate that S-PMSI.

In addition, a PE, say PE1, that desires to transmit multicast data

packets of a unidirectional C-flow on a MP2MP LSP MUST originate an
S-PMSI A-D route with an NLRI matching the C-flow (according to the
specification of [MVPN-WILDCARDS] Section 3.1). PE1 need not be the
root node of the MP2MP LSP, but if it is not, the same LSP MUST have
been advertised in the PTA of an S-PMSI A-D route originated by its
root node, and the root node MUST include a PE Distinguisher Labels
attribute that assigns a label to the IP address of PE1.

If any VRF of a given MVPN uses this method when instantiating an
S-PMSI with a bidirectional P-tunnel, all VRFs of that MVPN must use
this method.

When the Partitioned Method with LSP Hierarchy is used to instantiate
an S-PMSI, it may be used to implement the "Partitioned Sets of PEs"
method of supporting C-BIDIR, as discussed in section 11.2 of [MVPN]
and section 3.6 of [RFC6517]. A C-BIDIR flow MUST be carried only on
a (C-*,C-G), (C-*,C-BIDIR), or (C-*,C-*) S-PMSI. A PE MUST NOT
originate a (C-S,C-G) S-PMSI A-D route for any C-G that is a C-BIDIR
group.

4.3.1. When an S-PMSI is a 'Match for Transmission'

Given the need for a PE, say PE1, to transmit multicast data packets
of a particular C-flow, [MVPN-WILDCARDS] Section 3.1 gives a four-
step algorithm for determining the S-PMSI A-D route, if any, that
"matches" that C-flow for transmission.

If the C-flow is not a BIDIR-PIM C-flow, these rules apply unchanged.
Once PE1 finds the matching S-PMSI (if any) is found, PE1 may
transmit a packet of that C-flow on the P-tunnel advertised in that
route. The packet MUST carry the PE Distinguisher Label assigned by
the root node of that P-tunnel to the IP address of PE1.

If the C-flow is a BIDIR-PIM C-flow, the rules are given below.

Assume PE1 determines that the upstream PE for C-G's C-RPA is PE2.

   - If there is an installed S-PMSI A-D route, or an S-PMSI A-D route
     originated by PE1 itself, whose NLRI contains (C-*,C-G) and whose
     PTA identifies a bidirectional P-tunnel, then the C-flow matches
     that route

   - otherwise, if there is an installed S-PMSI A-D route, or an
     S-PMSI A-D route currently originated by PE1 itself, whose NLRI
     contains (C-*,C-BIDIR) and whose PTA identifies a bidirectional
     P-tunnel, then the C-flow matches that route

     - otherwise, if there is an installed S-PMSI A-D route (or an
       S-PMSI A-D route currently originated by PE1 itself) whose NLRI
       contains (C-*,C-*) and whose PTA identifies a bidirectional
       P-tunnel, then the C-flow matches that route

   PE1 MUST transmit the C-flow on the P-tunnel identified in the PTA of
   the matching S-PMSI A-D route. In constructing the packet's MPLS
   label stack, it must used the PE Distinguisher Label that was
   assigned by the P-tunnel's root node to the IP address of PE2.
   (Note: the PE Distinguisher Label is the one assigned to the address
   of PE2, not the one assigned to the address of PE1.)

4.3.2. When an S-PMSI is a 'Match for Reception'

   Given the need for a PE, say PE1, to receive multicast data packets
   of a particular C-flow, [MVPN-WILDCARDS] Section 3.2 specifies
   procedures for determining the S-PMSI A-D route, if any, that
   "matches" that C-flow for reception. Those rules require that the
   matching S-PMSI A-D route has been originated by the upstream PE for
   the C-flow. These rules are modified in this section, as follows.

   Consider a particular C-flow. Suppose either:

     - the C-flow is unidirectional, and PE1 determines that its
       upstream PE is PE2, or

     - the C-flow is bidirectional, and PE1 determines that the upstream
       PE for its C-RPA is PE2.

   Then the C-flow may match an installed S-PMSI A-D route that was not
   originated by PE2, as long as:

     1. the PTA of that A-D route identifies an MP2MP LSP, and

     2. there is an installed S-PMSI A-D route originated the root node
        of that LSP, or PE1 itself the root node of the LSP and there
        is a currently originated S-PMSI A-D route from PE1 whose PTA
        identifies that LSP, and

     3. the latter S-PMSI A-D route (the one identified in 2 just
        above) contains a PE Distinguisher Labels attribute that
        assigned an MPLS label to the IP address of PE2.

   However, a bidirectional C-flow never matches an S-PMSI A-D route
   whose NLRI contains (C-S,C-G).

   If a multicast data packet is received over a matching P-tunnel, but

does not carry the value of the PE Distinguisher Label that has been
assigned to the upstream PE for its C-flow, then the packet MUST be
discarded.

5. IANA Considerations

This document has no actions for IANA.

6. Security Considerations

There are no additional security considerations beyond those of
[MVPN] and [MVPN-BGP], or any that may apply to the particular
protocol used to set up the bidirectional tunnels ([BIDIR-PIM],
[mLDP]).

7. Acknowledgments

The authors wish to thank Karthik Subramanian, Rajesh Sharma, and
Apoorva Karan for their input. We also thank Yakov Rekhter for his
valuable critique.

8. Authors' Addresses

Arjen Boers
E-mail: arjen@boers.com

Yiqun Cai
Microsoft
1065 La Avenida
Mountain View, CA 94043
E-mail: yiqunc@microsoft.com

Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA, 01719
E-mail: erosen@cisco.com

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a Diegem 1831
Belgium
E-mail: ice@cisco.com

9. Normative References

[BIDIR-PIM] "Bidirectional Protocol Independent Multicast", Handley,
Kouvelas, Speakman, Vicisano, RFC 5015, October 2007

[L3VPN], "BGP/MPLS IP Virtual Private Networks", Rosen, Rekhter
(editors), RFC 4364, February 2006

[mLDP] "Label Distribution Protocol Extensions for
Point-to-Multipoint and Multipoint-to-Multipoint Label Switched
Paths", Wijnands, Minei, Kompella, Thomas, RFC 6388, November 2011

[MVPN] "Multicast in MPLS/BGP IP VPNs", Rosen, Aggarwal, et. al., RFC
6513, February 2012

[MVPN-BGP] "BGP Encodings and Procedures for Multicast in MPLS/BGP IP
VPNs", Aggarwal, Rosen, Morin, Rekhter, RFC 6514, February 2012

[MVPN-WILDCARDS] "Wild Cards in Multicast VPN Auto-Discovery Routes",
Rosen, Rekhter, Hendrickx, Qiu, RFC 6625, May 2012

[PIM] "Protocol Independent Multicast - Sparse Mode (PIM-SM):
Protocol Specification (Revised)", Fenner, Handley, Holbrook,
Kouvelas, RFC 4601, August 2006

[RFC2119] "Key words for use in RFCs to Indicate Requirement
Levels.", Bradner, March 1997

10. Informative References

[RFC6517] "Mandatory Features in a Layer 3 Multicast BGP/MPLS VPN
Solution", Morin, Niven-Jenkins, Kamite, Zhang, Leymann, Bitar, RFC
6517, February 2012