                   Multicast Virtual Private Network (MVPN):
                        Using Bidirectional P-Tunnels

Abstract

   A set of prior RFCs specify procedures for supporting multicast in
   BGP/MPLS IP VPNs. These procedures allow customer multicast data to
   travel across a service provider's backbone network through a set of
   multicast tunnels. The tunnels are advertised in certain BGP
   multicast auto-discovery routes, by means of a BGP attribute known
   as the "Provider Multicast Service Interface (PMSI) Tunnel"
   attribute. Encodings have been defined that allow the PMSI Tunnel
   attribute to identify bidirectional (multipoint-to-multipoint)
   multicast distribution trees. However, the prior RFCs do not provide
   all the necessary procedures for using bidirectional tunnels to
   support multicast VPNs. This document updates RFCs 6513, 6514, and
   6625 by specifying those procedures. In particular, it specifies the
   procedures for assigning customer multicast flows (unidirectional or
   bidirectional) to specific bidirectional tunnels in the provider
   backbone, for advertising such assignments, and for determining which
   flows have been assigned to which tunnels.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by
   the Internet Engineering Steering Group (IESG). Further
   information on Internet Standards is available in Section 2 of
   RFC 5741.

   Information about the current status of this document, any
   errata, and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc7582.

Table of Contents

1. Introduction

   The RFCs that specify multicast support for BGP/MPLS IP VPNs
   ([RFC6513], [RFC6514], and [RFC6625]) allow customer multicast data
   to be transported across a service provider's network though a set of
   multicast tunnels. These tunnels are advertised in BGP multicast
   auto-discovery (A-D) routes, by means of a BGP attribute known as the
   "Provider Multicast Service Interface (PMSI) Tunnel" attribute. The
   base specifications allow the use of bidirectional (multipoint-to-
   multipoint) multicast distribution trees and describe how to encode
   the identifiers for bidirectional trees into the PMSI Tunnel
   attribute. However, those specifications do not provide all the
   necessary detailed procedures for using bidirectional tunnels; the
   full specification of these procedures was considered to be outside
   the scope of those documents. The purpose of this document is to
   provide all the necessary procedures for using bidirectional trees in
   a service provider's network to carry the multicast data of VPN
   customers.

1.1. Terminology

   This document uses terminology from [RFC6513] and, in particular,
   uses the prefixes "C-" and "P-", as specified in Section 3.1 of
   [RFC6513], 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:

   o MVPN

      Multicast Virtual Private Network -- a VPN [RFC4364] in which
      multicast service is offered.

   o VRF

      VPN Routing and Forwarding table [RFC4364].

   o PE

      A Provider Edge router, as defined in [RFC4364].

   o SP

      Service Provider.

   o LSP

      An MPLS Label Switched Path.

   o P2MP

      Point-to-Multipoint.

   o MP2MP

      Multipoint-to-multipoint.

   o Unidirectional

      Adjective for a multicast distribution tree in which all traffic
      travels downstream from the root of the tree. Traffic can enter a
      unidirectional tree only at the root. A P2MP LSP is one type of
      unidirectional tree. Multicast distribution trees set up by
      Protocol Independent Multicast - Sparse Mode (PIM-SM) [RFC4601]
      are also unidirectional trees. Data traffic traveling along a
      unidirectional multicast distribution tree is sometimes referred
      to in this document as "unidirectional traffic".

   o Bidirectional

      Adjective for a multicast distribution tree in which traffic may
      travel both upstream (towards the root) and downstream (away from
      the root). Traffic may enter a bidirectional tree at any node.
      An MP2MP LSP is one type of bidirectional tree. Multicast
      distribution trees created by Bidirectional Protocol Independent
      Multicast (BIDIR-PIM) [RFC5015] are also bidirectional trees.

      Data traffic traveling along a bidirectional multicast
      distribution tree is sometimes referred to in this document as
      "bidirectional traffic".

   o P-tunnel

      A tunnel through the network of one or more SPs. In this
      document, the P-tunnels we speak of are instantiated as
      bidirectional multicast distribution trees.

   o SSM

      Source-Specific Multicast.   When SSM is being used, a multicast
      distribution tree carries traffic from only a single source.

   o ASM

      Any Source Multicast. When ASM is being used, some multicast
      distribution trees ("share trees") carry traffic from multiple
      sources.

   o C-S

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

   o C-G

      Multicast Group. A multicast group address (destination address)
      in the address space of a customer network. When used without
      qualification, "C-G" may refer to either a unidirectional group
      address or a bidirectional group address.

   o C-G-BIDIR

      A bidirectional multicast group address (i.e., a group address
      whose IP multicast distribution tree is built by BIDIR-PIM).

   o C-multicast flow or C-flow

      A customer multicast flow. A C-flow travels through VPN customer
      sites on a multicast distribution tree set up by the customer.
      These trees may be unidirectional or bidirectional, depending upon
      the multicast routing protocol used by the customer. A C-flow
      travels between VPN customer sites by traveling through P-tunnels.

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

      If a customer uses the ASM model, then some or all of the
      customer's C-flows may be traveling along the same "shared tree".
      In this case, we will speak of a "(C-*,C-G)" flow to refer to a
      set of C-flows that travel along the same shared tree in the
      customer sites.

   o C-BIDIR flow or bidirectional C-flow

      A C-flow that, in the VPN customer sites, travels along a
      bidirectional multicast distribution tree. The term "C-BIDIR
      flow" indicates that the customer's bidirectional tree has been
      set up by BIDIR-PIM.

   o RP

      A Rendezvous Point, as defined in [RFC4601].

   o C-RP

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

   o RPA

      A Rendezvous Point Address, as defined in [RFC5015].

   o C-RPA

      An RPA in the customer's address space.

   o P-RPA

      An RPA in the SP's address space.

   o Selective P-tunnel

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

   o Inclusive P-tunnel

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

   o PMSI

      Provider Multicast Service Interface. A PMSI is a conceptual
      overlay on a Service Provider backbone, allowing a PE in a given
      MVPN to multicast to other PEs in the MVPN. PMSIs are
      instantiated by P-tunnels.

   o I-PMSI

      Inclusive PMSI. Traffic multicast by a PE on an I-PMSI is
      received by all other PEs in the MVPN. I-PMSIs are instantiated
      by Inclusive P-tunnels.

   o S-PMSI

      Selective PMSI. Traffic multicast by a PE on an S-PMSI is
      received by some (but not necessarily all) of the other PEs in the
      MVPN. S-PMSIs are instantiated by Selective P-tunnels.

o Intra-AS I-PMSI A-D route

   Intra-AS (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. See [RFC6514], Section 4.1.

o 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 a particular C-flow or a particular set of
   C-flows is bound to (i.e., is traveling through) a particular
   P-tunnel. See [RFC6514], Section 4.3.

o (C-S,C-G) S-PMSI A-D route

   An S-PMSI A-D route whose NLRI (Network Layer Reachability
   Information) contains C-S in its "Multicast Source" field and C-G
   in its "Multicast Group" field.

o (C-*,C-G) S-PMSI A-D route

   An S-PMSI A-D route whose NLRI contains the wildcard (C-*) in its
   "Multicast Source" field and C-G in its "Multicast Group" field.
   See [RFC6625].

o (C-*,C-G-BIDIR) S-PMSI A-D route

   An S-PMSI A-D route whose NLRI contains the wildcard (C-*) in its
   "Multicast Source" field and C-G-BIDIR in its "Multicast Group"
   field. See [RFC6625].

o (C-*,C-*) S-PMSI A-D route

   An S-PMSI A-D route whose NLRI contains the wildcard C-* in its
   "Multicast Source" field and the wildcard C-* in its "Multicast
   Group" field. See [RFC6625].

o (C-*,C-*-BIDIR) S-PMSI A-D route

   An S-PMSI A-D route whose NLRI contains the wildcard C-* in its
   "Multicast Source" field and the wildcard "C-*-BIDIR" in its
   "Multicast Group" field. See Section 2 of this document.

   o (C-S,C-*) S-PMSI A-D route

      An S-PMSI A-D route whose NLRI contains C-S in its "Multicast
      Source" field and the wildcard C-* in its "Multicast Group" field.
      See [RFC6625].

   o Wildcard S-PMSI A-D route

      A (C-*,C-G) S-PMSI A-D route, a (C-*,C-*) S-PMSI A-D route, a
      (C-S,C-*) S-PMSI A-D route, or a (C-*,C-*-BIDIR) S-PMSI A-D route.

   o PTA

      PMSI Tunnel attribute, a BGP attribute that identifies a P-tunnel.
      See [RFC6514], Section 8.

   The terminology used for categorizing S-PMSI A-D routes will also be
   used for categorizing the S-PMSIs advertised by those routes. For
   example, the S-PMSI advertised by a (C-*,C-G) S-PMSI A-D route will
   be known as a "(C-*,C-G) S-PMSI".

   Familiarity with multicast concepts and terminology [RFC4601] is also
   presupposed.

   This specification uses the terms "match for transmission" and "match
   for reception" as they are defined in [RFC6625]. When it is clear
   from the context whether we are talking of transmission or reception,
   we will sometimes talk simply of a C-flow "matching" an I-PMSI or
   S-PMSI A-D route.

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

1.2. Overview

   The base documents for MVPN ([RFC6513] and [RFC6514]) define a "PMSI
   Tunnel attribute" (PTA). This is a BGP Path attribute that may be
   attached to the BGP "I-PMSI A-D routes" and "S-PMSI A-D routes" that
   are defined in those documents. The base documents define the way in
   which the identifier of a bidirectional P-tunnel is to be encoded in
   the PTA. However, those documents do not contain the full set of
   specifications governing the use of bidirectional P-tunnels; rather,
   those documents declare the full set of specifications for using
   bidirectional P-tunnels to be outside their scope. Similarly, the

use of bidirectional P-tunnels advertised in wildcard S-PMSI A-D
routes is declared by [RFC6625] to be "outside the scope" of that
document.

This document provides the specifications governing the use of
bidirectional P-tunnels to provide MVPN support. This includes the
procedures for assigning C-flows to specific bidirectional P-tunnels,
for advertising the fact that a particular C-flow has been assigned
to a particular bidirectional P-tunnel, and for determining the
bidirectional P-tunnel on which a given C-flow may be expected.

The C-flows carried on bidirectional P-tunnels may, themselves, be
either unidirectional or bidirectional. Procedures are provided for
both cases.

This document does not specify any new data encapsulations for
bidirectional P-tunnels. Section 12 ("Encapsulations") of [RFC6513]
applies unchanged.

With regard to the procedures for using bidirectional P-tunnels to
instantiate PMSIs, if there is any conflict between the procedures
specified in this document and the procedures of [RFC6513],
[RFC6514], or [RFC6625], the procedures of this document take
precedence.

The use of bidirectional P-tunnels to support extranets [MVPN-XNET]
is outside the scope of this document. The use of bidirectional
P-tunnels as "segmented P-tunnels" (see Section 8 of [RFC6513] and
various sections of [RFC6514]) is also outside the scope of this
document.

1.2.1. Bidirectional P-Tunnel Technologies

This document supports two different technologies for creating and
maintaining bidirectional P-tunnels:

o Multipoint-to-multipoint Label Switched Paths (MP2MP LSPs) that
   are created through the use of the Label Distribution Protocol
   (LDP) Multipoint-to-Multipoint extensions [RFC6388].

o Multicast distribution trees that are created through the use of
   BIDIR-PIM [RFC5015].

Other bidirectional tunnel technologies are outside the scope of this
document.

1.2.2. Reasons for Using Bidirectional P-Tunnels

   Bidirectional P-tunnels can be used to instantiate I-PMSIs and/or
   S-PMSIs.

   An SP may decide to use bidirectional P-tunnels to instantiate
   certain I-PMSIs and/or S-PMSIs in order to provide its customers with
   C-BIDIR support, using the "Partitioned Set of PEs" technique
   discussed in Section 11.2 of [RFC6513] and Section 3.6 of [RFC6517].
   This technique can be used whether the C-BIDIR flows are being
   carried on an I-PMSI or an S-PMSI.

   Even if an SP does not need to provide C-BIDIR support, it may still
   decide to use bidirectional P-tunnels, in order to save state in the
   network's transit nodes. For example, if an MVPN has n PEs attached
   to sites with multicast sources, and there is an I-PMSI for that
   MVPN, instantiating the I-PMSI with unidirectional P-tunnels (i.e.,
   with P2MP multicast distribution trees) requires n multicast
   distribution trees, each one rooted at a different PE. If the I-PMSI
   is instantiated by a bidirectional P-tunnel, a single multicast
   distribution tree can be used, assuming appropriate support by the
   provisioning system.

   An SP may decide to use bidirectional P-tunnels for either or both of
   these reasons. Note that even if the reason for using bidirectional
   P-tunnels is to provide C-BIDIR support, the same P-tunnels can also
   be used to carry unidirectional C-flows, if that is the choice of the
   SP.

   These two reasons for using bidirectional P-tunnels may appear to be
   somewhat in conflict with each other, since (as will be seen in
   subsequent sections) the use of bidirectional P-tunnels for C-BIDIR
   support may require multiple bidirectional P-tunnels per VPN. Each
   such P-tunnel is associated with a particular "distinguished PE", and
   can only carry those C-BIDIR flows whose C-RPAs are reachable through
   its distinguished PE. However, on platforms that support MPLS
   upstream-assigned labels ([RFC5331]), PE Distinguisher Labels
   (Section 4 of [RFC6513] and Section 8 of [RFC6514]) can be used to
   aggregate multiple bidirectional P-tunnels onto a single outer
   bidirectional P-tunnel, thereby allowing one to provide C-BIDIR
   support with minimal state at the transit nodes.

   Since there are two fundamentally different reasons for using
   bidirectional P-tunnels, and since many deployed router platforms do
   not support upstream-assigned labels at the current time, this
   document specifies several different methods of using bidirectional
   P-tunnels to instantiate PMSIs. We refer to these as "PMSI
   Instantiation Methods". The method or methods deployed by any

particular SP will depend upon that SP's goals and engineering trade-
offs and upon the set of platforms deployed by that SP.

The rules for using bidirectional P-tunnels in I-PMSI or S-PMSI A-D
routes are not exactly the same as the rules for using unidirectional
P-tunnels, and the rules are also different for the different PMSI
instantiation methods. Subsequent sections of this document specify
the rules in detail.

1.2.3. Knowledge of Group-to-RP and/or Group-to-RPA Mappings

If a VPN customer is making use of a particular ASM group address,
the PEs of that VPN generally need to know the group-to-RP mappings
that are used within the VPN. If a VPN customer is making use of
BIDIR-PIM group addresses, the PEs need to know the group-to-RPA
mappings that are used within the VPN. Commonly, the PEs obtain this
knowledge either through provisioning or by participating in a
dynamic "group-to-RP(A) mapping discovery protocol" that runs within
the VPN. However, the way in which this knowledge is obtained is
outside the scope of this document.

The PEs also need to be able to forward traffic towards the C-RPs
and/or C-RPAs and to determine whether the next-hop interface of the
route to a particular C-RP(A) is a VRF interface or a PMSI. This is
done by applying the procedures of [RFC6513], Section 5.1.

1.2.4. PMSI Instantiation Methods

This document specifies three methods for using bidirectional
P-tunnels to instantiate PMSIs: two partitioned methods (the Flat
Partitioned Method and the Hierarchical Partitioned Method) and the
Unpartitioned Method.

o Partitioned Methods

   In the Partitioned Methods, a particular PMSI is instantiated by a
   set of bidirectional P-tunnels. These P-tunnels may be aggregated
   (as inner P-tunnels) into a single outer bidirectional P-tunnel
   ("Hierarchical Partitioning"), or they may be unaggregated ("Flat
   Partitioning"). Any PE that joins one of these P-tunnels can
   transmit a packet on it, and the packet will be received by all
   the other PEs that have joined the P-tunnel. For each such
   P-tunnel (each inner P-tunnel, in the case of Hierarchical
   Partitioning) there is one PE that is its distinguished PE. When
   a PE receives a packet from a given P-tunnel, the PE can determine
   from the packet's encapsulation the P-tunnel it has arrived on,
   and it can thus infer the identity of the distinguished PE
   associated with the packet. This association plays an important

role in the treatment of the packet, as specified later on in this
document.

The number of P-tunnels needed (the number of inner P-tunnels
needed, if Hierarchical Partitioning is used) depends upon a
number of factors that are described later in this document.

The Hierarchical Partitioned Method requires the use of upstream-
assigned MPLS labels (PE Distinguisher Labels) and requires the
use of the PE Distinguisher Labels attribute in BGP. The Flat
Partitioned Method requires neither of these.

The Partitioned Method (either Flat or Hierarchical) is a
prerequisite for implementing the "Partitioned Sets of PEs"
technique of supporting C-BIDIR, as discussed in [RFC6513],
Section 11.2. The Partitioned Method (either Flat or
Hierarchical) is also a prerequisite for applying the "Discarding
Packets from Wrong PE" technique, discussed in [RFC6513], Section
9.1.1, to a PMSI that is instantiated by a bidirectional P-tunnel.

The Flat Partitioned Method is a prerequisite for implementing the
"Partial Mesh of MP2MP P-Tunnels" technique for carrying customer
bidirectional (C-BIDIR) traffic, as discussed in [RFC6513],
Section 11.2.3.

The Hierarchical Partitioned Method is a prerequisite for
implementing the "Using PE Distinguisher Labels" technique of
carrying customer bidirectional (C-BIDIR) traffic, as discussed in
[RFC6513], Section 11.2.2.

Note that a particular deployment may choose to use the
Partitioned Methods for carrying the C-BIDIR traffic on
bidirectional P-tunnels, while carrying other traffic either on
unidirectional P-tunnels or on bidirectional P-tunnels using the
Unpartitioned Method. Routers in a given deployment must be
provisioned to know which PMSI instantiation method to use for
which PMSIs.

There may be ways of implementing the Partitioned Methods with
PMSIs that are instantiated by unidirectional P-tunnels. (See,
e.g., [MVPN-BIDIR-IR].) However, that is outside the scope of the
current document.

o Unpartitioned Method

In the Unpartitioned Method, a particular PMSI can be instantiated
by a single bidirectional P-tunnel. Any PE that joins the tunnel
can transmit a packet on it, and the packet will be received by

all the other PEs that have joined the tunnel. The receiving PEs
can determine the tunnel on which the packet was transmitted, but
they cannot determine which PE transmitted the packet, nor can
they associate the packet with any particular distinguished PE.

When the Unpartitioned Method is used, this document does not
mandate that only one bidirectional P-tunnel be used to
instantiate each PMSI. It allows for the case where more than one
P-tunnel is used. In this case, the transmitting PEs will have a
choice of which such P-tunnel to use when transmitting, and the
receiving PEs must be prepared to receive from any of those
P-tunnels. The use of multiple P-tunnels in this case provides
additional robustness, but it does not provide additional
functionality.

If bidirectional P-tunnels are being used to instantiate the PMSIs of
a given MVPN, one of these methods must be chosen for that MVPN. All
the PEs of that MVPN must be provisioned to know the method that is
being used for that MVPN.

I-PMSIs may be instantiated by bidirectional P-tunnels using either
the Partitioned (either Flat or Hierarchical) Methods or the
Unpartitioned Method. The method used for a given MVPN 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 for the given MVPN's I-PMSI.

If a bidirectional P-tunnel is used to instantiate an S-PMSI
(including the case of a (C-*,C-*) S-PMSI), either the Partitioned
Methods (either Flat or Hierarchical) or the Unpartitioned Method may
be used. The method used by a given VRF is determined by
provisioning. It is desirable to be able to provision this on a per-
MVPN basis. All the VRFs of a single MVPN MUST be provisioned to use
the same method for those of their S-PMSIs that are instantiated by
bidirectional P-tunnels.

If one of the Partitioned Methods is used, all the VRFs of a single
MVPN MUST be provisioned to use the same variant of the Partitioned
Methods, i.e., either they must all use the Flat Partitioned Method
or they must all use the Hierarchical Partitioned Method.

It is valid to use the Unpartitioned Method to instantiate the
I-PMSIs, while using one of the Partitioned Methods to instantiate
the S-PMSIs.

It is valid to instantiate some S-PMSIs by unidirectional P-tunnels
and others by bidirectional P-tunnels.

The procedures for the use of bidirectional P-tunnels, specified in
subsequent sections of this document, depend on both the tunnel
technology and the PMSI instantiation method. Note that this
document does not specify procedures for every possible combination
of tunnel technology and PMSI instantiation method.

2. The All BIDIR-PIM Wildcard

   [RFC6514] specifies the method of encoding C-multicast source and
   group addresses into the NLRI of certain BGP routes. [RFC6625]
   extends that specification by allowing the source and/or group
   address to be replaced by a wildcard. 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 C-flows. To do this, it is
   necessary to have a way to encode a (C-*,C-*) wildcard that is
   restricted to BIDIR-PIM C-groups.

   Therefore, we 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. (This encoding is distinct from the group wildcard
   encoding defined in [RFC6625]). We will use the notation
   (C-*,C-*-BIDIR) to refer to the "all BIDIR-PIM groups" wildcard.

3. Using 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.

3.1. Procedures Specific to the Tunneling Technology

   This section discusses the procedures that are specific to a given
   tunneling technology (BIDIR-PIM or the MP2MP procedures of mLDP
   (Multipoint LDP)) but that are independent of the method
   (Unpartitioned, Flat Partitioned, or Hierarchical Partitioned) used
   to instantiate a PMSI.

3.1.1. BIDIR-PIM P-Tunnels

   Each BIDIR-PIM P-tunnel is identified by a unique P-group address
   ([RFC6513], Section 3.1). (The P-group address is called a
   "P-Multicast Group" in [RFC6514]). Section 5 of [RFC6514] 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.

   The use of BIDIR-PIM P-tunnels to support the Hierarchical
   Partitioned Method is outside the scope of this document.

   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 originator of the route SHOULD NOT
   include a PE Distinguisher Labels attribute. If it does, that
   attribute MUST be ignored. When we say the attribute is "ignored",
   we do not mean that its normal BGP processing is not done, but that
   the attribute has no effect on the data plane. However, it MUST be
   treated by BGP as if it were an unsupported optional transitive
   attribute. (PE Distinguisher Labels are used for the Hierarchical
   Partitioning Method, but this document does not provide support for
   the Hierarchical Partitioning Method with BIDIR-PIM P-tunnels.)

3.1.2. MP2MP LSPs

   Each MP2MP LSP is identified by a unique "MP2MP FEC (Forwarding
   Equivalence Class) element" [RFC6388]. 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;
   there is no need for different root nodes to use the same opaque
   value for a given MVPN.

   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 or types of tunnel identifier
   that is used in a given deployment. A given implementation is not
   expected to be able to advertise (in the PTAs of I-PMSI or S-PMSI A-D
   routes) tunnel identifiers of every possible type. However, an
   implementation SHOULD be able to accept and properly process a PTA
   that uses any legal type of tunnel identifier.

   Section 5 of [RFC6514] 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.

3.2. Procedures Specific to the PMSI Instantiation Method

   When either the Flat Partitioned Method or the Hierarchical
   Partitioned Method is used to implement the "Partitioned Sets of PEs"
   method of supporting C-BIDIR, as discussed in Section 11.2 of
   [RFC6513] and Section 3.6 of [RFC6517], a C-BIDIR flow MUST be
   carried only on an I-PMSI or on a (C-*,C-G-BIDIR), (C-*,C-*-BIDIR),
   or (C-*,C-*) S-PMSI. A PE MUST NOT originate any (C-S,C-G-BIDIR)
   S-PMSI A-D routes. (Though it may, of course, originate (C-S,C-G)
   S-PMSI A-D routes for C-G's that are not C-BIDIR groups.) Packets of
   a C-BIDIR flow MUST NOT be carried on a (C-S,C-*) S-PMSI.

   Sections 3.2.1 and 3.2.2 specify additional details of the two
   Partitioned Methods.

3.2.1. Flat Partitioning

   The procedures of this section and its subsections apply when (and
   only when) the Flat Partitioned Method is used. This method is
   introduced in [RFC6513], 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.

When a PE originates an I-PMSI or S-PMSI A-D route whose PTA
specifies a bidirectional P-tunnel, the PE MUST be the root node of
the specified P-tunnel.

If BIDIR-PIM P-tunnels are used, each advertised P-tunnel MUST have a
distinct P-group address. The PE advertising the tunnel will be
considered to be the root node of the tunnel. Note that this creates
a unique mapping from P-group address to root node. The assignment
of P-group addresses to MVPNs is by provisioning.

If MP2MP LSPs are used, each P-tunnel MUST have a distinct MP2MP FEC
(i.e., a distinct combination of root node and opaque value). The PE
advertising the tunnel MUST be the same PE identified in the root
node field of the MP2MP FEC that is encoded in the PTA.

It follows that two different PEs may not advertise the same
bidirectional P-tunnel. Any PE that receives a packet from the
P-tunnel can infer the identity of the P-tunnel from the packet's
encapsulation. Once the identity of the P-tunnel is known, the root
node of the P-tunnel is also known. The root node of the P-tunnel on
which the packet arrived is treated as the distinguished PE for that
packet.

The Flat Partitioned Method does not use upstream-assigned labels in
the data plane, and hence does not use the BGP PE Distinguisher
Labels attribute. When this method is used, I-PMSI and/or S-PMSI A-D
routes SHOULD NOT contain a PE Distinguisher Labels attribute; if
such an attribute is present in a received I-PMSI or S-PMSI A-D
route, it MUST be ignored. (When we say the attribute is "ignored",
we do not mean that its normal BGP processing is not done, but that
the attribute has no effect on the data plane. It MUST, however, be
treated by BGP as if it were an unsupported optional transitive
attribute.)

When the Flat Partitioned Method is used to instantiate the I-PMSIs
of a given MVPN, every PE in that MVPN that originates an Intra-AS
I-PMSI A-D route MUST include a PTA that specifies a bidirectional
P-tunnel. If the intention is to carry C-BIDIR traffic on the
I-PMSI, a PE MUST originate an Intra-AS I-PMSI A-D route if one of
its VRF interfaces is the next-hop interface on its best path to the
C-RPA of any bidirectional C-group of the MVPN.

When the Flat Partitioned Method is used to instantiate a (C-*,C-*)
S-PMSI, a (C-*,C-*-BIDIR) S-PMSI, or a (C-*,C-G-BIDIR) S-PMSI, a PE
that originates the corresponding S-PMSI A-D route MUST include in
that route a PTA specifying a bidirectional P-tunnel. Per the
procedures of [RFC6513] and [RFC6514], a PE will originate such an
S-PMSI A-D route only if one of the PE's VRF interfaces is the next-

hop interface of the PE's best path to the C-RPA of a C-BIDIR group
that is to be carried on the specified S-PMSI.

PMSIs that are instantiated via the Flat Partitioned Method may carry
customer bidirectional traffic AND customer unidirectional traffic.
The rules of Sections 3.2.1.1 and 3.2.1.2 determine when a given
customer multicast packet is a match for transmission to a given
PMSI. However, if the "Partitioned Set of PEs" method of supporting
C-BIDIR traffic is being used for a given MVPN, the PEs must be
provisioned in such a way that packets from a C-BIDIR flow of that
MVPN never match any PMSI that is not instantiated by a bidirectional
P-tunnel. (For example, if the given MVPN's (C-*,C-*) S-PMSI were
not instantiated by a bidirectional P-tunnel, one could meet this
requirement by carrying all C-BIDIR traffic of that MVPN on a
(C-*,C-*-BIDIR) S-PMSI.)

When a PE receives a customer multicast data packet from a
bidirectional P-tunnel, it associates that packet with a
distinguished PE. The distinguished PE for a given packet is the
root node of the tunnel from which the packet is received. The rules
of Sections 3.2.1.1 and 3.2.1.2 ensure that:

o 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.

o If the received 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 [RFC6513] for the C-RPA
   of the bidirectional C-group.

The rules of Sections 3.2.1.3 and 3.2.1.4 allow the receiving PEs to
determine the expected distinguished PE for each C-flow, and ensure
that a packet will be discarded if its distinguished PE is not the
expected distinguished PE for the C-flow to which the packet belongs.
This prevents duplication of data for both bidirectional and
unidirectional C-flows.

3.2.1.1. When an S-PMSI Is a 'Match for Transmission'

Suppose a given PE, say PE1, needs to transmit multicast data packets
of a particular C-flow. Section 3.1 of [RFC6625] 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, those rules apply unchanged;
the remainder of this section applies only to C-BIDIR flows. If a
C-BIDIR flow has group address C-G-BIDIR, the rules applied by PE1
are given below:

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

   * If there is a (C-*,C-G-BIDIR) S-PMSI A-D route currently
      originated by PE1, then the C-flow matches that route.

   * Otherwise, if there is a (C-*,C-*-BIDIR) S-PMSI A-D route
      currently originated by PE1, then the C-flow matches that
      route.

   * Otherwise, if there is a (C-*,C-*) S-PMSI A-D route currently
      originated by PE1, then the C-flow matches that route.

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

   * If there is an installed (C-*,C-G-BIDIR) S-PMSI A-D route
      originated by PE2, then the C-flow matches that route.

   * Otherwise, if there is an installed (C-*,C-*-BIDIR) S-PMSI A-D
      route originated by PE2, then the C-flow matches that route.

   * Otherwise, if there is an installed (C-*,C-*) S-PMSI A-D route
      originated by PE2, then the C-flow matches that route.

If there is an S-PMSI A-D route that matches a given C-flow, and if
PE1 needs to transmit packets of that C-flow or other PEs, then it
MUST transmit those packets on the bidirectional P-tunnel identified
in the PTA of the matching S-PMSI A-D route.

3.2.1.2. When an I-PMSI Is a 'Match for Transmission'

Suppose a given PE, say PE1, needs to transmit packets of a given
C-flow (of a given MVPN) to other PEs, but according to the
conditions of Section 3.2.1.1 and/or Section 3.1 of [RFC6625], that
C-flow does not match any S-PMSI A-D route. Then, the packets of the
C-flow need to be transmitted on the MVPN's I-PMSI.

If the C-flow is not a BIDIR-PIM C-flow, the P-tunnel on which the
C-flow MUST be transmitted is the one identified in the PTA of the
Intra-AS I-PMSI A-D route originated by PE1 for the given MVPN.

If the C-flow is a BIDIR-PIM C-flow with group address C-G-BIDIR, the
rules applied by PE1 are:

o Suppose that the C-RPA for C-G-BIDIR is a C-address of PE1, or
   that PE1's route to the C-RPA is via a VRF interface.   Then, if
   there is an I-PMSI A-D route currently originated by PE1, the
   C-flow MUST be transmitted on the P-tunnel identified in the PTA
   of that I-PMSI A-D route.

o If PE1 determines the upstream PE for C-G-BIDIR's C-RPA to be some
   other PE, say PE2, then if there is an installed I-PMSI A-D route
   originated by PE2, the C-flow MUST be transmitted on the P-tunnel
   identified in the PTA of that route.

If there is no I-PMSI A-D route meeting the above conditions, the
C-flow MUST NOT be transmitted.

3.2.1.3. When an S-PMSI Is a 'Match for Reception'

Suppose a given PE, say PE1, needs to receive multicast data packets
of a particular C-flow. Section 3.2 of [RFC6625] 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. The remainder of this section
applies only to C-BIDIR flows.

The rules of [RFC6625], Section 3.2.1, are not applicable to C-BIDIR
flows. The rules of [RFC6625], Section 3.2.2, are replaced by the
following rules.

Suppose PE1 needs to receive (C-*,C-G-BIDIR) traffic. Suppose also
that PE1 has determined that PE2 is the upstream PE [RFC6513] for the
C-RPA of C-G-BIDIR. Then:

o If PE1 is not the same as PE2, and PE1 has an installed (C-*,C-G-
   BIDIR) S-PMSI A-D route originated by PE2, then (C-*,C-G-BIDIR)
   matches this route.

o Otherwise, if PE1 is the same as PE2, and PE1 has currently
   originated a (C-*,C-G-BIDIR) S-PMSI A-D route, then
   (C-*,C-G-BIDIR) matches this route.

o Otherwise, if PE1 is not the same as PE2, and PE1 has an installed
   (C-*,C-*-BIDIR) S-PMSI A-D route originated by PE2, then
   (C-*,C-G-BIDIR) matches this route.

   o Otherwise, if PE1 is the same as PE2, and PE1 has currently
      originated a (C-*,C-*-BIDIR) S-PMSI A-D route, then
      (C-*,C-G-BIDIR) matches this route.

   o Otherwise, if PE1 is not the same as PE2, and PE1 has an installed
      (C-*,C-*) S-PMSI A-D route originated by PE2, then (C-*,C-G-BIDIR)
      matches this route.

   o Otherwise, if PE1 is the same as PE2, and PE1 has currently
      originated a (C-*,C-*) S-PMSI A-D route, then (C-*,C-G-BIDIR)
      matches this route.

   If there is an S-PMSI A-D route matching (C-*,C-G-BIDIR), according
   to these rules, the root node of that P-tunnel is considered to be
   the distinguished PE for that (C-*,C-G-BIDIR) flow. If a
   (C-*,C-G-BIDIR) packet is received on a P-tunnel whose root node is
   not the distinguished PE for the C-flow, the packet MUST be
   discarded.

3.2.1.4. When an I-PMSI Is a 'Match for Reception'

   Suppose a given PE, say PE1, needs to receive packets of a given
   C-flow (of a given MVPN) from another PE, but according to the
   conditions of Section 3.2.1.3 and/or Section 3.2 of [RFC6625], that
   C-flow does not match any S-PMSI A-D route. Then, the packets of the
   C-flow need to be received on the MVPN's I-PMSI.

   If the C-flow is not a BIDIR-PIM C-flow, the rules for determining
   the P-tunnel on which packets of the C-flow are expected are given in
   [RFC6513]. The remainder of this section applies only to C-BIDIR
   flows.

   Suppose that PE1 needs to receive (C-*,C-G-BIDIR) traffic from other
   PEs. Suppose also that PE1 has determined that PE2 is the upstream
   PE [RFC6513] for the C-RPA of C-G-BIDIR. Then, PE1 considers PE2 to
   be the distinguished PE for (C-*,C-G-BIDIR). If PE1 has an installed
   Intra-AS I-PMSI A-D route originated by PE2, PE1 will expect to
   receive packets of the C-flow from the tunnel specified in that
   route's PTA. (If all VRFs of the MVPN have been properly provisioned
   to use the Flat Partitioned Method for the I-PMSI, the PTA will
   specify a bidirectional P-tunnel.) Note that if PE1 is the same as
   PE2, then the relevant Intra-AS I-PMSI A-D route is the one currently
   originated by PE1.

   If a (C-*,C-G-BIDIR) packet is received on a P-tunnel other than the
   expected one, the packet MUST be discarded.

3.2.2. Hierarchical Partitioning

   The procedures of this section and its subsections apply when (and
   only when) the Hierarchical Partitioned Method is used. This method
   is introduced in [RFC6513], Section 11.2.2. This document only
   provides procedures for using this method when using MP2MP LSPs as
   the P-tunnels.

   The Hierarchical Partitioned Method provides the same functionality
   as the Flat Partitioned Method, but it requires a smaller amount of
   state to be maintained in the core of the network. However, it
   requires the use of upstream-assigned MPLS labels ("PE Distinguisher
   Labels"), which are not necessarily supported by all hardware
   platforms. The upstream-assigned labels are used to provide an LSP
   hierarchy, in which an outer MP2MP LSP carries multiple inner MP2MP
   LSPs. Transit routers along the path between PE routers then only
   need to maintain state for the outer MP2MP LSP.

   When this method is used to instantiate a particular PMSI, the
   bidirectional P-tunnel advertised in the PTA of the corresponding
   I-PMSI or S-PMSI A-D route is the outer P-tunnel. When a packet is
   received from a P-tunnel, the PE that receives it can infer the
   identity of the outer 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 outer
   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 [RFC6513], Section 12.3.) The PE Distinguisher Label
   may be thought of as identifying an inner MP2MP LSP whose root is the
   PE corresponding to that label.

   In the context of a given MVPN, if it is desired to use the
   Hierarchical Partitioned Method to instantiate an I-PMSI, a (C-*,C-*)
   S-PMSI, or a (C-*,C-*-BIDIR) S-PMSI, the corresponding A-D routes
   MUST be originated by some of the PEs that attach to that MVPN. The
   PEs that are REQUIRED to originate these routes are those that
   satisfy one of the following conditions:

   o There is a C-BIDIR group for which the best path from the PE to
      the C-RPA of that C-group is via a VRF interface.

   o The PE might have to transmit unidirectional customer multicast
      traffic on the PMSI identified in the route (of course this
      condition does not apply to (C-*,C-*-BIDIR) or to (C-*,C-G-BIDIR)
      S-PMSIs).

   o The PE is the root node of the MP2MP LSP that is used to
      instantiate the PMSI.

When the Hierarchical Partitioned method is used to instantiate a
(C-*,C-G-BIDIR) S-PMSI, the corresponding (C-*,C-G-BIDIR) S-PMSI
route MUST NOT be originated by a given PE unless either (a) that
PE's best path to the C-RPA for C-G-BIDIR is via a VRF interface, or
(b) the C-RPA is a C-address of the PE. Further, that PE MUST be the
root node of the MP2MP LSP identified in the PTA of the S-PMSI A-D
route.

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

Suppose that for a given MVPN, the Hierarchical Partitioned Method is
used to instantiate the I-PMSI. In general, more than one of the PEs
in the MVPN will originate an Intra-AS I-PMSI A-D route for that
MVPN. This document allows the PTAs of those routes to all specify
the same MP2MP LSP as the "outer tunnel". However, it does not
require that those PTAs all specify the same MP2MP LSP as the outer
tunnel. By having all the PEs specify the same outer tunnel for the
I-PMSI, one can minimize the amount of state in the transit nodes.
By allowing them to specify different outer tunnels, one uses more
state, but may increase the robustness of the system.

The considerations of the previous paragraph apply as well when the
Hierarchical Partitioned Method is used to instantiate an S-PMSI.

3.2.2.1. Advertisement of PE Distinguisher Labels

A PE Distinguisher Label is an upstream-assigned MPLS label [RFC5331]
that can be used, in the context of an MP2MP LSP, to denote a
particular PE that either has joined or may in the future join that
LSP.

In order to use upstream-assigned MPLS labels in the context of an
outer MP2MP LSP, there must be a convention that identifies a
particular router as the router that is responsible for allocating
the labels and for advertising the labels to the PEs that may join
the MP2MP LSP. This document REQUIRES that the PE Distinguisher
Labels used in the context of a given MP2MP LSP be allocated and
advertised by the router that is the root node of the LSP.

This convention accords with the rules of Section 7 of [RFC5331].
Note that according to Section 7 of [RFC5331], upstream-assigned
labels are unique in the context of the IP address of the root node;
if two MP2MP LSPs have the same root node IP address, the upstream-
assigned labels used within the two LSPs come from the same label
space.

This document assumes that the root node address of an MP2MP LSP is
an IP address that is uniquely assigned to the node. The use of an
"anycast address" as the root node address is outside the scope of
this document.

A PE Distinguisher Labels attribute SHOULD NOT be attached to an
I-PMSI or S-PMSI A-D route unless that route also contains a PTA that
specifies an MP2MP LSP. (While PE Distinguisher Labels could in
theory also be used if the PTA specifies a BIDIR-PIM P-tunnel, such
use is outside the scope of this document.)

The PE Distinguisher Labels attribute specifies a set of <MPLS label,
IP address> bindings. Within a given PE Distinguisher Labels
attribute, each such IP address MUST appear at most once, and each
MPLS label MUST appear only once. Otherwise, the attribute is
considered to be malformed, and the "treat-as-withdraw" error-
handling approach described in Section 2 of [BGP-ERROR] MUST be used.

When a PE Distinguisher Labels attribute is included in a given
I-PMSI or S-PMSI A-D route, it MUST assign a label to the IP address
of each of the following PEs:

o The root node of the MP2MP LSP identified in the PTA of the route.

o Any PE that is possibly the ingress PE for a C-RPA of any C-BIDIR
   group.

o Any PE that may need to transmit non-C-BIDIR traffic on the MP2MP
   LSP identified in the PTA of the route.

One simple way to meet these requirements is to assign a PE
Distinguisher label to every PE that has originated an Intra-AS
I-PMSI A-D route.

3.2.2.2. When an S-PMSI Is a 'Match for Transmission'

Suppose a given PE, say PE1, needs to transmit multicast data packets
of a particular C-flow. Section 3.1 of [RFC6625] 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, those rules apply unchanged.
If there is a matching S-PMSI A-D route, the P-tunnel on which the
C-flow MUST be transmitted is the one identified in the PTA of the
matching route. Each packet of the C-flow MUST carry the PE
Distinguisher Label assigned by the root node of that P-tunnel to the
IP address of PE1. See Section 12.3 of [RFC6513] for encapsulation
details.

The remainder of this section applies only to C-BIDIR flows. If a
C-BIDIR flow has group address C-G-BIDIR, the rules applied by PE1
are the same as the rules given in Section 3.2.1.1.

If there is a matching S-PMSI A-D route, PE1 MUST transmit the C-flow
on the P-tunnel identified in its PTA. Suppose PE1 has determined
that PE2 is the upstream PE for the C-RPA of the given C-flow. In
constructing the packet's MPLS label stack, PE1 must use the PE
Distinguisher Label that was assigned by the P-tunnel's root node to
the IP address of "PE2", not the label assigned to the IP address of
"PE1" (unless, of course, PE1 is the same as PE2). See Section 12.3
of [RFC6513] for encapsulation details. Note that the root of the
P-tunnel might be a PE other than PE1 or PE2.

3.2.2.3. When an I-PMSI Is a 'Match for Transmission'

Suppose a given PE, say PE1, needs to transmit packets of a given
C-flow (of a given MVPN) to other PEs, but according to the
conditions of Section 3.2.2.2 and/or Section 3.1 of [RFC6625], that
C-flow does not match any S-PMSI A-D route. Then the packets of the
C-flow need to be transmitted on the MVPN's I-PMSI.

If the C-flow is not a BIDIR-PIM C-flow, the P-tunnel on which the
C-flow MUST be transmitted is the one identified in the PTA of the
Intra-AS I-PMSI A-D route originated by PE1 for the given MVPN. Each
packet of the C-flow 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 with group address C-G-BIDIR, the
rules as applied by PE1 are the same as those given in Section
3.2.1.2.

If there is a matching I-PMSI A-D route, PE1 MUST transmit the C-flow
on the P-tunnel identified in its PTA. In constructing the packet's
MPLS label stack, it must use the PE Distinguisher Label that was
assigned by the P-tunnel's root node to the IP address of "PE2", not
the label assigned to the IP address of "PE1" (unless, of course, PE1
is the same as PE2). (Section 3.2.1.2 specifies the difference
between PE1 and PE2.) See Section 12.3 of [RFC6513] for
encapsulation details. Note that the root of the P-tunnel might be a
PE other than PE1 or PE2.

If, for a packet of a particular C-flow, there is no S-PMSI A-D route
or I-PMSI A-D route that is a match for transmission, the packet MUST
NOT be transmitted.

3.2.2.4. When an S-PMSI Is a 'Match for Reception'

   Suppose a given PE, say PE1, needs to receive multicast data packets
   of a particular C-flow. Section 3.2 of [RFC6625] 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. The rules are modified in this section, as follows:

   Consider a particular C-flow. Suppose either:

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

   o 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,

   2. there is an installed S-PMSI A-D route originated by the root node
      of that LSP, or PE1 itself is 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.

3.2.2.5. When an I-PMSI Is a 'Match for Reception'

   If a PE needs to receive packets of a given C-flow (of a given MVPN)
   from another PE, and if, according to the conditions of Section
   3.2.2.4, that C-flow does not match any S-PMSI A-D route, then the
   packets of the C-flow need to be received on the MVPN's I-PMSI. The
   P-tunnel on which the packets are expected to arrive is determined by
   the Intra-AS I-PMSI A-D route originated by the distinguished PE for
   the given C-flow. The PTA of that route specifies the "outer

P-tunnel" and thus determines the top label that packets of that
C-flow will be carrying when received. A PE that needs to receive
packets of a given C-flow must determine the expected value of the
second label for packets of that C-flow. This will be the value of a
PE Distinguisher Label, taken from the PE Distinguisher Labels
attribute of the Intra-AS I-PMSI A-D route of the root node of that
outer tunnel. The expected value of the second label on received
packets (corresponding to the "inner tunnel") of a given C-flow is
determined according to the following rules.

First, the distinguished PE for the C-flow is determined:

o If the C-flow is not a BIDIR-PIM C-flow, the distinguished PE for
   the C-flow is its upstream PE, as determined by the rules of
   [RFC6513].

o If the C-flow is a BIDIR-PIM C-flow, the distinguished PE for the
   C-flow is its upstream PE of the C-flow's C-RPA, as determined by
   the rules of [RFC6513].

The expected value of the second label is the value that the root PE
of the outer tunnel has assigned, in the PE Distinguisher Labels
attribute of its Intra-AS I-PMSI A-D route, to the IP address of the
distinguished PE.

Packets addressed to C-G that arrive on other than the expected inner
and outer P-tunnels (i.e., that arrive with unexpected values of the
top two labels) MUST be discarded.

3.2.3. Unpartitioned

   When a particular MVPN uses the Unpartitioned Method of instantiating
   an I-PMSI with a bidirectional P-tunnel, it MUST be the case that at
   least one VRF of that MVPN originates an Intra-AS I-PMSI A-D route
   that includes a PTA specifying a bidirectional P-tunnel. The
   conditions under which an Intra-AS I-PMSI A-D route must be
   originated from a given VRF are as specified in [RFC6514]. This
   document allows all but one of such routes to omit the PTA. However,
   each such route MAY contain a PTA. If the PTA is present, it MUST
   specify a bidirectional P-tunnel. As specified in [RFC6513] and
   [RFC6514], every PE that imports such an Intra-AS I-PMSI A-D route
   into one of its VRFs MUST, if the route has a PTA, join the P-tunnel
   specified in the route's PTA.

   Packets received on any of these P-tunnels are treated as having been
   received over the I-PMSI. The disposition of a received packet MUST
   NOT depend upon the particular P-tunnel over which it has been
   received.

When a PE needs to transmit a packet on such an I-PMSI, then if that
PE advertised a P-tunnel in the PTA of an Intra-AS I-PMSI A-D route
that it originated, the PE SHOULD transmit the on that P-tunnel.
However, any PE that transmits a packet on the I-PMSI MAY transmit it
on any of the P-tunnels advertised in any of the currently installed
Intra-AS I-PMSI A-D routes for its VPN.

This allows a single bidirectional P-tunnel to be used to instantiate
the I-PMSI, but also allows the use of multiple bidirectional
P-tunnels. There may be a robustness advantage in having multiple
P-tunnels available for use, but the number of P-tunnels used does
not impact the functionality in any way. If there are, e.g., two
P-tunnels available, these procedures allow each P-tunnel to be
advertised by a single PE, but they also allow each P-tunnel to be
advertised by multiple PEs. Note that the PE advertising a given
P-tunnel does not have to be the root node of the tunnel. The root
node might not even be a PE router, and it might not originate any
BGP routes at all.

In the Unpartitioned Method, packets received on the I-PMSI cannot be
associated with a distinguished PE, so duplicate detection using the
techniques of Section 9.1.1 of [RFC6513] is not possible; the
techniques of Sections 9.1.2 or 9.1.3 of [RFC6513] would have to be
used instead. Support for C-BIDIR using the "Partitioned set of PEs"
technique (Section 11.2 of [RFC6513] and Section 3.6 of [RFC6517]) is
not possible when the Unpartitioned Method is used. If it is desired
to use that technique to support C-BIDIR, but also to use the
Unpartitioned Method to instantiate the I-PMSI, then all the C-BIDIR
traffic would have to be carried on an S-PMSI, where the S-PMSI is
instantiated using one of the Partitioned Methods.

When a PE, say PE1, needs to transmit multicast data packets of a
particular C-flow to other PEs, and PE1 does not have an S-PMSI that
is a match for transmission for that C-flow (see Section 3.2.3.1),
PE1 transmits the packets on one of the P-tunnel(s) that instantiates
the I-PMSI. When a PE, say PE1, needs to receive multicast data
packets of a particular C-flow from another PE, and PE1 does not have
an S-PMSI that is a match for reception for that C-flow (see Section
3.2.3.2), PE1 expects to receive the packets on any of the P-tunnels
that instantiate the I-PMSI.

When a particular MVPN uses the Unpartitioned Method to instantiate a
(C-*,C-*) S-PMSI or a (C-*,C-*-BIDIR) S-PMSI using a bidirectional
P-tunnel, the same conditions apply as when an I-PMSI is instantiated
via the Unpartitioned Method. The only difference is that a PE need
not join a P-tunnel that instantiates the S-PMSI unless that PE needs
to receive multicast packets on the S-PMSI.

When a particular MVPN uses bidirectional P-tunnels to instantiate
other S-PMSIs, different S-PMSI A-D routes that do not contain
(C-*,C-*) or (C-*,C-*-BIDIR), 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.

While the Unpartitioned Method MAY be used to instantiate an S-PMSI
to which one or more C-BIDIR flows are bound, it must be noted that
the "Partitioned Set of PEs" method discussed in Section 11.2 of
[RFC6513] and Section 3.6 of [RFC6517] cannot be supported using the
Unpartitioned Method. C-BIDIR support would have to be provided by
the procedures of [RFC6513], Section 11.1.

3.2.3.1. When an S-PMSI Is a 'Match for Transmission'

Suppose a PE needs to transmit multicast data packets of a particular
customer C-flow. [RFC6625], 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-PIM groups are also 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:

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

When the Unpartitioned Method is used, the PE SHOULD transmit the
C-flow on the P-tunnel advertised in the in the matching S-PMSI A-D
route, but it MAY transmit the C-flow on any P-tunnel that is
advertised in the PTA of any installed S-PMSI A-D route that contains
the same (C-S,C-G) as the matching S-PMSI A-D route.

3.2.3.2. When an S-PMSI Is a 'Match for Reception'

Suppose a PE needs to receive multicast data packets of a particular
customer C-flow. Section 3.2 of [RFC6625] 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 [RFC6625] 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:

When the Unpartitioned Method is used, the PE MUST join the P-tunnel
that is advertised in the matching S-PMSI A-D route, and it MUST also
join the P-tunnels that are advertised in other installed S-PMSI A-D
routes that contain the same (C-S,C-G) as the matching S-PMSI A-D
route.

3.2.4. Minimal Feature Set for Compliance

Implementation of bidirectional P-tunnels is OPTIONAL. If
bidirectional P-tunnels are not implemented, the issue of compliance
to this specification does not arise. However, for the case where
bidirectional P-tunnels ARE implemented, this section specifies the
minimal set of features that MUST be implemented in order to claim
compliance to this specification.

In order to be compliant with this specification, an implementation
that provides bidirectional P-tunnels MUST support at least one of
the two P-tunnel technologies mentioned in Section 1.2.1.

A PE that does not provide C-BIDIR support using the "partitioned set
of PEs" method is deemed compliant to this specification if it
supports the Unpartitioned Method, using either MP2MP LSPs or BIDIR-
PIM multicast distribution trees as P-tunnels.

A PE that does provide C-BIDIR support using the "partitioned set of
PEs" method MUST, at a minimum, be able to provide C-BIDIR support
using the "Partial Mesh of MP2MP P-tunnels" variant of this method
(see Section 11.2 of [RFC6513]). An implementation will be deemed
compliant to this minimum requirement if it can carry all of a VPN's
C-BIDIR traffic on a (C-*,C-*-BIDIR) S-PMSI that is instantiated by a
bidirectional P-tunnel, using the Flat Partitioned Method.

4. Security Considerations

   There are no additional security considerations beyond those of
   [RFC6513] and [RFC6514], or any that may apply to the particular
   protocol used to set up the bidirectional tunnels ([RFC5015],
   [RFC6388]).

5. References

5.1. Normative References

   [RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119,
               DOI 10.17487/RFC2119, March 1997,
               <http://www.rfc-editor.org/info/rfc2119>.

   [RFC4364]   Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
               Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364,
               February 2006, <http://www.rfc-editor.org/info/rfc4364>.

   [RFC4601]   Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas,
               "Protocol Independent Multicast - Sparse Mode (PIM-SM):
               Protocol Specification (Revised)", RFC 4601,
               DOI 10.17487/RFC4601, August 2006,
               <http://www.rfc-editor.org/info/rfc4601>.

   [RFC5015]   Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano,
               "Bidirectional Protocol Independent Multicast (BIDIR-
               PIM)", RFC 5015, DOI 10.17487/RFC5015, October 2007,
               <http://www.rfc-editor.org/info/rfc5015>.

   [RFC6388]   Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B.
               Thomas, "Label Distribution Protocol Extensions for
               Point-to-Multipoint and Multipoint-to-Multipoint Label
               Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November
               2011, <http://www.rfc-editor.org/info/rfc6388>.

   [RFC6513]   Rosen, E., Ed., and R. Aggarwal, Ed., "Multicast in
               MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513,
               February 2012, <http://www.rfc-editor.org/info/rfc6513>.

   [RFC6514]   Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP
               Encodings and Procedures for Multicast in MPLS/BGP IP
               VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012,
               <http://www.rfc-editor.org/info/rfc6514>.

   [RFC6625]   Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R.
                Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes",
                RFC 6625, DOI 10.17487/RFC6625, May 2012,
                <http://www.rfc-editor.org/info/rfc6625>.

5.2. Informative References

   [BGP-ERROR] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K.
                Patel, "Revised Error Handling for BGP UPDATE Messages",
                Work in Progress, draft-ietf-idr-error-handling-19, April
                2015.

   [MVPN-BIDIR-IR]
                Zhang, Z., Rekhter, Y., and A. Dolganow, "Simulating
                'Partial Mesh of MP2MP P-Tunnels' with Ingress
                Replication", Work in Progress,
                draft-ietf-bess-mvpn-bidir-ingress-replication-00,
                January 2015.

   [MVPN-XNET] Rekhter, Y., Ed., Rosen, E., Ed., Aggarwal, R., Cai, Y.,
                and T. Morin, "Extranet Multicast in BGP/IP MPLS VPNs",
                Work in Progress, draft-ietf-bess-mvpn-extranet-02, May
                2015.

   [RFC5331]   Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream
                Label Assignment and Context-Specific Label Space", RFC
                5331, DOI 10.17487/RFC5331, August 2008,
                <http://www.rfc-editor.org/info/rfc5331>.

   [RFC6517]   Morin, T., Ed., Niven-Jenkins, B., Ed., Kamite, Y.,
                Zhang, R., Leymann, N., and N. Bitar, "Mandatory Features
                in a Layer 3 Multicast BGP/MPLS VPN Solution", RFC 6517,
                DOI 10.17487/RFC6517, February 2012,
                <http://www.rfc-editor.org/info/rfc6517>.

Authors' Addresses

   Eric C. Rosen
   Juniper Networks, Inc.
   10 Technology Park Drive
   Westford, MA 01886
   United States

   Email: erosen@juniper.net

   IJsbrand Wijnands
   Cisco Systems, Inc.
   De kleetlaan 6a
   Diegem 1831
   Belgium

   Email: ice@cisco.com

   Yiqun Cai
   Microsoft
   1065 La Avenida
   Mountain View, CA 94043
   United States

   Email: yiqunc@microsoft.com

   Arjen Boers

   Email: arjen@boers.com