

Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 9, 2015

Z. Li
Z. Zhuang
Huawei Technologies
March 8, 2015

BGP Extensions for Service-Oriented MPLS Path Programming (MPP)
draft-li-idr-mpls-path-programming-01

Abstract

Service-oriented MPLS programming is to provide customized service process based on flexible label combinations. BGP will play an important role for MPLS path programming to allocate MPLS segment, download programmed MPLS path and the mapping of the service path to the transport path. This document defines BGP extensions to support service-oriented MPLS path programming.

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].

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 September 9, 2015.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Terminology	2
3. MPLS Segment Allocations	3
4. Download of MPLS Path	3
5. Download of Mapping of Service Path to Transport Path	5
5.1. Extended Unicast Tunnel Attributes	6
5.2. Extended PMSI Tunnel Attribute	7
6. Route Flag Extended Community	9
7. Destination Node Attribute	9
8. Capability Negotiation	10
9. IANA Considerations	11
10. Security Considerations	11
11. References	11
11.1. Normative References	11
11.2. Informative References	12
Authors' Addresses	12

1. Introduction

Service-oriented MPLS programming proposed by [I-D.li-spring-mpls-path-programming] is to provide customized service process based on flexible label combinations. BGP will play an important role for MPLS path programming to allocate MPLS segment, download programmed MPLS path and the mapping of the service path to the transport path. This document defines BGP extensions to support service-oriented MPLS path programming.

2. Terminology

BGP: Border Gateway Protocol

EVPN: Ethernet VPN

L2VPN: Layer 2 VPN

L3VPN: Layer 3 VPN

MPP: MPLS Path Programming

MVPN: Multicast VPN

RR: Route Reflector

SDN: Software-Defined Network

S-EVPN: Segment-based EVPN

SR-Path: Segment Routing Path

NLRI: Network Layer Reachability Information

3. MPLS Segment Allocations

MPLS Segment is the component to compose the MPLS path. [I-D.li-spring-mpls-path-programming] proposes the use cases for service-oriented MPLS path programming which needs following MPLS segments:

1. MPLS path programming for unicast service
 - o MPLS Segment for VPN identification
 - o MPLS Segment for ECMP
 - o MPLS Segment for OAM (Source identification)
 - o MPLS Segment for Traffic Steering
2. MPLS path programming for multicast service
 - o MPLS Segment for MVPN identification
 - o MPLS Segment for Source identification
3. MPLS virtual network for services
 - o MPLS Segment for MPLS virtual network

These MPLS Segments are defined in individual drafts. It is out of the scope of this document.

4. Download of MPLS Path

According to the service requirements, the central controller can combine MPLS segments flexibly. Then it can download the service label combination for specific prefix related with the service.BGP

extensions are necessary to advertise label stacks for prefix in NLRI field.

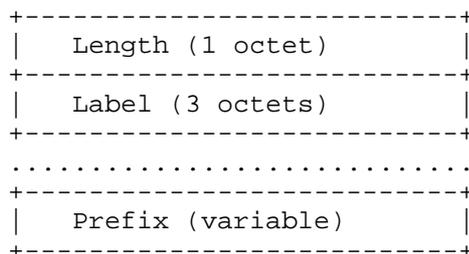


Figure 1: NLRI Definition in RFC3107

[RFC3107] defines above NLRI to advertise label binding for specific prefix. The label field can carry one or more labels. Each label is encoded as 3 octets, where the high-order 20 bits contain the label value, and the low order bit contains "Bottom of Stack". But for other AFI/SAFIs using label binding such as VPNv4, VPNv6, EVPN, MVPN, etc., it dose not support the capability to carry more labels for the specific prefix. Moreover for the AFI/SAFIs which do not support label binding capability originally, but may possibly adopt MPLS path programming now, there is no label field in the NLRI. In order to support flexible MPLS path programming, this document defines and uses a new BGP attribute called the "Extended Label attribute". This is an optional transitive BGP attribute. The format of this attribute is defined as follows:

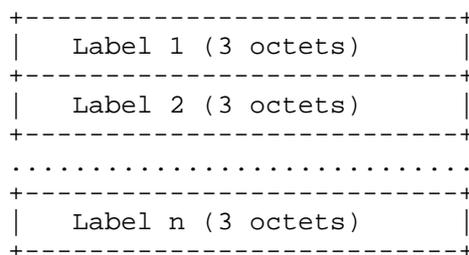


Figure 2: Extended Label Attribute

The Label field carries one or more labels (that corresponds to the stack of labels [[RFC3032]]). Each label is encoded as 3 octets, where the high-order 20 bits contain the label value, and the low order bit contains "Bottom of Stack" (as defined in [[RFC3032]]).

The Central Controller for MPLS path programming could build a route with Extended Label attribute and send it to the ingress routers.

Upon receiving such a route from the MPP Controller, the ingress router SHOULD select such a route as the best path. If a packet comes into the ingress router and uses such a path, the ingress router will encapsulate the stack of labels which gets from the Extended Label Attribute of the route into the packet and forward the packet along the path.

The "Extended Label attribute" can be used for various BGP address families. Before using this attribute, firstly, it is necessary to negotiate the capability between two nodes to support MPLS path programming for a specific BGP address family. If negotiation fails, a node MUST NOT send this attribute and MUST discard this attribute when it receives.

5. Download of Mapping of Service Path to Transport Path

Since the transport path is also to satisfy the service bearing the requirement, it can also be programmed according to traffic engineering requirements of service. Or the transport path can be set up according to general traffic engineering requirements. Then there needs to be implements the mapping of the service path to the transport path. BGP Extensions is necessary that through the community attribute of BGP, the identifier of the transport path can be carried when distribute label stack for a specific prefix.

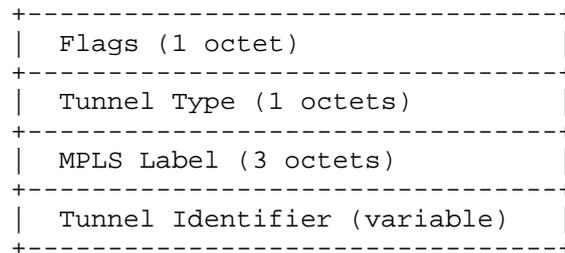


Figure 3: PMSI Tunnel Attribute in RFC6514

[RFC6514] defines the "P-Multicast Service Interface Tunnel (PMSI Tunnel) attribute". It is shown in the above figure. Since the attribute is always for the specific usage in BGP-based MVPN, it can not be applied to all possible use cases of service-oriented MPLS path programming. This document accordingly defines two new types of BGP attribute for both usage of unicast service path and the multicast service path: Extended Unicast Tunnel Attribute and Extended PMSI Tunnel Attribute.

5.1. Extended Unicast Tunnel Attributes

This document defines and uses a new BGP attribute called the "Extended Unicast Tunnel attribute". This is an optional transitive BGP attribute. The format of this attribute is defined as follows:

```

+-----+
| Flags (1 octet) |
+-----+
| Tunnel Type (1 octets) |
+-----+
| Tunnel Identifier (variable) |
+-----+
| Tunnel Specific Attributes (Variable)(Optional) |
+-----+

```

This document defines the following flags:

```

0 1 2 3 4 5 6 7
+++++
| reserved |S|
+++++

```

+ Unicast Tunnel Setup Required (S)

If the S flag is not set, the client node is just to map the service path to the corresponding tunnel. If the S flag is set, the client node MUST set up the tunnel according to the tunnel identifier and the tunnel specific attribute firstly. Then it maps the service path to the corresponding tunnel.

The Tunnel Type identifies the type of the tunneling technology used for the unicast service path. The type determines the syntax and semantics of the Tunnel Identifier field. This document defines the following Tunnel Types:

```

+ 0 - No tunnel information present
+ 1 - RSVP-TE LSP
+ 2 - LDP LSP
+ 3 - GRE Tunnel
+ 4 - MPLS-based Segment Routing Best-effort Path
+ 5 - MPLS-based Segment Routing Traffic Engineering Path

```

Tunnel Specific Attributes contains the attributes of the tunnel. The field is optional. The value depends on the tunnel type. It will be defined in the future versions.

When the Tunnel Type is set to "No tunnel information present", the Tunnel attribute carries no tunnel information (no Tunnel Identifier). when the type is used, the tunnel used for the service path is determined by the ingress router.

When the Tunnel Type is set to RSVP - Traffic Engineering (RSVP-TE) Label Switched Path (LSP), the Tunnel Identifier is <C-Type, Tunnel Sender Address, Tunnel ID, Tunnel End-point Address> as specified in [RFC3209]. If C-Type = 7, Tunnel Sender Address and Tunnel End-point Address are IPv4 address in 4 octets. If C-Type = 8, Tunnel Sender Address and Tunnel End-point Address are IPv6 address in 16 octets. The other fields in the RSVP-TE LSP Identifier are the same as specified in [RFC3209].

When the Tunnel Type is set to LDP LSP, the Tunnel Identifier is <Ingress Router's IP Address, Address Family, Prefix Length, Prefix> as specified in [RFC5036].

When the Tunnel Type is set to GRE Tunnel, the Tunnel Identifier is <Ingress Router's IP Address, Address Family, Source IP Address, Destination IP Address>.

When the Tunnel Type is set to MPLS-based Segment Routing Best-effort Path, the Tunnel Identifier is <Ingress Router's IP Address, Address Family, Destination Address>. When the ingress router receives a BGP route with MPLS-based Segment Routing Path Tunnel Identifier in the Extended Unicast Tunnel attribute, it will find the best-effort SR-path based on the destination address.

When the Tunnel Type is set to MPLS-based Segment Routing Traffic Engineering Path, the Tunnel Identifier is <C-Type, Tunnel Sender Address, Tunnel ID, Tunnel End-point Address>. If C-Type = 7, Tunnel Sender Address and Tunnel End-point Address are IPv4 address in 4 octets. If C-Type = 8, Tunnel Sender Address and Tunnel End-point Address are IPv6 address in 16 octets. The tunnel identifier is similar as that of RSVP-TE LSP.

5.2. Extended PMSI Tunnel Attribute

This document defines and uses a new BGP attribute called the "Extended PMSI Tunnel attribute". This is an optional transitive BGP attribute. The format of this attribute is defined as follows:

```

+-----+
| Flags (1 octet) |
+-----+
| Tunnel Type (1 octets) |
+-----+
| Tunnel Identifier (variable) |
+-----+
| Tunnel Specific Attributes (Variable)(Optional) |
+-----+

```

This document defines the following flags:

```

0 1 2 3 4 5 6 7
+-----+
| reserved |S|
+-----+

```

+ PMSI Tunnel Setup Required (S)

If the S flag is not set, the client node is just to map the service path to the corresponding tunnel. If the S flag is set, the client node MUST set up the tunnel according to the tunnel identifier and the tunnel specific attribute firstly. Then it maps the service path to the corresponding tunnel.

The Tunnel Type identifies the type of the tunneling technology used for the multicast service path. The type determines the syntax and semantics of the Tunnel Identifier field. This document defines the following Tunnel Types:

```

+ 0 - No tunnel information present
+ 1 - RSVP-TE P2MP LSP
+ 2 - mLDP P2MP LSP
+ 3 - PIM-SSM Tree
+ 4 - PIM-SM Tree
+ 5 - BIDIR-PIM Tree
+ 6 - Ingress Replication
+ 7 - mLDP MP2MP LSP

```

Tunnel Identifier: The definition of Tunnel Identifier is the same as those specified in section 5 of [RFC6514].

Tunnel Specific Attributes contains the attributes of the PMSI tunnel. The field is optional. The value depends on the PMSI tunnel type. It will be defined in the future versions.

6. Route Flag Extended Community

This document defines and uses a new BGP Extended Community called as the "Route Flag Extended Community" which Type value is to be assigned by IANA.

The Route Flag Extended Community is used to carry the flag appointed by a BGP route server (e.g., a central controller).

The format of this extended community is defined as follows:

0	1	2	3	4	5	6	7
Type		Reserved				Flag	

Flag = 0, Treat as normal route

Flag = 1, Treat as best route

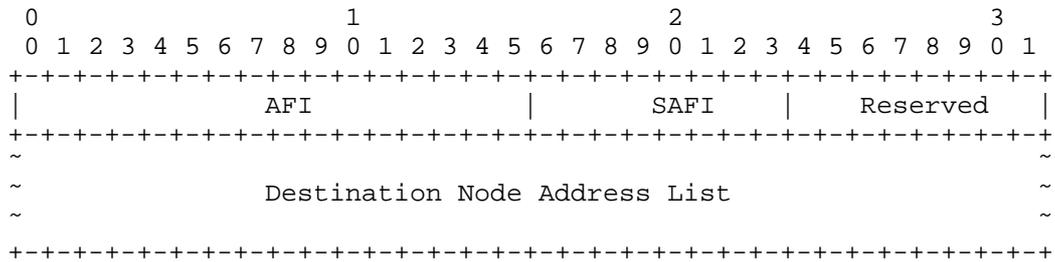
When a router receives a BGP route with a Route Flag Extended Community and the Flag set to "1", it SHOULD use the route as the best route when select the route from multiple routes for a specific prefix.

7. Destination Node Attribute

This document defines and uses a new BGP attribute called as the "Destination Node attribute" which Type value is to be assigned by IANA. The Destination Node attribute is an optional non-transitive attribute that can be applied to any address family.

The Destination Node attribute is used to carry a list of node addresses, which are intended to be used to determine the nodes where the route with such attribute SHOULD be considered. If a node receives a BGP route with a Destination Node attribute, it MUST check the node address list. If one address of the list belongs to this node, the route MUST be used in this node. Otherwise the route MUST be ignored silently.

The format of this attribute is defined as follows:



AFI: Address Family Identifier (16 bits).

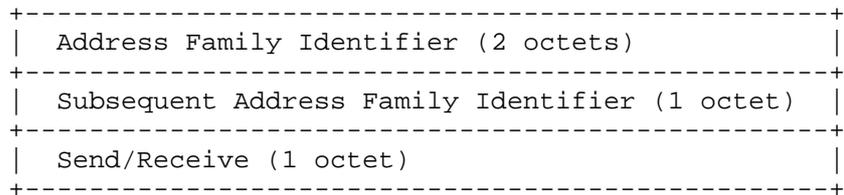
SAFI: Subsequent Address Family Identifier (8 bits).

Reserved: One octet reserved for special flags

Destination Node Address List: The list of IPv4 (AFI=1) or IPv6 (AFI=2) address.

8. Capability Negotiation

It is necessary to negotiate the capability to support MPLS path programming. The MPLS-Path-Programming Capability is a new BGP capability [RFC5492]. The Capability Code for this capability is to be specified by the IANA. The Capability Length field of this capability is variable. The Capability Value field consists of one or more of the following tuples:



The meaning and use of the fields are as follows:

Address Family Identifier (AFI): This field is the same as the one used in [RFC4760].

Subsequent Address Family Identifier (SAFI): This field is the same as the one used in [RFC4760].

Send/Receive: This field indicates whether the sender is (a) willing to receive programming MPLS paths from its peer (value 1), (b) would

like to send programming MPLS paths to its peer (value 2), or (c) both (value 3) for the <AFI, SAFI>.

9. IANA Considerations

TBD.

10. Security Considerations

TBD.

11. References

11.1. Normative References

[I-D.li-spring-mpls-path-programming]

Li, Z., "Use Cases and Framework of Service-Oriented MPLS Path Programming (MPP)", draft-li-spring-mpls-path-programming-00 (work in progress), July 2014.

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

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[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.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", RFC 5036, October 2007.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, February 2009.

11.2. Informative References

- [RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, May 2001.
- [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

Shunwan Zhuang
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: zhuangshunwan@huawei.com