

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

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

Use Cases and Framework of Service-Oriented MPLS Path Programming (MPP)
draft-li-spring-mpls-path-programming-01

Abstract

Source Packet Routing in Networking (SPRING) architecture for unicast traffic has been proposed to cope with the use cases in traffic engineering, fast re-reroute, service chain, etc. It can leverage existing MPLS dataplane without any modification. In fact, the label stack capability in MPLS would have been utilized well to implement flexible path programming to satisfy all kinds of requirements of service bearing. But in the distributed environment, the flexible programming capability is difficult to implement and always confined to reachability. As the introducing of central control in the network, the flexible MPLS programming capability becomes possible owing to two factors: 1. It becomes easier to allocate label for more purposes than reachability; 2. It is easy to calculate the MPLS path in a global network view. Moreover, the MPLS path programming capability can be utilized to satisfy more requirements of service bearing in the service layer which is defined as service-oriented MPLS path programming. This document defines the concept of MPLS path programming, then proposes use cases, architecture and protocol extension requirements in the service layer for the architecture of Source Packet Routing in Networking (SPRING).

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	3
2. Terminology	3
3. Programming Capability of MPLS Path	4
3.1. History Review	4
3.2. Gap Analysis of Segment Routing	5
4. Use Cases of Service-Oriented MPLS Path Programming	6
4.1. Use Cases for Unicast Service	6
4.1.1. Basic Reachability	6
4.1.2. VPN Identification	6
4.1.3. ECMP(Equal Cost Multi-Path)	6
4.1.4. Service OAM	7
4.1.5. Traffic Steering	7
4.2. Use Cases of Multicast Service	7
4.2.1. Basic Reachability	8
4.2.2. MVPN Identification	8
4.2.3. Source Identification	8
4.3. Use Cases of MPLS Virtual Network	8
4.4. Use cases of More Label Combinations	9
4.5. Use cases for Centralized Mapping of Service to Tunnels	9
5. Framework of Service-Oriented MPLS Path Programming	10
5.1. Central Control for MPLS Path Programming	10
5.2. BGP-based MPLS Segment Distribution	11
5.3. MPLS Service Path Programming	11
5.3.1. Label Combination and Download of MPLS Path	11

5.3.2. Mapping of Service Path to Service Path 12

5.4. Compatibility 12

5.5. Protocol Extensions Requirements 12

5.5.1. BGP 12

5.5.2. I2RS 13

6. IANA Considerations 13

7. Security Considerations 13

8. References 13

8.1. Normative References 13

8.2. Informative References 13

Authors' Addresses 15

1. Introduction

Source Packet Routing in Networking (SPRING) architecture for unicast traffic has been proposed to cope with the use cases in traffic engineering, fast re-reroute, service chain, etc. It can leverage existing MPLS dataplane without any modification. In fact, the label stack capability in MPLS would have been utilized well to implement flexible path programming to satisfy all kinds of requirements of service bearing. But in the distributed environment, the flexible programming capability is difficult to implement and always confined to reachability. As the introducing of central control in the network, the flexible MPLS programming capability becomes possible owing to two factors: 1. It becomes easier to allocate label for more purposes than reachability; 2. It is easy to calculate the MPLS path in a global network view. Moreover, the MPLS path programming capability can be utilized to satisfy more requirements of service bearing in the service layer which is defined as service-oriented MPLS path programming. This document defines the concept of MPLS path programming, then proposes use cases, architecture and protocol extension requirements in the service layer for the architecture of Source Packet Routing in Networking (SPRING).

2. Terminology

BGP: Border Gateway Protocol

BUM: Broadcast, Unknown unicast and Multicast

EVPN: Ethernet VPN

FRR: Fast Re-Route

L2VPN: Layer 2 VPN

L3VPN: Layer 3 VPN

MPP: MPLS Path Programming

MVPN: Multicast VPN

RR: Route Reflector

SDN: Software-Defined Network

SR-path: Segment Routing Path

3. Programming Capability of MPLS Path

MPLS path is composed by label stacks. Since in the label stack the labels in different layers can represent different meaning and the depth of the label stack can be unlimited in theory, it is possible can make up all kinks of MPLS paths based on the combination of labels. If we look on the combination of MPLS labels as programming, it is can be seen that the MPLS path has high programming capability.

3.1. History Review

The solutions based on MPLS label stack have been widely deployed. For example, in the scenario of Options C inter-AS VPN ([RFC4364]), we assume that LDP over TE is used as the transport tunnel and the TE tunnel starts at the ingress PE, following label stack can be composed by the ingress PE for MPLS path to bear VPN service:

```
+-----+-----+-----+-----+
|VPN Prefix|  BGP  |  LDP  | RSVP-TE |
|  Label  | Label | Label | Label  |
+-----+-----+-----+-----+
```

If facility FRR ([RFC4090]) is deployed for the MPLS TE tunnel, once the failure happens, additional label will be pushed for the label stack which is shown as follows:

```
+-----+-----+-----+-----+-----+
|VPN Prefix|  BGP  |  LDP  | RSVP-TE | BYPASS FRR |
|  Label  | Label | Label | Label  | Label  |
+-----+-----+-----+-----+-----+
```

The combination of labels in the above label stack is not simpler than the existing segment routing solution which composes the segment routing path through combination of segments. In fact, this is also a use case of source packet routing. But the combination is not as flexible as the segment routing since the combination of labels is always to cope with the reachability issue with limited capability in the distributed environment as follows:

1. Each label in the label stack is always binded with the reachability to a specific prefix. That is, the purpose of the label binding is limited.
2. It is difficult to implement flexible path calculation based on policy or constraints. For example, MPLS TE proposes rich set of traffic engineering attributes for transport. But it needs complex configurations in each ingress node in an unscalable way. That is, the path calculation and composition capability is limited.

As more concepts on MPLS label are proposed such as entropy label, source label, segment routing, etc., the purpose of label binding expands and the combination of labels can become more flexible. MPLS path programming capability becomes more realistic to satisfy more application scenarios.

3.2. Gap Analysis of Segment Routing

Segment Routing ([I-D.filsfils-spring-segment-routing]) is a typical example of MPLS path programming. The segment based on MPLS label is to represent nodes or agencies in the network. Through the collected information of network segments and path calculation based on the service requirement in the central controller, there will be flexible segment routing paths for the usage of traffic engineering. The SR-path can be advertised to the ingress node through PCE extensions. ([I-D.sivabalan-pce-segment-routing]).

Segment routing can implement source packet routing with high flexibility. On the other hand, there are multiple layers for MPLS path to bear services which is shown in the following figure:

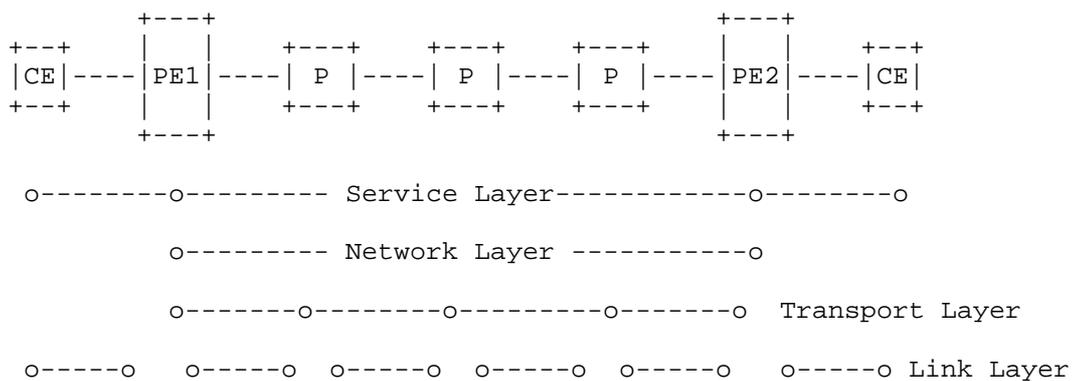


Figure 1: Multiple Layers of Service Bearing

Now the segment routing is to provide the source packet routing in the transport layer. We can call this type of source packet routing as Transport-Oriented MPLS path programming. There will be more application scenarios which needs the source packet routing in the service layer and network layer. We call these types of source packet routing as Service-Oriented MPLS path programming.

4. Use Cases of Service-Oriented MPLS Path Programming

4.1. Use Cases for Unicast Service

4.1.1. Basic Reachability

The basic reachability for VPN service is to allocate label to specific prefix including IP address or MAC address. MPLS path is as follows (using L3VPN as the example):

```
+-----+
|VPN Prefix| ---> Transport
|  Label   |      Tunnel
+-----+
```

4.1.2. VPN Identification

There are several use cases which need to indentify the VPN the packet belongs to in the forwarding plane such as the egress PE node protection for VPN ([I-D.zhang-l3vpn-label-sharing]). MPLS Path can be as follows:

```
+-----+-----+
|VPN Prefix|  VPN  | ---> Transport
|  Label   |  Label |      Tunnel
+-----+-----+
```

4.1.3. ECMP(Equal Cost Multi-Path)

In order to satisfy ECMP to take full advantage of link bandwidth in the network, the entropy label ([RFC6790]) can be encapsulated. MPLS path can be as follows:

```
+-----+-----+-----+
| Entropy |VPN Prefix|  VPN  | ---> Transport
|  Label  |  Label  |  Label |      Tunnel
+-----+-----+-----+
```

4.1.4. Service OAM

OAM is an important requirement for the service. The performance metrics should be measured against the Service Level Agreement(SLA) for the user. Now there are relatively complete and mature OAM mechanism for the point-to-point service. But for LDP LSP, owing to the MP2P model it is difficult to identify the flow from a specific PE based on the label. Source label has been proposed as a possible solution([I-D.chen-mpls-source-label]). When the source label is applied, MPLS path can be as follows:

```

+-----+-----+-----+-----+
| Entropy | VPN Prefix|   VPN   | Source   | ---> Transport
|  Label  |   Label   |  Label  |  Label   |           Tunnel
+-----+-----+-----+-----+

```

4.1.5. Traffic Steering

Service traffic may span multiple ASes. It is an important use case to steer traffic at ASBR in an AS to specific ASBR in neighboring AS. There are possible solutions for this type of traffic steering:

1. Traffic Steering based on Transport Tunnel

This method looks on the segment between two ASBRs as the extension of the transport tunnel in an AS. It can steer the traffic through the specific path to the neighboring AS.

2. Traffic Steering in Service/Network Layer

This method is to directly encapsulate the service flow with the steering label in the ingress PE before it enters into the transport tunnel. [I-D.filsfils-spring-segment-routing-central-epe] illustrates the application of Segment Routing to solve the Egress Peer Engineering (EPE) requirement. When this method is applied, the MPLS path can be as follows:

```

+-----+-----+-----+-----+-----+
| Entropy | Steering | VPN Prefix|   VPN   | Source   | ---> Transport
|  Label  |   Label  |   Label   |  Label  |  Label   |           Tunnel
+-----+-----+-----+-----+-----+

```

4.2. Use Cases of Multicast Service

4.2.1. Basic Reachability

When MPLS multicast tunnel is applied for the multicast service in BGP-based MVPN, VPLS or EVPN, the basic MPLS path can be as follows:

```
+-----+
| Multicast |---> Transport
| Payload   |      Multicast Tunnel
+-----+
```

4.2.2. MVPN Identification

When multiple MVPNs shares the MPLS multicast tunnel, it is necessary to encapsulate the label to identify specific MVPN([RFC6514]). The MPLS path can be as follows:

```
+-----+-----+
| Multicast | MVPN   | ---> Transport
| Payload   | Label  |      Multicast Tunnel
+-----+-----+
```

4.2.3. Source Identification

In order to implement the split horizon or C-MAC learning in the forwarding plane when MPLS multicast is to bear BUM traffic in L2VPN, it is necessary to introduce the label to identify the source of the BUM traffic([I-D.li-l2vpn-segment-evpn]). The MPLS path is as follows:

```
+-----+-----+-----+
| Multicast | EVPN   | Source | ---> Transport
| Payload   | Label  | Label  |      Multicast Tunnel
+-----+-----+-----+
```

4.3. Use Cases of MPLS Virtual Network

The framework of MPLS virtual network has been proposed in [I-D.li-mpls-network-virtualization-framework]. When the unicast service or the multicast service enters into the transport tunnel, it may take different MPLS virtual network identified by the MPLS label for the purpose of QoS routing, security or virtual operations. The MPLS path is as follows:

```
+-----+-----+-----+
| Service   | ---> | Virtual Network | ---> Transport
| Label Stack |      | Label           |      Tunnel
+-----+-----+-----+
```

4.4. Use cases of More Label Combinations

Service-oriented MPLS path programming can make full use of flexible combination of MPLS labels to satisfy different requirements for the service flow. Based on the above proposed use cases, MPLS path can be composed adopting part or whole labels for these use cases based on the service requirement. Besides this, more flexible MPLS label combination may be provided:

1. Hierarchical process or multiple repeated process: The label for the same usage can exist in different layers. Or the process identified by the label can exist in multiple nodes along the path. Then the labels for the same usage can be encapsulated several times in the label stack. The encapsulation can be as follows (using SERVICE LABEL to identify the label for the same service process in different layers):

SERVICE LABEL	VPN Prefix Label	SERVICE LABEL	VPN Label	SERVICE LABEL	Tunnel Label
---------------	------------------	---------------	-----------	---------------	--------------

2. Special-purpose label indicator: Since the label in the service-oriented MPLS programming is for special-purpose process, it may need a special purpose label to indicate the usage of the label followed the special-purpose labels. For example, the ELI(Entropy Label Indicator) is introduced for the entropy label. This may introduce more labels for the combination.

This document is not to define all possible use cases for the service-oriented path programming. The new use cases can be defined in the future independent document.

4.5. Use cases for Centralized Mapping of Service to Tunnels

In the transport layers, there can be multiple tunnels to one specific destination which satisfy different constraints. In the traditional way, the tunnel is set up by the distributed forwarding nodes. As the PCE-initiated LSP setup [I-D.ietf-pce-pce-initiated-lsp] is introduced, the tunnel can be setup in the central controlled way. In order to satisfy the different service requirements, it is necessary to provide the capability to flexibly map the service to different tunnels. Since the central control point has enough information based on the whole network view, it can be an effective way to map the service to the tunnel by the central point and advertise the mapping information to the end-points of the service to guide the mapping in the forwarding node.

5. Framework of Service-Oriented MPLS Path Programming

5.1. Central Control for MPLS Path Programming

Central control plays an important role in MPLS path programming. It can extend the MPLS path programming capability easily. There are two important functionalities for the central control:

1. Central controlled MPLS label allocation: Label can be allocated centrally for special usage other than reachability. These labels can be used to compose MPLS path. We call it as MPLS Segment.
2. Central controlled MPLS path programming: Central controller can calculate path in a global network view and implement the MPLS path programming based on the collected information of MPLS segments to satisfy different requirements of services.

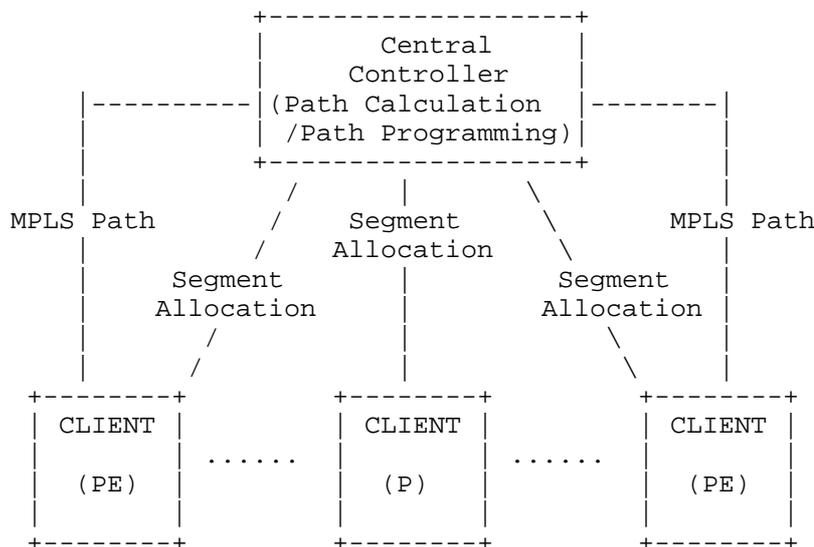


Figure 2 Central Control for MPLS Path Programming

There are two types of MPLS path: Transport-Oriented MPLS Path and Service-Oriented MPLS Path. For the transport-oriented MPLS path, segment routing is the typical solution: MPLS segment distribution is done by IGP extensions ([I-D.ietf-isis-segment-routing-extensions]) and [I-D.ietf-ospf-segment-routing-extensions]); the programmed MPLS path can be downloaded through PCEP extensions from PCE to PCC([I-D.sivabalan-pce-segment-routing]). For the service-oriented MPLS path programming, it not only includes composing the MPLS path in the service and network layer, but also includes determining the

mapping of the service path to the transport path. Since the process corresponding to the label in the service label stack is always located at the PE nodes, BGP extensions can be introduced for service-oriented path programming.

5.2. BGP-based MPLS Segment Distribution

1. Label Allocation

There are two types of label used for MPLS segments:

- 1) Local Label: The service process is done locally. The label can be allocated by the local PE which provides the process.
- 2) Global Label: The service process is common in multiple PEs. This means the label has global meaning. The label allocation can be done by the central controller. The global label work can refer to [I-D.li-mpls-global-label-framework].

2. Label Mapping Distribution

BGP extensions can be used to distribution label mapping. Regarding to the above two types of label allocation, the process is as follows:

- 1) Local Label Mapping: BGP can directly distribute the label mapping from the local PE to peer PEs. The local PE can also only distribute the label mapping to central controller. Then the central controller re-distribute the label mapping to other PEs. In this method, the central controller plays the role of traditional RR.
- 2) Global Label Mapping: The label mapping for the service can be directly distributed by the central controller to multiple PEs. It can be done by BGP extensions.

5.3. MPLS Service Path Programming

5.3.1. Label Combination and 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. The BGP extensions can be reused to download the programmed MPLS path.

5.3.2. Mapping of Service Path to Service Path

Since the transport path is also to satisfy the service bearing the requirement, it can reuse the existing MPLS tunnel technology or it can also be programmed according to traffic engineering requirements of service. Then there needs to be implements the mapping of the service path to the transport path. There are two ways to implement the mapping:

1. BGP Extensions: Through the community attribute of BGP, the identifier of the transport path can be carrier when distribute label stack for a specific prefix.
2. I2RS Extensions: I2RS can be used to download route policy to the client node. Based on the policy, the client node can implement the required mapping.

5.4. Compatibility

When the MPLS path programming is done the central controller and downloaded through BGP extensions to the Client node, the path SHOULD has higher priority than the path calculated on the Client node's own.

5.5. Protocol Extensions Requirements

5.5.1. BGP

REQ 01: BGP extensions SHOULD be introduced to distribute local label mapping for specific process to the central controller and other client nodes.

REQ 02: BGP extensions SHOULD be introduced to distribute global label mapping for specific process from the central controller to the client nodes .

REQ 03: BGP extensions SHOULD be introduced to download label stack for service-oriented MPLS path.

REQ 04: BGP extensions SHOULD be introduced to carry the identifier of the transport MPLS path with service MPLS path to implement the mapping.

REQ 05: BGP extensions SHOULD be introduced to specify the end-points to accept the prefix advertised by the central controller.

REQ 06: BGP extensions SHOULD be introduced to specify the priority for the prefix with attributes of MPLS path programming advertised by the central controller.

REQ 07: When route selection is done in the client node, the path advertised by the central controller SHOULD have higher priority than the path calculated on the client's own.

5.5.2. I2RS

REQ 11: I2RS clients SHOULD provide interface to I2RS agent to download policy to implement the mapping of the service path to the transport path.

6. IANA Considerations

This document makes no request of IANA.

7. Security Considerations

TBD.

8. References

8.1. Normative References

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

8.2. Informative References

[I-D.chen-mpls-source-label]
Chen, M., Xu, X., Li, Z., Fang, L., and G. Mirsky,
"MultiProtocol Label Switching (MPLS) Source Label",
draft-chen-mpls-source-label-06 (work in progress),
October 2014.

[I-D.filsfils-spring-segment-routing]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R.,
Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe,
"Segment Routing Architecture", draft-filsfils-spring-
segment-routing-04 (work in progress), July 2014.

- [I-D.filsfils-spring-segment-routing-central-epe]
Filsfils, C., Previdi, S., Patel, K., Aries, E., shaw@fb.com, s., Ginsburg, D., and D. Afanasiev, "Segment Routing Centralized Egress Peer Engineering", draft-filsfils-spring-segment-routing-central-epe-03 (work in progress), January 2015.
- [I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-03 (work in progress), October 2014.
- [I-D.ietf-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-04 (work in progress), February 2015.
- [I-D.ietf-pce-pce-initiated-lsp]
Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", draft-ietf-pce-pce-initiated-lsp-03 (work in progress), March 2015.
- [I-D.li-l2vpn-segment-evpn]
Li, Z., Yong, L., and J. Zhang, "Segment-Based EVPN(S-EVPN)", draft-li-l2vpn-segment-evpn-01 (work in progress), February 2014.
- [I-D.li-mpls-global-label-framework]
Li, Z., Zhao, Q., Chen, X., Yang, T., and R. Raszuk, "A Framework of MPLS Global Label", draft-li-mpls-global-label-framework-02 (work in progress), July 2014.
- [I-D.li-mpls-global-label-usecases]
Li, Z., Zhao, Q., Yang, T., and R. Raszuk, "Use Cases of MPLS Global Label", draft-li-mpls-global-label-usecases-02 (work in progress), July 2014.
- [I-D.li-mpls-network-virtualization-framework]
Li, Z. and M. Li, "Framework of Network Virtualization Based on MPLS Global Label", draft-li-mpls-network-virtualization-framework-00 (work in progress), October 2013.

- [I-D.sivabalan-pce-segment-routing]
Sivabalan, S., Medved, J., Filsfils, C., Crabbe, E.,
Raszuk, R., Lopez, V., and J. Tantsura, "PCEP Extensions
for Segment Routing", draft-sivabalan-pce-segment-
routing-03 (work in progress), July 2014.
- [I-D.zhang-l3vpn-label-sharing]
Zhang, M., Zhou, P., and R. White, "Label Sharing for Fast
PE Protection", draft-zhang-l3vpn-label-sharing-02 (work
in progress), June 2014.
- [RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute
Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May
2005.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
Networks (VPNs)", RFC 4364, February 2006.
- [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.
- [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and
L. Yong, "The Use of Entropy Labels in MPLS Forwarding",
RFC 6790, November 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