

MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 28, 2008

S. Matsushima, Ed.
Softbank Telecom
T. Murakami
Cisco Systems
K. Nagami
INTEC Netcore
February 25, 2008

BGP Point to Multipoint LSP
draft-satoru-mpls-bgp-multipoint-05

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of 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/lid-abstracts.txt>.

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

This Internet-Draft will expire on August 28, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2008).

Abstract

This document describes motivations and use cases to P2MP extension of BGP. This extension will allow to make both Inter-AS and Intra-AS P2MP LSP that fit to some use cases then also clarify required features.

Internet-Draft

BGP P2MP LSP

February 2008

Table of Contents

1.	Introduction	3
2.	Motivations	3
2.1.	Inter-AS P2MP	3
2.2.	Intra-AS	3
3.	Use cases	4
3.1.	Inter-AS MVPN Option C	4
3.2.	Carriers' Carrier P2MP routing	4
3.3.	Over-layered Intra-AS P2MP-LSP	5
4.	BGP P2MP LSP	6
5.	Scalability Considerations	7
6.	Security Considerations	7
7.	Acknowledgments	7
8.	References	7
8.1.	Normative References	7
8.2.	Informative References	8
	Authors' Addresses	8
	Intellectual Property and Copyright Statements	10

Internet-Draft

BGP P2MP LSP

February 2008

[1.](#) Introduction

For some reasons, BGP [[1](#)] is useful protocols to create Inter-AS and Intra-AS LSPs by extending to carry MPLS label informations[2] . In the case of Point-to-Multipoint (P2MP) LSP, suitable motivations exist that using BGP to fit some use cases then we clarify required features for BGP P2MP. In each use case, we assume that inter-AS link and network uses MPLS as P2MP transport as well as intra-AS networks.

[2.](#) Motivations

This section describes motivations of P2MP extension to BGP.

[2.1.](#) Inter-AS P2MP

Some applications that uses P2MP LSP has appeared. Service provider is requesting the P2MP LSP connection among ASes. Inter-AS Option C of [RFC4364](#) [[3](#)] is used for a lot of service provider's Inter-AS connections. In this case, it is also necessary to achieve Option C in the Inter-AS Multicast VPN [[6](#)] environment. At this time, the P2MP LSP connection is needed between ASes. Moreover, P2MP LSP connection between PE and CE is preferable in the case of Carriers' Carrier network of [RFC4364](#) [[3](#)] as core network of VPN provider.

The method of using PCE for the P2MP LSP connection with Inter-AS is examined. PCE offers passing calculation information necessary to connect P2MP TE-LSP[7] in the Inter-AS environment but there is a case where TE-LSP cannot be connected by policy of Inter-AS. In this case, neither P2MP-TE nor PCE are applicable. On the other hand, BGP is generally used for Inter-AS in many cases then service provider may want to use BGP as P2MP connection as well. However, P2MP LSP cannot be done by BGP.

[2.2.](#) Intra-AS

As for P2MP LSP that uses RSVP-TE or MLDP, signaling is done with hop-by-hop manner. All LSRs are in Intra-AS that P2MP-LSP passes should support these signaling then makes efficient P2MP tree. Oppositely, there are a lot of demands of LSR that specializes P2MP function such as signaling and packet replication be isolated from unicast LSR to keep simple network even if it sacrifices efficiency.

In this case, Service Provider may want to establish P2MP LSPs among P2MP LSRs as multi-hop, over-lay manner. they need to decide which LSR become appropriate branch LSR of P2MP LSP and which is backup in policy basis. Hence they want to decide the branch LSR explicitly in

Intra-AS.

[3.](#) Use cases

This section describes some use case to find out required features to BGP support P2MP ability.

[3.1.](#) Inter-AS MVPN Option C

When a VPN traverse two or more AS composing of Inter-AS Option C, ASBRs should make LSP to PE that exists in adjacent other AS. Here, advertising routes related to VPN customer between ASes must be done only between RRs. This means that ASBR must be agnostic from VPN routes. Also ASBR should maintain P-Tunnel to PE in Multicast VPN.

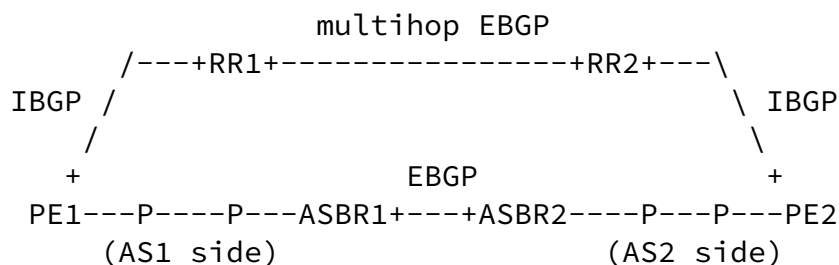


Figure 1: MVPN Option C

RRs exchanges routes related to Multicast VPN on multihop EBGP peer

in Figure 1 , and ASBR exchanges routes related to P-Tunnel. This route must not contain any VPN information. Routes that ASBR exchanges should have MPLS label that identifies P-Tunnel that should be distinct in each ASes. It is necessary to enhance BGP for EBGP between ASBR. Note that in the case of ingress replication P-Tunnel type is excepted. And whole of solution to achieve MVPN Option C is out of scope.

3.2. Carriers' Carrier P2MP routing

When a certain service provider is offering VPN MPLS backbone service by Carriers 'Carrier (CsC), service provider maintains only CE route in VPN. Therefore, CsC PE does not need to know VPN service routes. This means that CsC network can scale well.

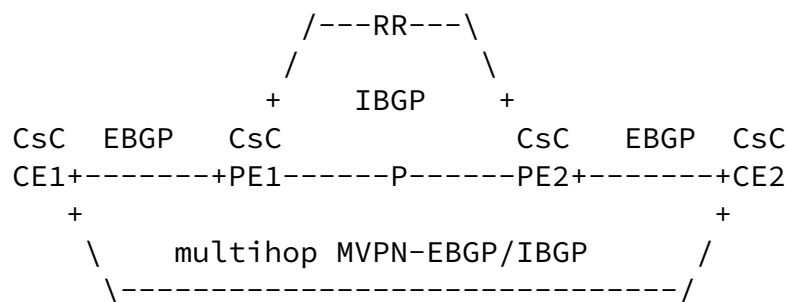


Figure 2: P2MP Carriers' Carrier

In CsC network shown in Figure 2, CsC-PEs should offer MPLS P2MP transport so that CsC CE may need MPLS P2MP connectivity with remote CE when CsC CE does Multicast service or Multicast VPN service. If CsC CE1 does source CE and MVPN-BGP[8], CsC CE1 advertises Auto-discovery route to CsC CE2 on MVPN-IBGP. At this time, it is necessary to deliver Tunnel identifier specified in PMSI Tunnel attribute to CsC-CE2 via CsC-PE1, RR, and CsC-PE2.

Similarly, it is necessary to deliver Tunnel identifier advertised with Leaf auto-discovery route on MVPN-IBGP to CsC-CE1 via CsC-PE2,

RR, and CsC-PE1 as for CsC-CE2. CsC PEs should bind P2MP tree of internal with LSP between CsC PE to CsC CEs. It is necessary to enhance BGP with these functions. Note that the case where CsC-CE1 and CsC-PE1 do ingress replication is excluded.

When CsC-PEs aggregate two or more CsC-CEs into a P2MP transport to improve scalability, leaf CsC-PE2 will receive even a packet not necessary as a result. De-multiplexer that takes out a necessary packet is needed from received packets. CsC-PE2 can forward packet to CsC-CE2 when received packet have label that local assigned to bind CsC-CE2. CsC-PE1 can forward packet when receive routes that include label to push onto shared P2MP transport. And then that label is bind with CsC-CE1, CsC-PE1 can forward packet that received from CsC-CE1 to CsC-PE2.

[3.3.](#) Over-layed Intra-AS P2MP-LSP

It might be difficult for SP that wants to keep simple own network in some cases to give backbone multicast and/or P2MP ability in existing MPLS network. As for such SP, the P2MP ability might be given to specific LSR, and may select way to separate service that needs packet replication from existing LSR though it somewhat sacrifices optimization of packet replication.

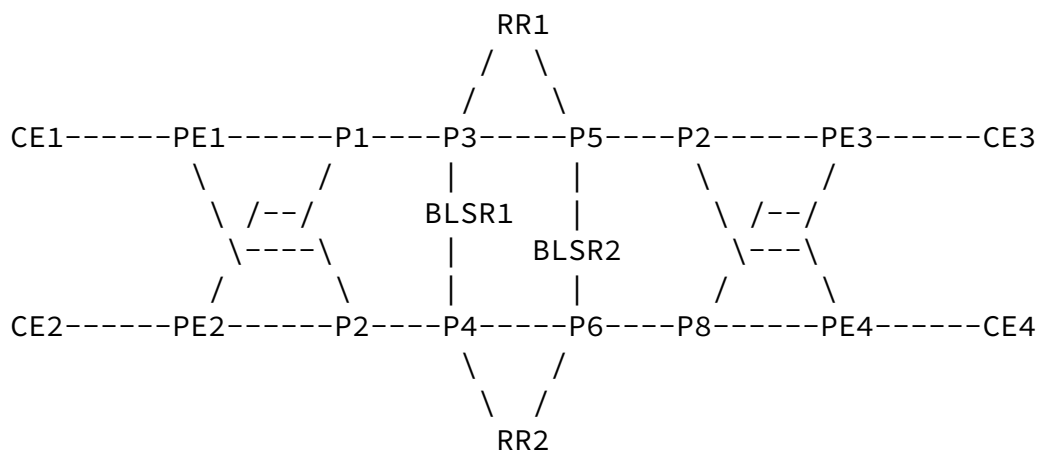


Figure 3: Over-layed Intra-AS P2MP LSPs

Network shown in Figure 3 starts offering CE1-CE4 Multicast VPN service by PE1-PE4, P1-P8, BLSR1, and 2 belonging and using P2MP LSP for same domain. In this network, only BLSR1 and 2 have P2MP branch capability, and P1-P8 doesn't have the P2MP ability. PE1-PE4 can become source and leaf of P2MP LSP. When PE1 becomes source PE and P2MP LSP is configured to PE3 and PE4, LSP is as follows.

```

PE1---->BLSR1---->PE3
      |
      +----->PE4

```

Unicast LSP overlays P2MP LSP between each PE and BLSR. For instance, PE1--> BLSR1 exists on unicast LSP that passes PE1->P1->P3->BLSR1. MPLS label that shows P2MP LSP becomes bottom stack label. If unicast LSP is TE-LSP, becomes possible to protection by Fast Reroute and to do the traffic engineering.

BGP may be used to establish P2MP LSP onto unicast LSP of each PE and BLSR in stack. It is necessary to enhance BGP for P2MP LSP. Because one unicast LSP can carry two or more P2MP LSP in stack, scalability of P router can be improved than hop-by-hop P2MP protocols. Moreover, it is possible to prepare for outage of P2MP node by policy based operation to which BLSR and/or source PE that becomes backup because it uses BGP are provided beforehand. Detail of BGP extensions are described with following section.

[4.](#) BGP P2MP LSP

Concrete P2MP informations that is carry by BGP and procedures are further study.

[5.](#) Scalability Considerations

The number of EBGp peering of ASBR and IBGP full mesh in AS give scalability impact even though BGP is originally excellent protocol in scalability. In use case described in [Section 3](#), Route reflector (RR) [\[4\]](#) be able to use instead of full mesh of IBGP. At that time, necessary routes must not be summarize by RR. Moreover, BGP maintains control plane only, and is not used for data plane.

As for BGP P2MP routes, the number of BGP P2MP routes may be increased by increasing LSPs and LSRs. To avoid or mitigate such increasing that may affect to scalability, RT Constraint [9] may appropriate for some cases.

6. Security Considerations

As well as [RFC3107](#) [2] describes, P2MP LSP using BGP has the "label spoofing" problem. For this, BGP LSR which also support P2MP LSP should not make the adjacencies of both trusted and untrusted system on the same interface.

7. Acknowledgments

Thanks to Miya Kohno and Paul Doolan for their support and helpful comments.

8. References

8.1. Normative References

- [1] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), January 2006.
- [2] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", [RFC 3107](#), May 2001.
- [3] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), February 2006.
- [4] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", [RFC 4456](#), April 2006.
- [5] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", [RFC 4360](#), February 2006.

8.2. Informative References

- [6] Aggarwal, R., Bandi, S., Cai, Y., Morin, T., Rekhter, Y., Rosen, E., Wijnands, I., and S. Yasukawa, "Multicast in MPLS/BGP IP VPNs", [draft-ietf-l3vpn-2547bis-mcast-06](#) (work in progress), January 2008.
- [7] Aggarwal, R., Papadimitriou, D., and S. Yasukawa, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", [RFC 4875](#), May 2007.
- [8] Aggarwal, R., "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", [draft-ietf-l3vpn-2547bis-mcast-bgp-04](#) (work in progress), November 2007.
- [9] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", [RFC 4684](#), November 2006.

Authors' Addresses

Satoru Matsushima (editor)
Softbank Telecom Corp.
9-1, Higashi-Shinbashi 1-chome
Minato-ku
Tokyo 105-7316
Japan

Email: satoru@ft.solteria.net

Tetsuya Murakami
Cisco Systems
2-1-1, Nishi-Shinjuku
Shinjuku-ku
Tokyo 163-0409
Japan

Email: tmurakam@cisco.com

Kenichi Nagami
INTEC Netcore, Inc.
1-3-3, Shin-suna
Koto-ku
Tokyo 135-0075
Japan

Email: nagami@inetcore.com

Internet-Draft

BGP P2MP LSP

February 2008

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at

ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

Matsushima, et al.

Expires August 28, 2008

[Page 10]