



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 7 August 2023.

## Copyright Notice

Copyright (c) 2023 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

- [1. Introduction](#)
- [2. Motivation](#)
- [3. Constructs and building blocks](#)
  - [3.1. Context Protocol Nexthop Address](#)
  - [3.2. MPLS context FIB](#)
  - [3.3. Context Label](#)
  - [3.4. Roles of nodes in a MPLS plane](#)
    - [3.4.1. Edge-nodes \(PLER\)](#)
    - [3.4.2. Transit-nodes \(PLSR\)](#)
  - [3.5. Sending traffic into the MPLS plane](#)
- [4. Terminology](#)
- [5. BGP families, routes and encoding](#)
  - [5.1. New address-families for "MPLS namespace signaling"](#)
    - [5.1.1. AFI: 16399, SAFI: 128](#)
    - [5.1.2. AFI: 16399, SAFI: 1](#)
  - [5.2. Routes and Operational procedures](#)
    - [5.2.1. "Context-Nexthop" discovery route](#)
    - [5.2.2. MPLS namespace "Private Label" routes](#)
- [6. Example of Usecases](#)
  - [6.1. Improve scaling and convergence of a Seamless MPLS network](#)
  - [6.2. Service Forwarding Helper usecase](#)
  - [6.3. Standard BGP API to a MPLS network's forwarding-plane](#)
  - [6.4. Traffic engineering and Security advantages](#)
- [7. IANA Considerations](#)
- [8. Security Considerations](#)
- [9. Acknowledgements](#)
- [10. Normative References](#)

## 1. Introduction

The MPLS forwarding layer in a core network is a shared resource. The MPLS FIB at nodes in this layer contains labels that are dynamically allocated and locally significant at that node. These labels are scoped in context of the global loopback address. Let us call this the global MPLS namespace.

For some usecases like upstream label allocation, it is useful to create private MPLS namespaces (virtual MPLS FIB) over this shared MPLS forwarding layer. This allows installing deterministic label values in the private FIBs created at nodes participating in the private MPLS namespace, while preserving the "locally significant" nature of the underlying shared global MPLS FIB.

This document defines new address families (AFI: 16399, SAFI: 128, or 1) and associated signaling mechanisms to create and use MPLS forwarding contexts in a network.

The mechanism described in this document reuse [\[RFC4364\]](#) and [\[RFC8277\]](#) procedures to implement Upstream label allocation. The MPLS Namespace family uses BGP VPN style NLRI where the FEC is a MPLS Label, instead of IP prefix. The concepts of MPLS Context tables and upstream allocation are described in [\[RFC5331\]](#).

A BGP speakers participating in a private MPLS namespace creates instance of "MPLS forwarding context" FIB, which is identified using a "Context Protocol Nexthop (CPNH)". A Context label MAY be advertised for the Context Protocol Nexthop (CPNH) using a transport layer protocol or BGP family to other nodes.

## 2. Motivation

A provider's core network consists of a global-domain (default forwarding-tables in P and PE nodes) that is shared by all tenants in the network and may also contain multiple private user-domains (e.g. VRF route tables).

The global MPLS forwarding-layer can be viewed as the collection of all default MPLS forwarding-tables. This global MPLS Fib layer contains labels locally significant to each node. The "local-significance of labels" gives the nodes freedom to participate in MPLS-forwarding with whatever label-ranges they can support in forwarding hardware.

In emerging usecases some applications using the MPLS-network may benefit from a "static labels" view of the MPLS-network. In some

other usecases, a standard mechanism to do Upstream label-allocation is beneficial.

It is desirable to leave the global MPLS FIB layer intact, and build private MPLS FIB-layers on top of it to achieve these requirements. The private MPLS FIBs can then be used by the applications as desired. The private MPLS FIBs need to be created only at the nodes in the network where predictable label-values (external label allocation) is desired. E.g. P-routers that need to act as a "Detour-nodes" or "Service-Forwarding-Helpers" that need to mirror service-labels.

In other words, provisioning of these private MPLS FIBs can be gradual and can co-exist with nodes not supporting the feature described in this document. These private MPLS FIBs can be stitched together using either the Context labels over the existing shared MPLS-network tunnels, or 'private' context-interfaces - to form the "private MPLS FIB layer".

An application can then install the routes with desired label-values in the private forwarding contexts with desired forwarding-semantics.

### **3. Constructs and building blocks**

The building-blocks that construct a private MPLS plane are described in this section.

#### **3.1. Context Protocol Nexthop Address**

A private MPLS plane (just "MPLS plane" here-after) is identified by an IP-address called Context Protocol Nexthop (CPNH). This address is unique in the core-network, like any other loopback address.

A loopback-address uniquely identifies a specific node in the network, and we call it Global Protocol Nexthop (GPNH) in this document. The CPNH address uniquely identifies a MPLS plane, aka "MPLS namespace".

Each node that has forwarding context for a MPLS plane MUST be configured with the same CPNH but a different RD, such that the RD:CPNH will uniquely identify that node in the MPLS plane.

#### **3.2. MPLS context FIB**

An instance of a MPLS forwarding-table at a node in the private MPLS plane. This Private MPLS FIB contains the private label routes.

A node can have context FIB for multiple MPLS planes. The same label-value can have a different forwarding-semantic in each MPLS

plane. Thus the applications using that MPLS plane get a deterministic label-value independent of other applications using other MPLS planes.

The terms "MPLS Namespace", "MPLS FIB-layer" and "MPLS plane" are used interchangeably in this document.

### **3.3. Context Label**

A Context label is a non-reserved dynamically allocated label, that is installed in the global MPLS FIB, and points to a MPLS-context FIB. The Context Label have forwarding semantics as follows in the global MPLS FIB:

Context Label -> Pop and Lookup in MPLS-context FIB

Advertising the "context label in conjunction with the GPNH" tells the network how to reach a "RD:CPNH".

### **3.4. Roles of nodes in a MPLS plane**

The node roles in a MPLS plane can be classified into "edge nodes" (call them PLER) or "transit-nodes" (call them PLSR).

#### **3.4.1. Edge-nodes (PLER)**

Private Label Edge-routers (PLER) have MPLS context FIB that belong to the MPLS plane. They advertise the presence of this context FIB using transport layer address families like [BGP-CT](#) [[BGP-CT](#)] or BGP-LU, and private label routes from this FIB are advertised using new BGP AFI/SAFI described in this document.

#### **3.4.2. Transit-nodes (PLSR)**

These are just Border-nodes that do label-swap forwarding for the context labels they see in the Context-Protocol-Nexthop advertisement routes (BGP-CT or BGP-LU) going thru them. They basically stitch/extend the label switched path to a PLER's CPNH when they re-advertise the CPNH routes with nexthop-self.

PLSRs dont have MPLS context FIBs. PLSRs dont have Context Protocol-Nexthop. Because they dont have Private label routes to originate.

However a node in the network can play both roles, of PLER and PLSR.

### **3.5. Sending traffic into the MPLS plane**

At a PLER, MPLS-traffic arriving with private label hits the correct private MPLS FIB by virtue of either arriving on a "private network-interface" that is attached to the MPLS context FIB, or arriving

with a "Context label" on a network-interface attached to the global MPLS FIB.

To send data traffic into this private MPLS plane, the sender MUST use as handle either a "Context label" advertised by a node or a "Private interface" owned by the MPLS context FIB at the node. The MPLS context FIB is created for an application that needs a private MPLS plane.

The Context label is the only dynamic label-value the application needs to learn from the network (PLER node it is connected to), to be able to use the private MPLS plane. The application can chose predictable value for the labels to be programmed in the private MPLS FIBs.

Once the packet enters the private MPLS plane at an edge-node (PLER), the node will forward the packet to the next node (PLSR or PLER), by pushing the Context label advertised by that next-node, and the transport-label to reach that node's GPNH. This will repeat until the packet reaches the PLER's private MPLS FIB that originated that private MPLS-label.

At each PLER in the MPLS plane, the private label value remains the same, and points towards the same resource attached to the MPLS plane. This allows the applications using the MPLS-network a static-labels view of the resources attached to the private MPLS plane.

At each PLSR in the MPLS plane, the Context label value will change (be swapped in forwarding), but is transparent to the application.

#### **4. Terminology**

P-router : A Provider core router, also called a LSR

LSR : Label Switch Router (pure transport node speaking LDP, RSVP etc)

PLSR: a BGP-CT or BGP-LU transit node in a private MPLS plane, that does label-swap forwarding for Context label.

PLER: an edge node in a private MPLS plane. It has a forwarding context for private labels.

Detour-router : A BGP border node that is used as a loose-hop in a traffic-engineered path

PE-router : Provider Edge router, that hosts a service (Internet, L3VPN etc)

SE-router : Service Edge router. Same as PE.

SFH-router : Service Forwarding Helper. A node helping an SE-router with service-traffic forwarding, using service routes mirrored by the SE.

MPLS FIB : MPLS Forwarding table

Global MPLS FIB : Global MPLS Forwarding table, to which shared-interfaces are connected

Private MPLS FIB : Private MPLS Forwarding table, to which private interfaces are connected

Private MPLS FIB Layer (Private MPLS plane): The group of Private MPLS FIBs in the network, connected together via Context labels

Context label : Locally-significant Non-reserved label pointing to a private MPLS FIB

Context nexthop IP-address (CPNH) : An IP-address that identifies the "Private MPLS FIB Layer". RD:CPNH identifies a Private MPLS FIB at a specific BGP node.

Global nexthop IP-address (GPNH) : Global Protocol Nexthop address. E.g. a loopback address used as transport tunnel end-point.

## **5. BGP families, routes and encoding**

This section describes the new constructs defined by this document.

### **5.1. New address-families for "MPLS namespace signaling"**

This document defines a new AFI: "MPLS Namespaces" (IANA code 16399). And two new address-families, using SAFIs 128 and 1. These address families are used to signal "MPLS namespaces" in BGP. To send or receive routes of these address families, these AFI, SAFI pair of values MUST be negotiated in Multiprotocol Extensions capability described in [RFC4760](#) [[RFC4760](#)]

#### **5.1.1. AFI: 16399, SAFI: 128**

This address-family is used to exchange private label-routes in private MPLS FIBs at routers that are connected using a common network interface. The private label route has NLRI prefix format "RD:PrivateLabel" and contains Route-Target extended-community identifying the private FIB Layer (VPN) the route belongs to. The nexthop of these routes is set to either the GPNH or the CPNH of the BGP-speaker advertising the RFC-8277 label.

Any transport layer protocol is used to advertise the Context label that the receiving router uses to send traffic into the private MPLS

FIB. The Context label installed in the global MPLS FIB points to the private MPLS FIB. The Context label is required when the connecting-interface is a shared common interface that terminates into the global MPLS FIB.

Routes of this address-family can be sent with either IPv4 or IPv6 nexthop. The type of nexthop is inferred from the length of the nexthop.

When the length of Next Hop Address field is 24 (or 48) the nexthop address is of type VPN-IPv6 with 8-octet RD set to zero (potentially followed by the link-local VPN-IPv6 address of the next hop with an 8-octet RD).

When the length of Next Hop Address field is 12 the nexthop address is of type VPN-IPv4 with 8-octet RD.

#### **5.1.2. AFI: 16399, SAFI: 1**

This address-family is used to exchange private label-routes in private MPLS FIBs to routers that are connected using a private network-interface.

Because the interface is private, and terminates directly into the private MPLS FIB, a Context label is not required to access the private MPLS FIB and NLRI prefix format is just "PrivateLabel/24", without the RD.

Routes of this address-family can be sent with either IPv4 or IPv6 nexthop. The type of nexthop is inferred from the length of the nexthop.

When the length of Next Hop Address field is 16 (or 32) the nexthop address is of type IPv6 (potentially followed by the link-local IPv6 address of the next hop).

When the length of Next Hop Address field is 4 the nexthop address is a 4 octet IPv4 address.

### **5.2. Routes and Operational procedures**

#### **5.2.1. "Context-Nexthop" discovery route**

The Context-NH discovery route may be a BGP-LU or [[BGP-CT](#)] family route that carries CPNH in the "Prefix" portion of the NLRI. And the Context label is carried in the "Label" field in the [[RFC8277](#)] format NLRI.



This route is advertised with the following path-attributes:

- \*BGP Nexthop attribute (code 14, MP\_REACH) carrying GPNH address.

- \*Route-Target extended community, identifying the Transport class, if applicable.

The "Context-Nexthop discovery route" is originated by each speaker who acts as a PLER. The "RD:Context-nexthop" uniquely identifies the private MPLS FIB at the speaker. The "Context-nexthop address" uniquely identifies the private MPLS plane in the network. The Context label advertised in this route has a local forwarding semantic of "Pop, Lookup in Private MPLS FIB".

A BGP speaker readvertising a BGP-CT Context-Nexthop for RD:CPNH discovery-route MUST follow the mechanisms described in [\[BGP-CT\]](#). Specifically when re-advertising with "next-hop self" MUST allocate a new Label with a forwarding semantic of "Swap Received-Context-Label, Forward to Received-GPNH". This extends reachability to the CPNH across tunnel domains.

#### **5.2.2. MPLS namespace "Private Label" routes**

The Private Label routes are carried in the new address-family "MPLS VpnUnicast" (AFI:16399, SAFI:128) aka "MPLS namespace signaling", defined in this document.

The NLRI format follows the specifications in [\[RFC8277\]](#), with the "Prefix" portion of the NLRI comprising of the RD and "Private MPLS Label" encoded as shown below.

In a MP\_REACH\_NLRI attribute whose AFI/SAFI is MPLS/128, the "Length" field will be 112 bits or less, comprising of the Label, RD and "Private MPLS Label".

In a MP\_REACH\_NLRI attribute whose AFI/SAFI is MPLS/1, the "Length" field will be 48 bits or less, comprising of the Label, and "Private MPLS Label".

NLRI Prefix (Private Label route, AFI:16399, SAFI:128)

This picture shows NLRI format when the RFC-8277 Multiple Labels Capability is not used:

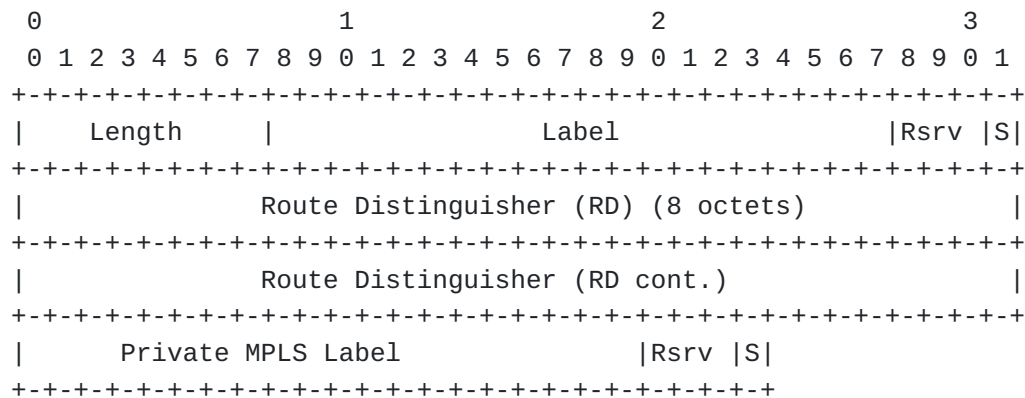


Fig 1: RFC-8277 NLRI with one Label.

- Length:

The Length field consists of a single octet. It specifies the length in bits of the remainder of the NLRI field.

In a MP\_REACH\_NLRI attribute whose AFI/SAFI is MPLS/128, the "Length" field will be 112 bits or less, comprising of the Label, RD and "Private MPLS Label".

As specified in [RFC4760], the actual length of the NLRI field will be the number of bits specified in the Length field, rounded up to the nearest integral number of octets.

- Label:

The Label field is a 20-bit field containing an MPLS label value (see [RFC3032]). This label is locally significant, downstream allocated at the speaker identified in the BGP Nexthop field in MP\_REACH\_NLRI (code 14). This label is pushed in nexthop of the route installed in MPLS context FIB at receiving router.

- Route Distinguisher (RD):

The 8 byte Route Distinguisher as specified in [RFC4760].

- Private MPLS Label:

The "Private MPLS Label" field is a 20-bit field containing an MPLS label value (see [RFC3032]). This is an upstream assigned MPLS label, used as destination of route installed in MPLS context FIB at the receiving router.

- Rsrv:

This 3-bit field SHOULD be set to zero on transmission and MUST be ignored on reception.

- S:

This 1-bit field MUST be set to one on transmission and MUST be ignored on reception.

Attributes on this route:

- \*BGP Nexthop attribute (code 14, MP\_REACH) carrying a GPNH address. (OR)

- \*The Multi-nexthop attribute [[MULTI-NH](#)] with forwarding-semantic:

  - "Forward to RD:CPNH"

- \*Route-Target extended-community, identifying the private FIB-layer

MultiNexthop BGP-attribute (Private Label route)

```
+-----+
| MultiNH.Num-Nexthops = 1                |
+-----+
| FwdSemanticsTLV.FwdAction = Forward      |
+-----+
| NHDescrTLV.NhopDescrType = RD:CPNH or GPNH |
+-----+
```

Fig 2: MultiNexthop attr of Private Label route

A speaker MAY readvertise a private label route without changing the Nexthop (RD:CPNH) carried in it, if the speaker is a pure PLSR.

If it does alter the nexthop to SelfRD:CPNH, it SHOULD act as a PLER, and for e.g. originate a "Context-Nexthop discovery route" for prefix "SelfRD:CPNH".

Even if the speaker sets nexthop-address to Self because of regular BGP readvertisement-rules, Label Prefix MUST NOT be altered, and the received NLRI "RD:Private-Label1" MUST be re-advertised as-is. Such that value of label "Private-Label1" doesn't change while the packet traverses multiple nodes in the private MPLS FIB layer.

The Route target attached to the route is the one identifying the private MPLS FIB layer (VPN). The Private label routes resolve over the Context-nexthop route that belong to the same VPN.

A node receiving a "Private Label route" RD:L1 MUST install the label L1 in the private MPLS Forwarding-context identified by the Route-Target attached to the route.

The label route MUST be installed with forwarding-semantic as specified in the received Multi-nexthop attribute. As an example, a Detour node MAY receive the private label route with a forwarding-semantic of "Forward to RD:CPNH" operation. And an Egress node MAY

receive a private label route with a forwarding-semantic pointing to a resource it houses. Note that such a Private label BGP route MAY be received from external-application also.

#### **5.2.2.1. Resolving received Private Label-routes**

A node receiving a "Context-nexthop discovery route" MUST be capable of using either the CPNH or the RD:CPNH carried in the NLRI, to resolve other routes received with this CPNH address or RD:CPNH in the "Nexthop-attributes".

The receiver of a private label route MUST recursively resolve the received nexthop (RD:CPNH) over the Context-Nexthop discovery-route for prefix "RD:CPNH" to determine the label stack "Context Label, Transport Label" to push, so that the MPLS packet with private label reaches the private MPLS FIB originating the route.

If a node receives multiple "Context-nexthop discovery route" for a CPNH, it SHOULD run path-selection after stripping the RD, to find the closest ingress to the private MPLS plane identified by the CPNH. This best path SHOULD be used to resolve a received private label route.

### **6. Example of Usecases**

#### **6.1. Improve scaling and convergence of a Seamless MPLS network**

MPLS Namespaces can be used to improve scaling and convergence properties of a scaled BGP MPLS network. It acts like a Mezanine transport layer that decouples the service layer from the actual transport layer.

Typically service routes in a MPLS network bind to the following entities that identify point-of-presence of a service:

- \*Protocol Nexthop - PE loopback address (GPNH)

- \*Service Label - PE advertised locally significant label that identifies the service

In this model, whenever a PE is taken out of service the GPNH changes, and Service-Label changes - which causes maintenance a heavy convergence event. Because the service routes with massive-scale need to be readvertised with new service-label or PE-address.

An alternate model could be: to advertise the service routes with a protocol-nexthop of CPNH (without RD), with a forwarding-semantic of:

- \*"Push <Private-Label>, and Forward to CPNH"

This model fully decouples the service-layer from the transport-layer identifiers, by making the Service routes refer to the CPNH and Private Labels. Thus the underlying transport layer can change (nodes representing a Private label can be added or removed) without any changes to the service routes. Which present good scaling properties for the network.

This model also allows anycast traffic forwarding to any resource in the network. Multiple PEs can advertise the same Private label to identify a specific service (e.g. peering with an AS) they are offering.

Once the service route traffic enters the private FIB layer, at the closest entry-point determined by path-selection of CPNH auto-discovery routes; then the Private Labels (with pre determined values) pushed will determine the loose hop path taken by the traffic and also the destination-resource.

## **6.2. Service Forwarding Helper usecase**

In a virtualized environment a Service-PE node (that comprises of a vCP and multiple vFPs) can mirror MPLS labels (GL1) in its global MPLS FIB to a private forwarding context at an upstream node (SFH) with information on which vFPs are optimal exit-points for that label. Such that the SFH can optimally forward traffic to GL1 to the right vFPs, thus avoiding intra fabric traffic hops.

To do this, the service PE advertises a private label route with RD:GL1 to the SFH node. The route is advertised with a Multi-nexthop attribute with one or more legs that have a "Forward to SEPx" semantics. Where SEPx is one of many exit-points at the Service-PE node.

## **6.3. Standard BGP API to a MPLS network's forwarding-plane**

This mechanism facilitates predictable (external-allocator determined) label-values, using a standard BGP-family as the API. It gives the external applications a separate MPLS FIB to play with, totally separate from other applications.

This also avoids vendor specific-API dependencies for external-allocators (controller softwares), and vice-versa.

This mechanism also increases the overall MPLS label-space available in the network, because it creates per application label forwarding contexts (namespaces), instead of reserving/splitting the global MPLS FIB among various applications.

#### 6.4. Traffic engineering and Security advantages

- \*Ability of ingress to steer mpls-traffic thru specific detour loose-hop nodes using predictable-labels' stack.
- \*Provide label-spoofing protection at edge-nodes - by virtue of using separate mpls forwarding contexts
- \*Allow private MPLS label usage to spread across multiple-domains/ AS and work seamlessly with existing technologies like Inter-AS VPN option C.

#### 7. IANA Considerations

This document makes following requests of IANA.

New BGP AFI code ("Address Family Numbers" registry):

\*16399 for "MPLS Namespaces"

Note to RFC Editor: this section may be removed on publication as an RFC.

#### 8. Security Considerations

Using separate mpls forwarding contexts for separate applications and stitching them into separate MPLS planes increases the security attributes of the MPLS network.

#### 9. Acknowledgements

The authors thank Jeffrey (Zhaohui) Zhang, Ron Bonica, Jeff Haas and John Scudder for the valuable discussions.

#### 10. Normative References

- [BGP-CT] Vairavakkalai, K., "BGP Classful Transport Planes", 6 September 2022, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ct#name-limiting-scope-of-visibility>>.
- [MULTI-NH] Vairavakkalai, K., "BGP MultiNexthop attribute", 5 November 2022, <<https://tools.ietf.org/html/draft-kaliraj-idr-multinexthop-attribute-04>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/

RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

**[RFC4364]** Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<https://www.rfc-editor.org/info/rfc4364>>.

**[RFC4760]** Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <<https://www.rfc-editor.org/info/rfc4760>>.

**[RFC5331]** Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, DOI 10.17487/RFC5331, August 2008, <<https://www.rfc-editor.org/info/rfc5331>>.

**[RFC8277]** Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <<https://www.rfc-editor.org/info/rfc8277>>.

#### **Authors' Addresses**

Kaliraj Vairavakkalai  
Juniper Networks, Inc.  
1133 Innovation Way,  
Sunnyvale, CA 94089  
United States of America

Email: [kaliraj@juniper.net](mailto:kaliraj@juniper.net)

Minto Jeyananth  
Juniper Networks, Inc.  
1133 Innovation Way,  
Sunnyvale, CA 94089  
United States of America

Email: [minto@juniper.net](mailto:minto@juniper.net)