Network Working Group   Rahul Aggarwal (Juniper Networks)
Internet Draft   Wim Henderickx (Alcatel-Lucent)
Expiration Date: August 2011   Praveen Muley (Alcatel-Lucent)
Intended Status: Proposed Standard   Ray Qiu (Huawei)
                               Yakov Rekhter (Juniper Networks)
                               February 14, 2011

Use of Wildcard in S-PMSI Auto-Discovery Routes

draft-rekhter-mvpn-wildcard-spmsi-04.txt

Status of this Memo

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

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF), its areas, and its working groups. Note that other
   groups may also distribute working documents as Internet-Drafts.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   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.

Internet Draft draft-rekhter-mvpn-wildcard-spmsi-04.txtFeburary 14, 2011

Abstract

The current MVPN specifications do not define encoding and procedures
   for advertising in a single route binding of multiple multicast
   streams of a given MVPN customer to a single provider's tunnel. This
   document defines such encoding and procedures. These procedures allow
   in certain situations to reduce MVPN control plane load (note though
   that these procedures have no impact on the data plane load). The
   procedures specified in this document assume that BGP is used for
   transmission of MVPN customers' routing information within the
   service provider(s) infrastructure.

Specification of Requirements

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in [RFC 2119](RFC 2119) [[RFC2119](RFC2119)].

[1](1). Introduction

   An S-PMSI auto-discovery route (A-D route), as defined in [MVPN-BGP],
   advertises binding of a given MVPN customer multicast flow (C-
   multicast flow) to a particular provider tunnel (P-tunnel). While the
   definition and procedures specified in [MVPN-BGP] support binding of
   multiple C-multicast flows to the same P-tunnel (by having multiple
   S-PMSI A-D routes advertise the same P-tunnel), they do not support
   the ability to advertise such a binding in a single S-PMSI A-D route.

   The ability of a PE to advertise binding of multiple C-multicast
   flows, all originating from the site(s) of a given MVPN connected to
   that PE, to a single P-tunnel in a single S-PMSI A-D route, rather
   than in multiple S-PMSI A-D routes (one per each C-multicast flow),
   improves control plane scalability, as it reduces the number of S-
   PMSI A-D routes. Note however, that the ability to advertise binding
   of multiple C-multicast flows to a single P-tunnel in a single S-PMSI
   A-D route has no impact on the forwarding/data plane scalability, as
   it does not reduces the number of P-tunnels, relative to the scenario
   where each C-multicast flow is advertised via its own S-PMSI A-D
   route, while all these routes advertise the same P-tunnel.

   One possible application of advertising binding of multiple C-
   multicast flows to a single P-tunnel in a single S-PMSI A-D route is
   when these are PIM-SM in ASM mode flows. In this case a PE router
   connected to an MVPN customer's site that contains customer's RP (C-
   RP) could bind all the C-multicast flows traveling along a customer's
   RPT tree to a single P-tunnel, and advertise such binding in a single

S-PMSI A-D route. Likewise, a PE router connected to an MVPN customer's site that contains multiple (multicast) sources, all sending to the same (multicast) group, could bind all the C-mulicast flows for that group originated by these sources to a single P-tunnel, and advertise such binding in a single S-PMSI A-D route.

Another possible application of advertising binding of multiple C-multicast flows to a single P-tunnel in a single S-PMSI A-D route is when these are PIM-Bidir flows. In this case a PE router could bind to a single P-tunnel all the C-multicast flows for the same (multicast) group that have been originated within the site(s) of a given MVPN connected to that PE, and advertise such binding in a single S-PMSI A-D route.

Another possible application of advertising binding of multiple C-multicast flows to a single P-tunnel in a single S-PMSI A-D route is when these are PIM-SM in SSM mode flows. In this case a PE router could bind to a single P-tunnel all the C-multicast flows coming from a given (multicast) source located in a site connected to that PE.

Yet another possible application of advertising binding of multiple C-multicast flows to a single P-tunnel in a single S-PMSI A-D route is to carry in that P-tunnel all the C-multicast flows originated within the site(s) of a given MVPN connected to a given PE.

This document defines OPTIONAL extensions to the procedures specified in [MVPN-BGP]. These extensions allow to advertise in a single S-PMSI A-D route binding of multiple C-multicast flows to a single P-tunnel. The extensions are based on the notion of a "wildcard".

In order to use the extensions specified in this document with a particular MVPN, all the PEs that have sites of that MVPN MUST support these extensions.

The procedures specified in this document assume that BGP is used for transmission of MVPN customers' multicast (C-multicast) routing information within the service provider(s) infrastructure among the PE routers ([MVPN-BGP]).

2. Encoding of wildcard in S-PMSI A-D routes

As specified in [MVPN-BGP], the NLRI of an S-PMSI A-D route has the
following format:

```
            +----------------------------------+
            |      RD   (8 octets)             |
            +----------------------------------+
            | Multicast Source Length (1 octet) |
            +----------------------------------+
            |  Multicast Source (Variable)     |
            +----------------------------------+
            | Multicast Group Length (1 octet) |
            +----------------------------------+
            |  Multicast Group   (Variable)    |
            +----------------------------------+
            |   Originating Router's IP Addr   |
            +----------------------------------+
```

This document uses a zero value in Mutlicast Source Length or
Multicast Group Length field to indicate a wildcard value for the
respective field. This document defines procedures for the following
two combinations of wildcard S-PMSI encodings:

   + (C-*, C-G): Source Wildcard, Group specified

   + (C-S, C-*): Source specified, Group Wildcard

   + (C-*, C-*): Source Wildcard, Group Wildcard

3. Procedures for (C-*, C-G) S-PMSI A-D routes

This document covers the use of (C-*, C-G) S-PMSI A-D routes for only
the C-multicast flows when C-G is not in the SSM range. Use of (C-*,
C-G) S-PMSI A-D routes for other C-multicast flows is outside the
scope of this document.

When a PE advertises an S-PMSI A-D route whose NLRI specifies (C-*,
C-G), the PE MUST use the P-tunnel advertised in this route for
sending any PIM-SM in ASM mode or PIM-Bidir C-multicast flows for
that C-G that it needs to send (downstream) to other PEs, except for
the C-multicast flows that the PE already bound to specific (C-S, C-
G)s S-PMSIs. Just like with (C-S, C-G) S-PMSI A-D routes, the
criteria for originating (C-*, C-G) S-PMSI A-D routes is local to the
originating PE.

When a PE receives an S-PMSI A-D route whose NLRI specifies (C-*, C-

G), the PE follows the procedures specified in [MVPN-BGP], except for
the case where the PE does not originate a Shared Tree Join C-
multicast route for (C-*, C-G), and for every Source Tree Join C-
multicast route for (C-S, C-G) originated by the PE, the PE already
accepted a (specific) (C-S, C-G) S-PMSI A-D route. In that case the
PE need not take any further action upon receiving the S-PMSI A-D
route with (C-*, C-G) NLRI.

If an implementation supports (C-*, C-G) S-PMSI A-D routes, then the
implementation MUST support receiving (C-S, C-*) and (C-*, C-*) S-
PMSI A-D routes, and MAY support originating (C-S, C-*) and (C-*,
C-*) S-PMSI A-D routes.

[4](4). Procedures for (C-S, C-*) S-PMSI A-D routes

This document covers the use of (C-S, C-*) S-PMSI A-D routes for only
the C-multicast flows when C-G is in the SSM range. Use of (C-S,
C-*) S-PMSI A-D routes for other C-multicast flows is outside the
scope of this document.

When a PE advertises an S-PMSI A-D route whose NLRI specifies (C-S,
C-*), the PE MUST use the P-tunnel advertised in this route for
sending any PIM-SM in SSM mode C-multicast flows for that C-S that it
needs to send (downstream) to other PEs, except for the C-multicast
flows that the PE already bound to specific (C-S, C-G)s S-PMSIs.
Just like with (C-S, C-G) S-PMSI A-D routes, the criteria for
originating (C-S, C-*) S-PMSI A-D routes is local to the originating
PE.

When a PE receives an S-PMSI A-D route whose NLRI specifies (C-S,
C-*), the PE follows the procedures specified in [MVPN-BGP], except
for the case where for every Source Tree Join C-multicast route for
(C-S, C-G) originated by the PE, the PE already accepted a (specific)
(C-S, C-G) S-PMSI A-D route. In that case the PE need not take any
further action upon receiving the S-PMSI A-D route with (C-S, C-*)
NLRI.

If an implementation supports (C-S, C-*) S-PMSI A-D routes, then the
implementation MUST support receiving (C-*, C-G) and (C-*, C-*) S-
PMSI A-D routes, and MAY support originating (C-*, C-G) and (C-*,
C-*) S-PMSI A-D routes.

[5](5). Procedures for (C-*, C-*) S-PMSI A-D routes

   (C-*, C-*) S-PMSI A-D routes are expected to be used when for a given
   MVPN a PE has a policy not to use I-PMSI for carrying multicast
   traffic originated in the MVPN's site(s) connected to that PE (this
   is known as "S-PMSI only" mode).

   When a PE advertises an S-PMSI A-D route whose NLRI specifies (C-*,
   C-*), the PE MUST use the P-tunnel advertised in this route for
   sending any C-multicast flows that it needs to send (downstream) to
   other PEs, except for the C-multicast flows that the PE already bound
   to either specific (C-*, C-G)s S-PMSIs, or specific (C-S, C-G)s S-
   PMSIs.

   Just like with (C-S, C-G) S-PMSI A-D routes, the criteria for
   originating (C-*, C-*) S-PMSI A-D routes is local to the originating
   PE. However, the following criteria must be implemented:

     + An implementation MUST support the ability to trigger origination
       of a (C-*, C-*) S-PMSI A-D route from a given VRF on a given PE
       when this PE receives (from some other PE(s)) either a Source
       Tree Join C-multicast route with the C-S carried in the route
       being reachable via one of the PE-CE interfaces of that VRF, or a
       Shared Tree Join C-multicast route with the C-RP carried in that
       route being reachable via one of the PE-CE interfaces on that
       VRF.

     + An implementation MUST also support the ability to trigger
       origination of a (C-*, C-*) S-PMSI A-D route from a given VRF on
       a given PE at provisioning time on that PE.

   To facilitate description of the procedures for receiving (C-*, C-*)
   S-PMSI A-D routes, we introduce the following definitions:

     + We say that an (C-S, C-G) S-PMSI A-D route received by a PE
       "matches" a Source Tree Join C-multicast route for (C-S, C-G)
       originated by that PE if the upstream PE of that route is the PE
       that originates the S-PMSI A-D route.

     + We say that an (C-*, C-G) S-PMSI A-D route received by a PE
       "matches" a Source Tree Join C-multicast route for (C-S, C-G)
       originated by that PE if the upstream PE of that route is the PE
       that originates the S-PMSI A-D route, and C-G is not in the SSM
       range.

+ We say that an (C-*, C-G) S-PMSI A-D route received by a PE
  "matches" a Shared Tree Join C-multicast route for (C-*, C-G)
  originated by that PE if the upstream PE of that route is the PE
  that originates the S-PMSI A-D route, and C-G is not in the SSM
  range.

+ We say that an (C-S, C-*) S-PMSI A-D route received by a PE
  "matches" a Source Tree Join C-multicast route for (C-S, C-G)
  originated by that PE if the upstream PE of that route is the PE
  that originates the S-PMSI A-D route, and C-G is in the SSM
  range.

When a PE receives an S-PMSI A-D route whose NLRI specifies (C-*,
C-*), the PE follows the procedures specified in [MVPN-BGP], except
when:

+ for all the Source Tree Join C-multicast routes originated by the
  PE, the PE already accepted either a matching (C-S, C-G), or a
  matching (C-*, C-G), or a matching (C-S, C-*) S-PMSI A-D route,
  AND

+ for all the Shared tree Join C-multicast routes originated by the
  PE, the PE already accepted a matching (C-*, C-G) S-PMSI A-D
  route,

in which case the PE need not take any further action upon receiving
the S-PMSI A-D route with NLRI (C-*, C-*).

If an implementation supports (C-*, C-*) S-PMSI A-D routes, then the
implementation MUST support receiving (C-*, C-G) and (C-S, C-*) S-
PMSI A-D routes, and MAY support originating (C-*, C-G) and (C-S,
C-*) S-PMSI A-D routes.

[6](6). **IANA Considerations**

This document introduces no new IANA Considerations.

7. Security Considerations

This document introduces no new Security Considerations, above and beyond what is already specified in [MVPN] and [MVPN-BGP].

8. Acknowledgements

Many thanks to Thomas Morin for his contributions, review, and comments.

9. Normative References

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

[MVPN] E. Rosen, R. Aggarwal [Editors], "Multicast in MPLS/BGP IP VPNs", draft-ietf-l3vpn-2547bis-mcast, work in progress

[MVPN-BGP], R. Aggarwal, E. Rosen, T. Morin, Y. Rekhter, "BGP Encodings for Multicast in MPLS/BGP IP VPNs", draft-ietf-l3vpn-2547bis-mcast-bgp, work in progress

10. Non-normative References

11. Authors' Addresses

Rahul Aggarwal
Juniper Networks, Inc.
e-mail: rahul@juniper.net

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

Praveen Muley
Alcatel-Lucent
e-mail: Praveen.Muley@alcatel-lucent.com

Ray (Lei) Qiu
2330 Central Expressway
Santa Clara, CA 95050

      USA
      e-mail: rayq@huawei.com

      Yakov Rekhter
      Juniper Networks, Inc.
      e-mail: yakov@juniper.net