Network Working Group Z. Li
Internet-Draft Q. Zhao
Intended status: Informational Huawei Technologies
Expires: August 18, 2014 T. Yang
China Mobile
February 14, 2014

A Framework of MPLS Global Label
draft-li-mpls-global-label-framework-01

Abstract

The document defines the framework of MPLS global label including: representation of MPLS global label, process of control plane for MPLS global label, and process of data plane for MPLS global label.

Requirements Language

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].

Table of Contents

1. Introduction 2
2. Terminology 3
3. Representation of MPLS Global Label 3
  3.1. Option A -- Traditional MPLS Label 3
  3.2. Option B -- Expansions of MPLS Label 3
4. Control Plane 4
  4.1. Overview 4
    4.1.1. Shared MPLS Global Label Range Calculation 4
    4.1.2. MPLS Global Label Allocation 4
    4.1.3. MPLS Global Label Withdraw 5
    4.1.4. Error Process 5
    4.1.5. Redundancy 5
  4.2. BGP-Based Control Plane 6
  4.3. IGP-Based Control Plane 6
  4.4. PCE-Based Control Plane 8
5. IANA Considerations 9
6. Security Considerations 9
7. References 9
  7.1. Normative References 9
  7.2. Informative References 9
Authors' Addresses 10

1. Introduction

[I-D.li-mpls-global-label-usecases] proposes possible usecases of MPLS global label. MPLS global label can be used for identification of the location, the service and the network in different application scenarios. The new solutions based on MPLS global label can gain advantage over the existing solutions to facilitate service provisions.

This document defines the framework for MPLS global label. The framework includes the representation of MPLS global label, the process of control plane and data plane for MPLS global label.

2. Terminology

BDR: Backup Designated Router

DR: Designated Router

FEC: Forward Equivalence Class

MVPN: Multicast VPN

NBMA: Non-broadcast multi-access

PCE: Path Computation Element

PCC: Path Computation Client

RR: Route Reflector

3. Representation of MPLS Global Label

3.1. Option A -- Traditional MPLS Label

Current MPLS label uses 20 bits to represent the label range from 0 to 2^20 - 1. Since the existing MPLS label is always allocated locally, in order to guarantee a specific label is allocated globally, the available label values of the network nodes should be reported to a central control point. The central control point can calculate the COMMON label space which is available for all network nodes. Then the network nodes must reserve the common label space for the global label. When the global label is requested to allocate for specific service, the central control point can allocate the label from the common label space.

3.2. Option B -- Expansions of MPLS Label

[I-D.mpls-big-label-ucase-req] proposes the usecases and requirements for MPLS big label. It could also be a reasonable way to define a new label range or segment for MPLS global label which is independent from the existing MPLS label range. The label stack mechanism can be introduced to expand the MPLS label range. For example, the MPLS global label can be represented as the following figure. The global label value is achieved by adding the actual base label value indicated by the base label and the remainder label value.

```
0                                     32                                    64
+-------------------+----+-+---------+-------------------+----+-+---------+
|    Base    Label  | TC |S|   TTL   |  Remainder Label  | TC |S|   TTL   |
+-------------------+----+-+---------+-------------------+----+-+---------+
```

Figure 1 Representation of MPLS Global Label

If the new label range is used for the global label, it can reduce the possible confliction with the existing label range.

4. Control Plane

4.1. Overview

MPLS global label should be allocated concentratedly to guarantee all nodes can understand the same meaning for a specific global label. It should adopt a server/client model in the control plane for MPLS global label allocation. The procedures for the global label are described as follows.

4.1.1. Shared MPLS Global Label Range Calculation

1. Clients nodes should report MPLS global label capability to the central controller.

2. The central controller collects MPLS global label capability and MPLS global label range of all nodes. Then it can calculate the shared MPLS global label range for all nodes.

3. The centralized controller should notify the shared global label range to all client nodes. Client nodes reserve the shared global label range.

4.1.2. MPLS Global Label Allocation

1. The client node should send the global label request for specific usage to the central controller. FEC(Forward Equivalence Class) should be incorporated in the MPLS global label request message.

2. When the central controllers receives the MPLS global label request, it should allocate the label from the shared MPLS global label segment of all nodes.

3. The central controller sends the MPLS global label mapping message to all client nodes. Thus the MPLS global label for specific usage can be understood by all client nodes.

4. The client node receives the MPLS global label mapping message and installs the corresponding MPLS forwarding entry for the global label.

4.1.2.1. Process of Duplicate MPLS Global Label Request

Since MPLS global label is used for specific usage globally, it is possible that multiple MPLS global label requests for the same usage are sent by different client nodes to the central controller. The controller needs to count the MPLS global label requests for the same usage. It can send multiple global label mapping messages to respond these requests. It can also send only one copy of the global label mapping message to all nodes in order to reduce the unnecessary protocol operation. If the controller sends multiple copies of the global label mapping message to respond each label request, client nodes need to ignore the subsequent messages.

4.1.3. MPLS Global Label Withdraw

TBD.

4.1.4. Error Process

TBD.

4.1.5. Redundancy

Since MPLS global label is allocated concentratedly, it is important to guarantee the high availability of the central controller. Redundant backup needs to be introduced for the high availability. The backup central controller needs to learn global label requests sent by client nodes and corresponding label mapping sent by the primary central controller. The backup central controller will not send any global label mapping to client nodes until failure happens for the primary central controller.

The primary role and the backup role of the central controllers can be specified administratively. They can also be elected dynamically based on auto-discovery of these controllers.

The failure detection mechanism also needs to be introduced between the primary controller and the backup controller. It can be the keepalive-like mechanism, the fast detection mechanism like BFD, or mixing use of both mechanisms.

4.2. BGP-Based Control Plane

```
                     +---------------------+
                     |      IP/MPLS        |
                     |      Network        |
   +----+     +----+ |                     | +----+     +----+
   | CE1|-----|    | |                     | |    |-----| CE2|
   +----+\    | PE1|\|      +----------+    |/| PE3|     +----+
          \   +----+ \______|   BGP    |    / +----+
           \                |Controller|___/|
            \ +----+ /------|  (RR+)   |    |
             \|    |/|      +----------+    |
              | PE2| |                      |
              +----+ |                      |
                     +---------------------+
```

Figure 2: BGP-based Control Plane

Many types of services such as VPLS[RFC4761], Multicast VPN[RFC6514] and E-VPN[I-D.ietf-l2vpn-evpn] are based on MP-BGP. If new solutions based on MPLS global label are introduced for such services, the architecture shown in the figure 2 can be applied.

In the BGP-based control plane, Route Reflector (RR) of BGP [RFC4456] can act as the role of the central controller. We call this type of RR as RR+. For VPLS, Multicast VPN and E-VPN services, auto-discovery mechanisms based on MP-BGP are introduced. So the RR+ can learn the necessary membership information through these A-D routes. RR+ can also learn the MPLS label capability information through necessary MP-BGP extensions. When MPLS global label is necessary, the BGP client on the PE node can send label request to RR+ and the label mapping for the allocated MPLS global label will be advertised to all PEs. Thus all PEs can learn the binding between the service and the MPLS global and the forwarding entry for the MPLS global label can be installed accordingly.

4.3. IGP-Based Control Plane

```
+------------------------------+    +------------------------------+
|         IGP Domain 1         |    |         IGP Domain 2         |
|         +----------+         |    |         +----------+         |
|         |   IGP    |         |    |         |   IGP    |         |
|         |Controller|------------------------|Controller|         |
|         |  (DR+)   |         |    |         |  (DR+)   |         |
|         |          |         |    |         |          |         |
|         +----------+         |    |         +----------+         |
|          /         \         |    |          /         \         |
|         /           \        |    |         /           \        |
|        /             \       |    |        /             \       |
| +--------+        +--------+ |    | +--------+        +--------+ |
| | NODE 1 |        | NODE n | |    | | NODE 1 |        | NODE n | |
| |        | ...... |        | |    | |        | ...... |        | |
| |  IGP   |        |  IGP   | |    | |  IGP   |        |  IGP   | |
| | CLIENT |        | CLIENT | |    | | CLIENT |        | CLIENT | |
| +--------+        +--------+ |    | +--------+        +--------+ |
|                              |    |                              |
+------------------------------+    +------------------------------+
```

Figure 3: IGP-based Control Plane

If the internal nodes of the network support MPLS global label, IGP-based control plane can be introduced. IGP has ever introduced the DR(Designated Router) and BDR(Backup Designated Router) concepts for broadcast and NBMA network([RFC2328]). The Designated Router is elected in the broadcast or NBMA network to act as a centralized control point to advertise adjacencies among DR and DR others. In the IGP-base control plane for MPLS global label, we can adopt the DR concept which can act as the central controller for the MPLS global label. We called this type of DR ad DR+. The DR+ can collect the MPLS global label capability of all client nodes. If MPLS global label is necessary for specific usage, the MPLS global label will be allocated by the DR+ and the corresponding label mapping can be advertised to all network nodes through IGP extensions. Thus all network nodes in the IGP area can learn the label binding between the specific usage and the MPLS global label and the forwarding entry for the MPLS global label can be installed accordingly.

MPLS global label binding information should be always advertised in a specific IGP domains. There may be multiple IGP domains and nodes in other IGP domains may be necessary to learn the MPLS global label information. There are two possible solutions:

1. The global label information can be advertised by IGP to span multiple domains. It is like leaking the information from the native area to other areas.

2. There can exist direct connections between IGP DR+. The MPLS global label information can be advertised from the native IGP DR+ to the other IGP DR+ using possible protocol extensions other than IGP(e.g. PCEP extensions or BGP extensions). The other IGP DR+ can learn the MPLS global label information and advertise it in its own area through IGP extensions.

4.4. PCE-Based Control Plane

```
   +------------------------------+    +------------------------------+
   |         PCE DOMAIN 1         |    |         PCE DOMAIN 2         |
   |          +--------+          |    |          +--------+          |
   |          |        |          |    |          |        |          |
   |          |  PCE   |--------------------------|  PCE   |          |
   |          | Server |          |    |          | Server |          |
   |          |        |          |    |          |        |          |
   |          +--------+          |    |          +--------+          |
   |         /          \         |    |         /          \         |
   |        /            \        |    |        /            \        |
   |       /              \       |    |       /              \       |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   | | NODE 1 |        | NODE n | |    | | NODE 1 |        | NODE n | |
   | |        | ...... |        | |    | |        | ...... |        | |
   | |  PCC   |        |  PCC   | |    | |  PCC   |        |  PCC   | |
   | |        |        |        | |    | |        |        |        | |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   |                              |    |                              |
   +------------------------------+    +------------------------------+
```

Figure 4: PCE-based Control Plane

PCE[RFC4655] is another choice to implement the control plane for MPLS global label. The PCE servers can act as the role of the centralized controller and the PCC can act the role of the client for process of MPLS global label. PCE servers can collect the MPLS global label capability of all nodes through PCEP extensions. If MPLS global label is necessary for specific usage, the label request can be sent from PCC to PCE server. MPLS global label will be allocated by the PCE server and the corresponding label mapping will be advertised to all network nodes through PCEP extensions. Thus all network nodes in the domain can learn the label binding between the specific usage and the MPLS global label and the forwarding entry for the MPLS global label can be installed accordingly.

If MPLS global label information needs to be advertised in different domain, it can be advertised from the native PCE server to other PCE servers through PECP extensions. Then other PCE servers can advertise the MPLS global information to PCC through PCEP in its own domain.

5. IANA Considerations

This document makes no request of IANA.

6. Security Considerations

TBD.

7. References

7.1. Normative References

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

7.2. Informative References

[I-D.ietf-l2vpn-evpn]
Sajassi, A., Aggarwal, R., Henderickx, W., Isaac, A., and J. Uttaro, "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-05 (work in progress), February 2014.

[I-D.li-mpls-global-label-usecases]
Li, Z., Zhao, Q., and T. Yang, "Usecases of MPLS Global Label", draft-li-mpls-global-label-usecases-01 (work in progress), February 2014.

[I-D.mpls-big-label-ucase-req]
Li, R. and K. Zhao, "MPLS Big Label Usecases and Requirements", draft-mpls-big-label-ucase-req-00 (work in progress), October 2013.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, April 2006.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

[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.

Authors' Addresses

Zhenbin Li
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: lizhenbin@huawei.com

Quintin Zhao
Huawei Technologies
125 Nagog Technology Park
Acton, MA 01719
US

Email: quintin.zhao@huawei.com

Tianle Yang
China Mobile
32, Xuanwumenxi Ave.
Beijing 01719
China

Email: yangtianle@chinamobile.com