Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: June 7, 2015

E. Rosen, Ed.
Juniper Networks, Inc.
M. Sivakumar
IJ. Wijnands
Cisco Systems, Inc.
S. Aldrin
Huawei Technologies
A. Dolganow
Alcatel-Lucent
T. Przygienda
Ericsson

December 4, 2014

Multicast VPN Using BIER
draft-rosen-l3vpn-mvpn-bier-02

Abstract

The Multicast Virtual Private Network (MVPN) specifications require the use of multicast tunnels ("P-tunnels") that traverse a Service Provider's backbone network. The P-tunnels are used for carrying multicast traffic across the backbone. A variety of P-tunnel types are supported. Bit Index Explicit Replication (BIER) is a new architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. This document specifies the protocol and procedures that allow MVPN to use BIER as the method of carrying multicast traffic over an SP backbone network.

Table of Contents

1. Introduction

[RFC6513] and [RFC6514] specify the protocols and procedures that a Service Provider (SP) can use to provide Multicast Virtual Private Network (MVPN) service to its customers. Multicast tunnels are created through an SP's backbone network; these are known as "P-tunnels". The P-tunnels are used for carrying multicast traffic across the backbone. The MVPN specifications allow the use of several different kinds of P-tunnel technology.

Bit Index Explicit Replication (BIER) ([BIER_ARCH]) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. The purpose of the current document is to specify the protocols and procedures needed in order to provide MVPN service using BIER to transport the multicast traffic over the backbone.

Although BIER does not explicitly build and maintain multicast tunnels, one can think of BIER as using a number of implicitly created tunnels through a "BIER domain". In particular, one can think of there as being one Point-to-Multipoint (P2MP) tunnel from each "Bit Forwarding Ingress Router" (BFIR) to all the "Bit Forwarding Egress Routers" (BFERs) in the BIER domain, where a BIER domain is generally co-extensive with an IGP network. These "tunnels" are not specific to any particular VPN. However, the MVPN architecture provides protocols and procedures that allow the traffic of multiple MVPNs to be aggregated on a single P-tunnel. In this document, we specify how to use these multi-VPN aggregation procedures to enable BIER to transport traffic from multiple MVPNs.

MVPN traffic must sometimes traverse more than one IGP domain, whereas BIER only carries multicast traffic within a single IGP domain. However, the MVPN specifications allow P-tunnels to be "segmented", where the segmentation points may either be Autonomous System Border Routers (ASBRs), as described in [RFC6514], or Area Border Routers (ABRs), as described in [SEAMLESS_MCAST]. As long as the segmentation points are capable of acting as BFIRs and BFERs, BIER can be used to provide some or all of the segments of a P-tunnel.

This revision of the document does not specify the procedures necessary to support MVPN customers that are using BIDIR-PIM. Those procedures will be added in a future revision.

This document uses the following terminology from [BIER_ARCH]:

- BFR: Bit-Forwarding Router.
- BFIR: Bit-Forwarding Ingress Router.
- BFER: Bit-Forwarding Egress Router.

This document uses the following terminology from [RFC6513]:

- MVPN: Multicast Virtual Private Network -- a VPN [RFC4364] in which multicast service is offered.
- P-tunnel. A multicast tunnel through the network of one or more SPs. P-tunnels are used to transport MVPN multicast data
- C-S: A multicast source address, identifying a multicast source located at a VPN customer site.
- C-G: A multicast group address used by a VPN customer.

- 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). Sets of C-flows can be identified by the use of the "C-*" wildcard (see [RFC6625]), e.g., (C-*,C-G).

- I-PMSI A-D Route: Inclusive Provider Multicast Service Interface Auto-Discovery route. Carried in BGP Update messages, these routes are used to advertise the "default" P-tunnel for a particular MVPN.

- 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.

- PMSI Tunnel attribute (PTA). This BGP attribute carried is used to identify a particular P-tunnel. When C-flows of multiple VPNs is carried in a single P-tunnel, this attribute also carries the information needed to multiplex and demultiplex the C-flows.

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 [RFC2119].

2. Use of the PMSI Tunnel Attribute

As defined in [RFC6514], the PMSI Tunnel attribute is used to identify the particular P-tunnel to which one or more multicast flows are being assigned.

The PMSI Tunnel attribute (PTA)contains the following fields:

- "Tunnel Type". IANA is requested to assign a new tunnel type codepoint for "BIER". This codepoint will be used to indicate that the PMSI is instantiated by BIER.

- "Tunnel Identifier". When the "tunnel type" field is "BIER", this field contains two subfields:

  1. The first subfield is a single octet, containing the sub-domain-id of the sub-domain to which the BFIR will assign the packets that it transmits on the PMSI identified by the NLRI of the BGP I-PMSI or S-PMSI A-D route that contains this PTA. (How that sub-domain is chosen is outside the scope of this document.)

2. The second subfield is the BFR-Prefix (see [BIER_ARCH]) of the originator of the route that is carrying this PTA. This will either be a /32 IPv4 address or a /128 IPv6 address. Whether the address is IPv4 or IPv6 can be inferred from the total length of the PMSI Tunnel attribute.

- "MPLS label". This field contains an upstream-assigned MPLS label. It is assigned by the router that originates the BGP route to which the PTA is attached. Constraints on the way in which the originating router selects this label are discussed below.

- "Leaf Info Required Bit". The setting of this bit depends upon the type of route and the NLRI of the route that carries the PTA.

  * In an I-PMSI A-D route or a (C-*,C-*) S-PMSI A-D route, the bit SHOULD be clear.

  * In other S-PMSI A-D routes, the bit SHOULD be set.

Note that if a PTA specifying "BIER" is attached to an I-PMSI or S-PMSI A-D route, the route MUST NOT be distributed beyond the boundaries of a BIER domain. That is, any routers that receive the route must be in the same BIER domain as the originator of the route. If the originator is in more than one BIER domain, the route must be distributed only within the BIER domain in which the BFR-Prefix in the PTA uniquely identifies the originator. As with all MVPN routes, distribution of these routes is controlled by the provisioning of Route Targets.

Suppose an ingress PE originates two x-PMSI A-D routes, where we use the term "x-PMSI" to mean "I-PMSI or S-PMSI". Suppose both routes carry a PTA, and the PTA of each route specifies"BIER".

- If the two routes do not carry the same set of Route Targets (RTs), then their respective PTAs MUST contain different MPLS label values.

- If the ingress PE is supporting MVPN extranet ([EXTRANET]) functionality, and if the two routes originate from different VRFs, then the respective PTAs of the two routes MUST contain different MPLS label values.

- If the ingress PE is supporting the "Extranet Separation" feature of MVPN extranet (see Section 7.3 of [EXTRANET], section), and if one of the routes carries the "Extranet Separation" extended community and the other does not, then the respective PTAs of the two routes MUST contain different MPLS label values.

When segmented P-tunnels are being used, an ABR or ASBR may receive, from a BIER domain, an x-PMSI A-D route whose PTA specifies "BIER". This means that BIER is being used for one segment of a segmented P-tunnel. The ABR/ASBR may in turn need to originate an x-PMSI A-D route whose PTA identifies the next segment of the P-tunnel. The next segment may also be "BIER". Suppose an ASBR receives x-PMSI A-D routes R1 and R2, and as a result originates x-PMSI A-D routes R3 and R4 respectively, where the PTAs of each of the four routes specify a BIER.. Then the PTAs of R3 and R4 MUST NOT specify the same MPLS label, UNLESS both of the following conditions hold:

- o R1 and R2 have the same "originating router" in their respective NLRIs.

- o R1 and R2 specify the same MPLS label in their respective PTAs.

3. Explicit Tracking

[Editor's note: The procedures of this section are still under discussion, and significant changes may be expected in the next revision.]

When using BIER to transport an MVPN data packet through a BIER domain, an ingress PE functions as a BFIR (see [BIER_ARCH]). The BFIR must determine the set of BFERs to which the packet needs to be delivered. This is done by using the explicit tracking mechanism specified in [RFC6513] and [RFC6514].

To determine the set of BFERs to which a given MVPN data packet needs to be delivered, the BFIR originating an S-PMSI A-D route sets the LIR bit in the route's PTA. Per [RFC6514], the BFERs will respond with Leaf A-D routes. By matching the received Leaf A-D routes to the originated S-PMSI A-D routes, the originator of the S-PMSI A-D route determines the set of BFERs that need to receive the multicast data flow (or flows) that is (are) identified in the NLRI of the of the S-PMSI A-D route.

This requires that each BFIR originate an S-PMSI A-D route for each C-flow for which it serves as BFIR. The BFIR MAY include, in each such route, a PTA as described in Section 2. However, if the BFIR has originated an I-PMSI A-D route or a wildcard S-PMSI A-D route that "matches" (according to the rules of [RFC6625]) a particular C-flow, then it may do explicit tracking for that C-flow by originating an S-PMSI A-D route for that C-flow, but including a PTA that specifies "no tunnel type".

4. Data Plane

The MVPN application plays the role of the "multicast flow layer" as described in [BIER_ARCH].

To transmit an MVPN data packet, an ingress PE follows the rules of [RFC6625] to find the S-PMSI A-D route or I-PMSI A-D route that is a "match for transmission" for that packet. (In applying the rules of [RFC6625], any S-PMSI A-D route with a PTA specifying "no tunnel information" is ignored.) If the matching route has a PTA specifying a "BIER", the (upstream-assigned) MPLS label from that PTA is pushed on the packet's label stack. Then the packet is forwarded according to the procedures of [BIER_ARCH] and [BIER_ENCAPS]. (See especially Section 4, "Imposing and Processing the BIER Encapsulation", of [BIER_ENCAPS].)

When a BFER receives an MVPN multicast data packet that has been BIER-encapsulated, the BIER layer passes the following information to the multicast flow layer:

- The BFR-prefix corresponding to the sub-domain-id and BFIR-id in the BIER header.

- The "payload", which is an MPLS packet whose top label is an upstream-assigned label. The BFR-prefix provides the "context" in which the upstream-assigned label is interpreted.

  Note that per [RFC5331], the context for an upstream-assigned label is the IP address of the label assigner, which in this case is the BFR-prefix of the BFIR.

5. Acknowledgments

The authors wish to thank Jeffrey Zhang for his ideas and contributions to this work.

6. IANA Considerations

IANA is requested to assign a value for "BIER" from the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry. The reference should be this document.

7. Security Considerations

The security considerations of [BIER_ARCH], [BIER_ENCAPS], [RFC6513] and [RFC6514] are applicable.

8. References

8.1. Normative References

[BIER_ARCH] Wijnands, IJ., "Multicast using Bit Index Explicit Replication Architecture", internet-draft draft-wijnands-bier-architecture-02, December 2014.

[BIER_ENCAPS] Wijnands, IJ., "Multicast using Bit Index Explicit Replication Architecture", internet-draft draft-wijnands-mpls-bier-encapsulation-02, December 2014.

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

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

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, August 2008.

[RFC6513] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", RFC 6513, February 2012.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, February 2012.

[RFC6625] Rosen, E., Rekhter, Y., Hendrickx, W., and R. Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes", RFC 6625, May 2012.

8.2. Informative References

[EXTRANET] Rekhter, Y. and E. Rosen, "Extranet Multicast in BGP/IP MPLS VPNs", internet-draft draft-ietf-l3vpn-mvpn-extranet-05, July 2014.

[SEAMLESS_MCAST] Rekhter, Y., Aggarwal, R., Morin, T., Grosclaude, I., Leymann, N., and S. Saad, "Inter-Area P2MP Segmented LSPs", internet-draft draft-ietf-mpls-seamless-mcast-14, June 2014.

Authors' Addresses

Eric C. Rosen (editor)
Juniper Networks, Inc.
10 Technology Park Drive
Westford, Massachusetts 01886
US

Email: erosen@juniper.net

Mahesh Sivakumar
Cisco Systems, Inc.
510 McCarthy Blvd
Milpitas, California 95035
US

Email: masivaku@cisco.com

IJsbrand Wijnands
Cisco Systems, Inc.
De Kleetlaan 6a
Diegem 1831
BE

Email: ice@cisco.com

Sam K Aldrin
Huawei Technologies
2330 Central Express Way
Santa Clara, California
US

Email: aldrin.ietf@gmail.com

Andrew Dolganow
Alcatel-Lucent
600 March Rd.
Ottawa, Ontario K2K 2E6
CA

Email: andrew.dolganow@alcatel-lucent.com

Tony Przygienda
Ericsson
300 Holger Way
San Jose, California 95134
US

Email: antoni.przygienda@ericsson.com