Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: July 2, 2012

X. Tu
Z. Fu
K. Li
ZTE Corporation

December 30, 2011

Resource Reservation Protocol - Traffic Engineering(RSVP-TE) Extensions for Inter-AS P2MP TE LSPs

draft-tfl-mpls-rsvp-te-inter-as-p2mp-00

Abstract

[RFC4875] describes extensions to RSVP-TE protocol for building a P2MP TE LSP in MPLS/GMPLS network environment.However, [RFC4875] doesn't specify path selecting problem of inter-AS P2MP TE LSP.This document specifies an inter-AS P2MP path computation method based on extentions to RSVP-TE Protocol.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on July 2, 2012.

Table of Contents

1. Introduction

[RFC4875] describes extensions to RSVP-TE protocol for building a P2MP TE LSP in MPLS/GMPLS network environment. However, [RFC4875] doesn't specify path selecting problem of inter-AS P2MP TE LSP. In inter-AS scenario, node maybe have not topology and resources of the whole network, so it may be not able to compute an inter-AS P2MP TE LSP.

This document specifies an inter-AS P2MP path bulding method based on extentions to RSVP-TE Protocol. Inter-AS P2MP LSP is comprised of multiple source-to-leaf(S2L) sub-LSPS, which are set up between the ingress and egress LSRs located in different ASes.

2. Terminology

AS: Autonomous System, A collection of connected routing prefixes under the control of one or more network operators that presents a common, clearly defined routing policy to the Internet.

ASBR: Autonomous System Border Router. Routers used to connect together ASes of the same or different service providers via one or more inter-AS links.

Ingress ASBR of AS(n): an ASBR connecting AS(n-1) to AS(n) on a path.

Egress ASBR of AS(n): an ASBR connecting AS(n) to AS(n+1) on a path.

3. Procedures Overview

3.1. Overview

This document specifies an inter-AS P2MP path bulding method using RSVP-TE Protocol Extension. Using this method, the ingress node will start the intra-AS P2MP TE LSP building process, and distribute the PATH message according to RSVP-TE protocol extension specified in [RFC4875]. Egress ASBRs will transmit the PATH request message of a P2MP TE LSP to the ingress ASBRs of the neighbor ASes, which will processes the message as an request for intra-AS P2MP TE LSP building request, and continuing the building process. Following text describes this mechanism.

1. Ingress node treats egress ASBRs and leaf nodes in the same AS as leaf nodes of the intra-AS P2MP tree, and computes P2MP path to those leaf nodes.. Since it is intra-AS path computation, current existing mature solutions could be used here.

2. Ingress node uses RSVP-TE protocol to construct Path messages and send them through the path that has been computed to leaf nodes in the same AS. The Path message should contain path information to leaf nodes and cost information of S2L sub-LSPs.

3. When egress ASBR node receives path request message, it should first decide whether or not cost value of path in the message is smaller than that exists on the node. If yes, egress ASBR node should save the Path message and update cost value, then transmit the Path message to ingress ASBR nodes of downstream neighbor AS. Otherwise, egress ASBR node should desert the Path message without any process.

4. When ingress ASBR node of neighbor AS receives Path message, it should first decide whether or not cost value of path in the message is smaller than that exists on the node. If yes, ingress ASBR node should save the later received path message and compute path to all egress ASBRs and leaf nodes in the same AS, after computation is finished, ingress node construct Path messages and send them through the path that has been computed to leaf nodes in the same AS. This is similar to step 1 and 2.

5. When egress node receives Path message, it should first decide whether or not cost value of path in the message is smaller than that exists on the node, if yes, then it should further decide the interface that received the newly path request message is not the same interface that has been used to reply, if yes, egress node should save the newly Path message, waits for a Delay time and then reserves resources and assigns label, replies Resv

message to corresponding upstream node.

6. Resv message passes through the path that Path message used, reaches ingress ASBR node and upstream neighbor AS's egress ASBR node, those ASBR nodes need reserve resources, assign label for upstream node, and transmit new Resv message to upstream node.

7. Ingress node receives Resv message, configures received label to the node. Then, an inter-AS S2L sub-LSP for one leaf node is set up.

3.2. Egress ASBR Node

When egress ASBR receivers PATH messages, it should process the message according to following procedures.

It should first decide whether or not the cost value of S2L sub-LSP in the messages is smaller than the pre-saved one. if yes, egress ASBR node should desert the Path message without any further process. Otherwise, egress ASBR must replace the pre-saved PATH request messages with the newly received messages. if links that connect to these ingress ASBR nodes could satisfy TE attributions, egress ASBR node will transmit Path message to these ingress ASBR nodes. The Path message treats egress ASBR node as root and ingress ASBR node of next neighbor AS as leaf node, and cost value of path in the Path message equals cost value of path in the Path message that egress ASBR node received plus cost value of the link between egress ASBR node and ingress ASBR node.

When Egress ASBR receives Resv message, it should process the message according to following procedures.

If it has replied the corresponding P2MP TE LSP, it only needs to configure label to the node. If not, it should also reserve resources and assign label, construct new Resv message and reply Resv message to corresponding upstream node. If this node receives more than one Resv messages, it should merge flow descriptors of those messages and reply single Resv message to corresponding upstream node.

3.3. Ingress ASBR Node

Ingress ASBR node receives Path message, it should process the message according to following procedures.

It should first decide whether or not cost value of path in the message isn't smaller than that exists on the node, if yes, ingress ASBR node should desert the path message without any further process.

Otherwise, ingress ASBR node should save the Path message. If this is first Path message received, ingress ASBR node would compute optimal multicast tree between it and every egress ASBR node or egress PE node in the same AS; if this is not the first message received, ingress ASBR node could use computation result which has been computed. Then, ingress ASBR node computes cost value of path between it and each egress ASBR node or egress PE node, constructs new Path message, sends the message to every egress ASBR node and egress PE node.

Ingress ASBR node receives Resv message, it should process the message according to following procedures.

If it has replied the corresponding P2MP TE LSP, it only needs to configure label to the node. If not, it should also reserve resources and assign label, construct new Resv message and reply Resv message to corresponding egress ASBR node of upstream neighbor AS. If this node receives more than one Resv messages, it should merge flow descriptors of those messages and reply single Resv message to corresponding egress ASBR node of upstream neighbor AS.

3.4. Transit Node

Transit node receives Path message, it should process the message according to following procedures.

It should first decide whether or not cost value of path in the message isn't smaller than that exists on the node, if yes, Transit node should desert the path message without any further process. Otherwise, transit node should update the message that saved on the node, and analyze the message and process it according to procedures defined in RFC4875, then transmit it to downstream node.

Transit node receives Resv message, it should process the message according to following procedures.

If it has replied the corresponding P2MP TE LSP, it only needs to configure label to the node. If not, it should also reserve resources and assign label, construct new Resv message and reply Resv message to corresponding upstream node. If this node receives more than one Resv messages, it should merge flow descriptors of those messages and reply single Resv message to corresponding upstream node.

3.5. Leaf Node

Leaf node receives Path message, it should process the message according to following procedures.

Leaf node should first decide whether or not this is the first message received, if yes, leaf node should start timer.

If this isn't the first message, and timer doesn't expire, leaf node should decide if cost value of path in the message is smaller than that exists on the node, if yes, leaf node should update the message that saved on the node. If cost value in the message is not smaller, leaf node should desert the message.

If this is not the first message but timer expires, leaf node should desert the message.

If timer expires, leaf node should reserve resources and assign label, and reply Resv message to upstream node.

4. Messages Format

4.1. Path Message

Path message is used to transmit path setup request, [RFC4875] gives a detailed definition. This document gives extensions to Path message, and Path message could also transmit cost or hops of path which is from root node to certain transit node or leaf node. The format of Path message is as follows.

```
  <Path Message> ::= <Common Header> [ <INTEGRITY> ]
                        [ [<MESSAGE_ID_ACK> | <MESSAGE_ID_NACK>] ...]
                        [ <MESSAGE_ID> ]
                        <SESSION> <RSVP_HOP>
                        <TIME_VALUES>
                        [ <EXPLICIT_ROUTE> ]
                        <LABEL_REQUEST>
                        [ <PROTECTION> ]
                        [ <LABEL_SET> ... ]
                        [ <SESSION_ATTRIBUTE> ]
                        [ <NOTIFY_REQUEST> ]
                        [ <ADMIN_STATUS> ]
                        [ <POLICY_DATA> ... ]
                        <sender descriptor>
                        [<S2L sub-LSP descriptor list>]

 <S2L sub-LSP descriptor list> ::= <S2L sub-LSP descriptor>
                        [ <S2L sub-LSP descriptor list> ]


 <S2L sub-LSP descriptor> ::= <S2L_SUB_LSP>
                              <S2L_SUB_LSP_COST>|     <S2L_SUB_LSP_HOPS>
                         [ <P2MP SECONDARY_EXPLICIT_ROUTE> ]
```

In each <S2L sub-LSP descriptor> object, there is a <S2L_SUB_LSP_COST> or <S2L_SUB_LSP_HOPS> object, they are used to describe cost of a path.

4.2. Resv message

Resv message is used to reply Path message and assign label for upstream node, [RFC4875] gives a detailed definition. This document gives extentions to Resv message, and Resv message could also transmit cost or hops of path which is from root node to leaf node. The format of Resv message is as follows.

```
      <Resv Message> ::=    <Common Header> [ <INTEGRITY> ]
                          [ [<MESSAGE_ID_ACK> | <MESSAGE_ID_NACK>] ... ]
                          [ <MESSAGE_ID> ]
                          <SESSION> <RSVP_HOP>
                          <TIME_VALUES>
                          [ <RESV_CONFIRM> ]  [ <SCOPE> ]
                          [ <NOTIFY_REQUEST> ]
                          [ <ADMIN_STATUS> ]
                          [ <POLICY_DATA> ... ]
                          <STYLE> <flow descriptor list>

   <flow descriptor list> ::= <FF flow descriptor list>
                          | <SE flow descriptor>

   <FF flow descriptor list> ::= <FF flow descriptor>
                          | <FF flow descriptor list>
                                  <FF flow descriptor>

   <SE flow descriptor> ::= <FLOWSPEC> <SE filter spec list>

   <SE filter spec list> ::= <SE filter spec>
                          | <SE filter spec list> <SE filter spec>

   <FF flow descriptor> ::= [ <FLOWSPEC> ] <FILTER_SPEC> <LABEL>
                          [ <RECORD_ROUTE> ]
                          [ <S2L sub-LSP flow descriptor list> ]

   <SE filter spec> ::= <FILTER_SPEC> <LABEL> [ <RECORD_ROUTE> ]
                          [ <S2L sub-LSP flow descriptor list> ]

   <S2L sub-LSP flow descriptor list> ::=
                          <S2L sub-LSP flow descriptor>
                          [ <S2L sub-LSP flow descriptor list> ]

   <S2L sub-LSP flow descriptor> ::= <S2L_SUB_LSP>
                            <S2L_SUB_LSP_COST>|<S2L_SUB_LSP_HOPS>
                          [ <P2MP_SECONDARY_RECORD_ROUTE> ]
```

The structure of <S2L sub-LSP flow descriptor> is similar to that of <S2L sub-LSP descriptor> ; <S2L_SUB_LSP_COST> and <S2L_SUB_LSP_HOPS> objects are used to describe cost of a path. The root of P2MP LSP could know cost value of path to each leaf node through those objects that contained in Resv message.

5. Object Format

5.1. S2L_SUB_LSP_COST Object

This object is used to describe cost value of a path. The format of this object is as follows.

S2L_SUB_LSP_ATTR Class = 55 S2L_SUB_LSP_COST C-Type = 1

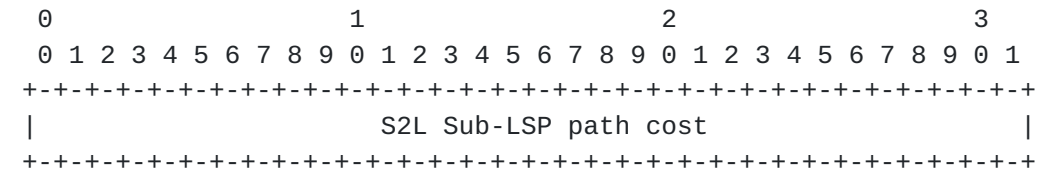

Figure 1: Inter-AS P2MP S2L_SUB_LSP_COST Object Format

5.2. S2L_SUB_LSP_HOPS Object

This object is used to describe hops of a path. The format of this object is as follows.

S2L_SUB_LSP_ATTR Class = 55 S2L_SUB_LSP_HOPS C-Type = 2

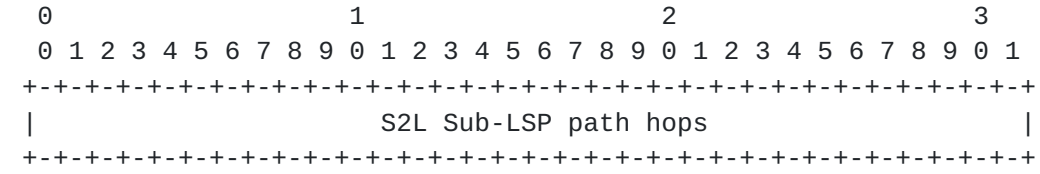

Figure 2: Inter-domain P2MP S2L_SUB_LSP_HOPS Object Format

6. Security Considerations

TBD

7. IANA Considerations

7.1. Class Numbers and C-Types

This document adds new class named S2L_SUB_LSP_ATTR, and class number is 55. This class have two C-Types. C-type 1 for cost value, C-Type 2 for hops value.

8. References

8.1. Normative References

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC4875] 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.

8.2. Informative References

[RFC6037] Rosen, E., Cai, Y., and IJ. Wijnands, "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs", RFC 6037, October 2010.

Authors' Addresses

XiaoPing Tu
ZTE Corporation
6F, RD Building 3, ZTE Industrical Park, XiLi LiuXian Road
Nanshan District, Shenzhen 518055
P.R.China

Phone: +86 755 26773624
Email: tu.xiaoping@zte.com.cn

ZhenTao Fu
ZTE Corporation
No.50, RuanJian Avenue
Yuhuatai District, Nanjing, Jiangsu
P.R.China

Phone: +86 25 88014227
Email: fu.zhentao@zte.com.cn

KeJun Li
ZTE Corporation
6F, RD Building 3, ZTE Industrical Park, XiLi LiuXian Road
Nanshan District, Shenzhen 518055
P.R.China

Phone: +86 755 26773611
Email: li.kejun@zte.com.cn