Workgroup: Network Working Group
Internet-Draft:
draft-kaliraj-idr-multinexthop-attribute-04
Published: 5 November 2022
Intended Status: Standards Track
Expires: 9 May 2023
Authors: K. Vairavakkalai, Ed. M. Jeyananth
Juniper Networks, Inc. Juniper Networks, Inc.
G. Mishra
Verizon Communications Inc.

BGP MultiNexthop Attribute

Abstract

Today, a BGP speaker can advertise one nexthop for a set of NLRIs in an Update. This nexthop can be encoded in either the BGP-Nexthop attribute (code 3), or inside the MP_REACH_NLRI attribute (code 14).

For cases where multiple nexthops need to be advertised, BGP-Addpath is used. Though Addpath allows basic ability to advertise multiple-nexthops, it does not allow the sender to specify desired relationship between the multiple nexthops being advertised e.g., relative-preference, type of load-balancing. These are local decisions at the receiving speaker based on local configuration and path-selection between the various additional-paths, which may tie-break on some arbitrary step like Router-Id or BGP nexthop address.

Some scenarios with a BGP-free core may benefit from having a mechanism, where egress-node can signal multiple-nexthops along with their relationship, in one BGP route, to ingress nodes. This document defines a new BGP attribute "MultiNexthop (MNH)" that can be used for this purpose.

This attribute can be used for both labeled and unlabled BGP families. The MNH can be used to advertise MPLS label along with nexthop for unlabeled families (e.g. Inet Unicast, Inet6 Unicast). Such that, mechanisms at the transport layer can work uniformly on labeled and unlabled BGP families. Service route scale can be confined closer to the service edge nodes, making the transport layer nodes light and nimble. They dont have any service route state, only have service end-point state.

The MNH plays different role in "downstream allocation" scenario than "upstream allocation" scenario. E.g. for [RFC8277] families that advertise downstream allocated labels, the MNH can play the "Label Descriptor" role, describing the forwarding semantics of the label being advertised. This can be useful in network visualization and controller based traffic engineering (e.g. EPE).

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 https://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 9 May 2023.

Copyright Notice

Copyright (c) 2022 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. Terminology
3. Use-cases examples
   3.1. Signaling optimal forwarding exit-points to ingress-node
   3.2. Choosing a received label based on it's forwarding-semantic at advertising node
   3.3. Signaling desired forwarding behavior when installing MPLS Upstream labels at receiving node
   3.4. Load-balancing over EBGP parallel links
   3.5. Flowspec routes with multiple Redirect-IP nexthops

3.6. Color-Only resolution nexthop
3.7. Use of Local Preference within Cooperating AS domains
3.8. Avoid label advertisement oscillation between multihomed PEs.
4. Protocol Operations
4.1. BGP Capability for MNH attribute
4.2. Scope of use, and propagation
4.3. Interaction of MNH with Nexthop (in attr-code 3, 14)
4.4. Interaction with Addpath
4.5. Path-selection considerations
4.5.1. Determining IGP cost
4.5.2. DOMAIN_LOCAL_PREF
4.6. Denoting upstream/downstream semantics
5. The "MultiNexthop (MNH)" BGP attribute encoding
5.1. Propagation Scope checker
5.2. MNH TLV
5.2.1. Upstream signaled Primary forwarding path.
5.2.2. Upstream signaled Backup forwarding path.
5.2.3. Domain Local Preference (DOMAIN_LOCAL_PREF)
5.2.4. Downstream signaled Label Descriptor.
5.3. Nexthop Forwarding Information TLV
5.4. Forwarding Instruction TLV
5.5. Forwarding Argument TLV
5.5.1. Endpoint Identifier
5.5.2. Path Constraints
5.5.3. Payload encapsulation info signaling
5.5.4. Endpoint attributes advertisement
6. Error handling procedures
7. Scaling considerations
8. IANA Considerations
8.1. BGP Path Attribute Code: MULTI_NEXT_HOP
8.2. BGP Capability Code: MULTI_NEXT_HOP
8.3. Registries for MULTI_NEXT_HOP
9. Security Considerations
10. Acknowledgements
11. References
11.1. Normative References
11.2. References
Authors' Addresses

1. Introduction

Today, a BGP speaker can advertise one nexthop for a set of NLRIs in an Update. This nexthop can be encoded in either the top-level BGP-Nexthop attribute (code 3), or inside the MP_REACH_NLRI attribute (code 14).

For cases where multiple nexthops need to be advertised, BGP-Addpath is used. Though Addpath allows basic ability to advertise multiple-

nexthops, it does not allow the sender to specify desired relationship between the multiple nexthops being advertised e.g., relative-ordering, type of load-balancing, fast-reroute. These are local decision at the receiving node based on local configuration and path-selection between the various additional-paths, which may tie-break on some arbitrary step like Router-Id or BGP nexthop address.

Some scenarios with a BGP-free core may benefit from having a mechanism, where egress-node can signal multiple-nexthops along with their relationship to ingress nodes. This document defines a new BGP attribute "MultiNexthop (MNH)" that can be used for this purpose.

This attribute can be used for both labeled and unlabled BGP families. The MNH can be used to advertise MPLS label along with nexthop for unlabeled families (e.g. Inet Unicast, Inet6 Unicast). Such that, mechanisms at the transport layer can work uniformly on labeled and unlabled BGP families. Service route scale can be confined closer to the service edge nodes, making the transport layer nodes light and nimble. They dont have any service route state, only have service end-point state.

The MNH plays different role in "downstream allocation" scenario than "upstream allocation" scenario. E.g. for [RFC8277] families that advertise downstream allocated labels, the MNH can play the "Label Descriptor" role, describing the forwarding semantics of the label being advertised. This can be useful in network visualization and controller based traffic engineering (e.g. EPE).

A new BGP capability ([RFC3392]) called "MultiNexthop (MNH" is defined with type code: IANA TBD. This capability is used to express the ability to send and receive MNH attribute.

2. Terminology

PNH address: Protocol Nexthop address carried in a BGP Update message.

MULTI_NEXT_HOP (aka MNH): BGP MultiNexthop attribute. The new attribute defined by this document.

MNH TLV: MultiNexthop TLV contained in a MNH attribute.

NFI TLV: Nexthop Forwarding Information TLV, contained in a MNH TLV.

FI TLV: Forwarding Instruction TLV, contained in a NFI TLV.

FA TLV: Forwarding Argument TLV, contained in a FI TLV.

3. Use-cases examples

3.1. Signaling optimal forwarding exit-points to ingress-node

In a BGP free core, one can dynamically signal to the ingress-node, how traffic should be load-balanced towards a set of exit-nodes, in one BGP-route containing this attribute.

Example, for prefix1, perform equal cost load-balancing towards exit-nodes A, B; where-as for prefix2, perform unequal-cost load-balancing (40%, 30%, 30%) towards exit-nodes A, B, C.

Example, for prefix1, use PE1 as primary-nexthop and use PE2 as a backup-nexthop.

3.2. Choosing a received label based on it's forwarding-semantic at advertising node

In Downstream label allocation case, the MNH plays role of "Label descriptor" and describes the forwarding treatment given to the label at the advertising speaker. The receiving speaker can benefit from this information as in the following examples:

- For a Prefix, a label with FRR enabled nexthop-set can be preferred to another label with a nexthop-set that doesn't provide FRR.

- For a Prefix, a label pointing to 10g nexthop can be preferred to another label pointing to a 1g nexthop

- Set of labels advertised can be aggregated, if they have same forwarding semantics (e.g. VPN per-prefix-label case)

3.3. Signaling desired forwarding behavior when installing MPLS Upstream labels at receiving node

In Upstream label allocation case, the receiving speaker's forwarding-state can be controlled by the advertising speaker, thus enabling a standardized API to program desired MPLS forwarding-state at the receiving node. This is described in the [MPLS-NAMESPACES]

3.4. Load-balancing over EBGP parallel links

Consider N parallel links between two EBGP speakers. There are different models possible to do load balancing over these links:

N single-hop EBGP sessions over the N links. Interface addresses are used as next-hops. N copies of the RIB are exchanged to form N-way ECMP paths. The routes advertised on the N sessions can be attached with Link bandwidth comunity to perform weighted ECMP.

1 multi-hop EBGP session between loopback addresses, reachable via static route over the N links. Loopback addresses are used as next-hops. 1 copy of the RIB is exchanged with loopback address as nexthop. And a static route can be configured to the loopback address to perform desired N-way ECMP path. M loopbacks are configured in this model, to achieve M different load balancing schemes: ECMP, weighted ECMP, Fast-reroute enabled paths etc.

1 multi-hop EBGP session between loopback addresses, reachable via static route over the N links. Interface addresses are used as next-hops, without using additional loopbacks. 1 copy of the RIB is exchanged with MNH attribute to form N-way ECMP paths, weighted ECMP, Fast-reroute backup paths etc. BFD may be used to these directly connected BGP nexthops to detect liveness.

3.5. Flowspec routes with multiple Redirect-IP nexthops

There are existing protocol machinery which can benefit from the ability of MNH to clearly specify fallback behavior when multiple nexthops are involved. One example is the scenario described in [FLWSPC-REDIR-IP] where multiple Redirect-to-IP nexthop addresses exist for a Flowspec prefix. In such a scenario, the receiving speakers may redirect the traffic to different nexthops, based on variables like IGP-cost. If instead, the MNH was used to specify the redirect-to-IP nexthop, then the order of preference between the different nexthops can be clearly specified using one flowspec route carrying a MNH containing those different nexthop-addresses specifying the desired preference-order. Such that, irrespective of IGP-cost, the receiving speakers will redirect the flow towards the same traffic collector device.

3.6. Color-Only resolution nexthop

Another existing protocol machinery that manufactures nexthop addresses from overloaded extended color community is specified in [SRTE-COLOR-ONLY]. In a way, the color field is overloaded to carry one anycast BGP next-hop with pre-specified fallback options. This approach gives us only two next-hops to play with. The 'BGP nexthop address' and the 'Color-only nexthop'

Instead, the MNH could be used to achieve the same result with more flexibility. Multiple BGP nexthops can be carried, each resolving over a desired Transport class (Color), and with customizable fallback order. And the solution will work for non-SRTE networks as-well.

3.7. Use of Local Preference within Cooperating AS domains

LOCAL_PREF defined in [RFC4271] is "AS Local" in scope, not allowed to propagate across EBGP boundaries. Only allowed to be sent over IBGP and Confed-EBGP sessions.

In some deployments where multiple AS are part of single administrative control (Inter-AS option C), it is desirable to use a similar construct across EBGP boundaries but still confining propagation within the Inter-AS option C administrative domain. The MNH attempts to solve this problem by introducing "Domain Local Preference (DOMAIN_LOCAL_PREF)".

3.8. Avoid label advertisement oscillation between multihomed PEs.

In a MPLS network, a router may be multihomed to two PEs. The PEs may re-advertise routes received from the router to the IBGP core with self as nexthop and a "per nexthop" label. The PEs may also protect failure of primary path to the router by using the IBGP path via the other multihomed PE as a backup path.

In this scenario, label allocation oscillation may occur when one PE advertises a new label to the other PE. Reception of a new label results in change of nexthop, as the label is used as back nexthop leg, and per-nexthop label allocation is in use. Thus a new label is allocated and advertised. And when this new label is received by the first PE, it allocates a new label in turn. This process repeats.

This oscillation can be stopped only if the primary path label allocated by a PE does not depend on the primary path label advertised by other PE. A PE needs to be able to advertise multiple labels, one for use as primary path and another to be used as bacakup path by the receiver.

MNH attribute allows to advertise a Backup forwarding path label in addition to Primary forwarding path label. Section 5.2.2.

4. Protocol Operations

4.1. BGP Capability for MNH attribute

A new BGP capability [RFC3392] called "BGP MultiNexthop Attribute (MULTI_NEXT_HOP)" is defined with type code: IANA TBD. The MNH attribute MUST NOT be sent to a BGP speaker that has not advertise the MNH capability. A BGP speaker MUST ignore the MNH attribute received from a peer which has not advertised the MNH capability.

4.2. Scope of use, and propagation

The MNH attribute is intended to be used in a BGP free core, between egress and ingress BGP speakers that understand this attribute.

Also, it is required to avoid un-intentionally leaking it to other AS on an EBGP session, via a BGP speaker that does not understand MNH attribute.

To achieve this, the attribute is defined as "optional non-transitive", and uses a new BGP capability. If a MNH-attribute is received by a PE BGP-speaker that does not understand it, the optional non-transitive nature avoids unintentionally propagating it towards EBGP-peers.

This also means that a RR needs to be upgraded to support this attribute before any PEs in the network can make use of it. When a RR receives the MNH-attribute from a client that supports the attribute, it propagates the attribute as-is when reflecting the route with nexthop unchanged.

When a BGP speaker receives the MNH-attribute from another speaker that did not advertise support of the attribute, the attribute is ignored.

The MNH attribute capability provides additonaly protection against receiving this attribute from EBGP peers, when not intended.

Further, the MNH attribute contains a 'Propagation Scope Checker' that enables propagating it across EBGP boundaries to AS that are under the same administrative control, but prohibits advertisement to an AS outside this administrative control

4.3. Interaction of MNH with Nexthop (in attr-code 3, 14)

When adding a MultiNexthop attribute to an advertised BGP route, the speaker MUST put the same next-hop address in the Advertising PNH field as it put in the Nexthop field inside NEXT_HOP attribute or MP_REACH_NLRI attribute.

A speaker that recognizes the MNH attribute and does not change the PNH while re-advertising the route, e.g. a Route Reflector MUST propagate the MultiNexthop attribute in the re-advertisement, satisfying the constraints in 'Propagation Scope Checker'.

A speaker that recognizes this attribute and changes the PNH while re-advertising the route MUST remove the MultiNexthop attribute in the re-advertisement. The speaker MAY however add a new MultiNexthop attribute to the re-advertisement; while doing so the speaker MUST

record in the "Advertising-PNH" field the same next-hop address as used in NEXT_HOP field or MP_REACH_NLRI attribute.

A speaker receiving a MNH attribute SHOULD ignore it if the next-hop address contained in Advertising-PNH field is not the same as the next-hop address contained in NEXT_HOP field or MP_REACH_NLRI field.

In case of [RFC2545], the global (non link-local) IPv6 address should be used for this purpose.

4.4. Interaction with Addpath

[ADDPATH-GUIDELINES] suggests the following:

"Diverse path: A BGP path associated with a different BGP next-hop and BGP router than some other set of paths. The BGP router associated with a path is inferred from the ORIGINATOR_ID attribute or, if there is none, the BGP Identifier of the peer that advertised the path."

When selecting "diverse paths" for ADD_PATH as specified above, the MNH attribute should also be compared if it exists, to determine if two routes have "different BGP next-hop".

4.5. Path-selection considerations

4.5.1. Determining IGP cost

While tie breaking in the path-selection as described in [RFC4271], 9.1.2.2. step (e) viz. the "IGP cost to nexthop", consider the highest cost among the nexthop-legs present in this attribute.

The IGP cost thus calculated is also used when constructing AIGP TLV ([RFC7311])

4.5.2. DOMAIN_LOCAL_PREF

DOMAIN_LOCAL_PREF is defined in section 5.2.3

When LOCAL_PREF is not available on a route, the DOMAIN_LOCAL_PREF if present is used to tie-break in same position in the path selection.

Procedures described in this document ensure that advertisement of DOMAIN_LOCAL_PREF is confined within cooperating AS domains (Inter AS option C) that are under single administrative control.

4.6. Denoting upstream/downstream semantics

MultiNexthop attribute may describe to a receiving speaker what the forwarding semantics of an Upstream-allocated label should be. This can be used with either labeled or unlabled BGP families.

A MultiNexthop attribute may also play "Downstream signaled Label Descriptor" role. A BGP speaker advertising a route carrying downstream allocated MPLS label MAY add this attribute to the BGP route, to "describe" to the receiving speaker what the label's forwarding semantics is at the Egress node.

Today semantics of a downstream-allocated label is known only to the egress node advertising the label. The speaker receiving the label-binding doesn't know what the label's forwarding semantic at the advertiser is. In some environments, it may be useful to convey this information to the receiving speaker. This may help in better debugging and manageability, or enable the receiving speaker, which could also be some centralized controller, make better decisions about which label to use, based on the label's forwarding-semantic.

While doing upstream-label allocation, this attribute can be used to convey the forwarding-semantics at the receiving node should be. Details of the BGP protocol extensions required for signaling upstream-label allocation are out of scope of this document, and are described in [MPLS-NAMESPACES].

In rest of this document, the use of term "Label" will mean downstream allocated label, unless specified otherwise as upstream-allocated label.

When using the MultiNexthop attribute for IP-routes, the Upstream role is used. Since IP prefixes are by nature upstream allocated, global scope.

5. The "MultiNexthop (MNH)" BGP attribute encoding

"MultiNexthop (MNH)" is a new BGP optional non-transitive attribute (code TBD), that can be used to convey one or more nexthops to a BGP-speaker. This attribute describes forwarding instructions using TLVs described in this document.

This section describes the organization and encoding of the MNH attribute.

```
 MNH Attribute: {
    Propagation Scope Checker,
    Num[MNH TLV]
 }

 MNH TLV: {
     { Nexthop Forwarding Information TLV }
 }

 Nexthop Forwarding Information TLV: {
     Num[Forwarding Instruction TLV]
 }

 Forwarding Instruction TLV: {
     {FwdAction, Forwarding Argument TLVs}
 }
```

Fig 1: Overview of MNH Attribute Layout - Eye candy summary.

A MNH attribute consists of a "Propagation Scope checker" and one of more "MNH TLVs". The Propagation Scope checker confines advertisement scope of a MNH attribute. A MNH TLV contains one Nexthop Forwarding Information (NFI) TLV. A NFI TLV contains one or more Forwarding Instructions (FI) TLV. A FI TLV contains a Forwarding-Action and one more Forwarding Argument TLVs. The Forwarding Argument describe the parameters required to complete the Forwarding Action.

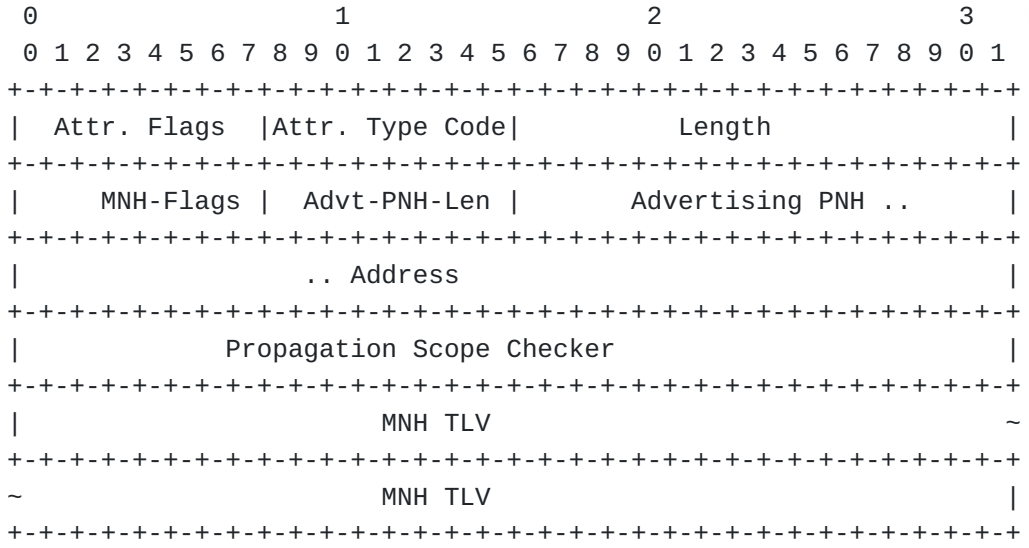


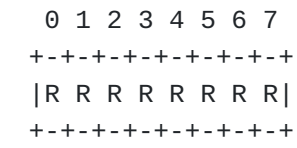
Fig 2: MultiNexthop - BGP Attribute.

- Attr. Flags (1 octet)
  BGP Path-attribute flags. indicating an Optional Non-Transitive attribute. i.e. Optional bit set, Transitive bit reset.

- Attr. Type Code (1 octet)
  Type code allotted by IANA. TBD.

- Length (1 or 2 octets)
  One or Two bytes field stating length of attribute value in bytes

- MNH-Flags (1 octet)



  All bits are reserved.

  R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Advt-PNH-Len (1 octet)
  Length in octets (4 for IPv4, 16 for IPv6, 12 for VPN-IPv4, 24 for VPN-IPv6) of Advertising PNH Address.

- Advertising PNH Address (Advt-PNH-Len octets)
  BGP Protocol Nexthop address advertised in NEXT_HOP or MP_REACH_N Used to sanity-check the MNH attribute. In case of RFC-2545, this the global (non link-local) IPv6 address.

- Propagation Scope Checker: confines advertisement scope of a MNH attr
  described in next section.

- MNH TLVs: One or more MNH TLVs are carried in a MNH attr.
  MNH TLV is described in subsequent sections.

5.1. Propagation Scope checker

The Propagation Scope Check controls the propagation scope of MNH attribute.

By default, MNH attr is not advertised. Setting up the Scope checker appropriately allows advertisement of the attribute within desired boundary.

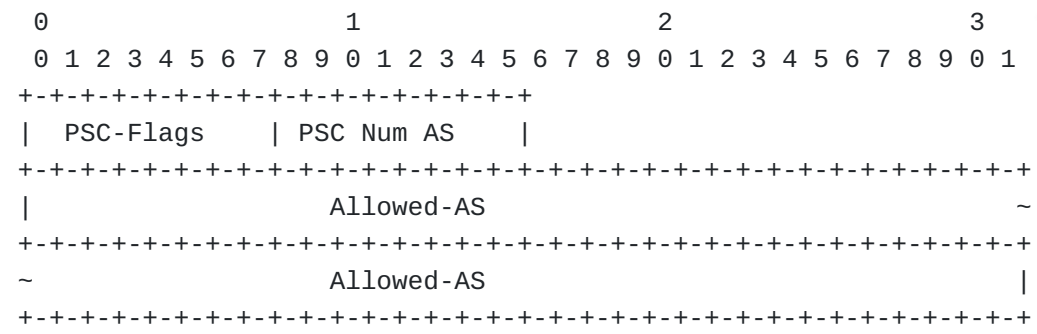
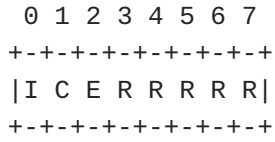


Fig 3: MNH Propagation Scope Checker

By default, MNH attr is not advertised. The PSC flags allow it be a

- PSC Flags (1 octet)



I: When Set allow advertisement to IBGP peers.
C: When Set allow advertisement to Confed-EBGP.
E: When Set allow advertisement to EBGP peers in Allowed-AS l
R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- PSC Num AS: number of AS numbers listed in following field.
  If this value is 0, E bit is considered Clear.
  If E bit is Set, this value should be at least 1.

- Allowed-AS: list of (4 octect) AS numbers that are under same adminis

When the I, C, E bits in PSC Flags are Clear, the MNH attribute MUST NOT be advertised. A speaker originating a MNH-attribute SHOULD set these bits based on desired scope of propagation.

To allow propagation across multiple AS domains, that are under single administrative control, the E bit is Set and "Allowed AS" field contains the list of AS numbers under same administrative control.

5.2. MNH TLV

The type of MNH TLV describes how the forwarding information carried in the MNH TLV is used.

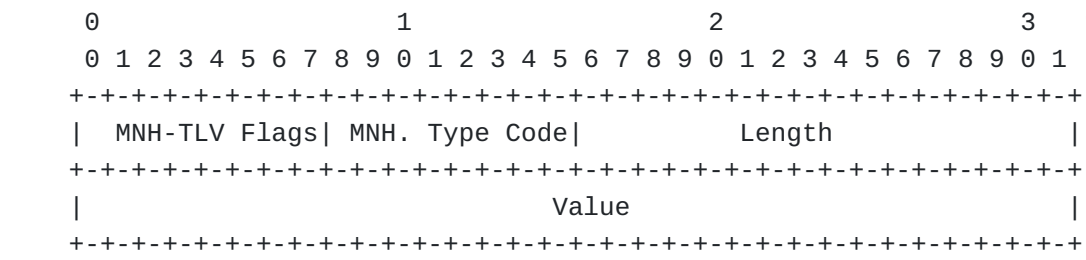
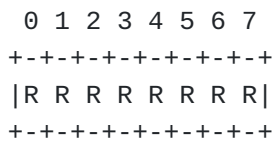


Fig 3: MNH TLV

- MNH-TLV Flags (1 octet)



All bits are reserved.

R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

MNH Type Code	Meaning
0	None
1	Upstream signaled primary forwarding path.
2	Upstream signaled backup forwarding path.
3	Domain Local Preference (DOMAIN_LOCAL_PREF)
4	Downstream signaled Label Descriptor.

- Length
  Length of Value portion in octects.

Type codes 1 and 2 are applicable for upstream allocated prefixes, example IP, MPLS, Flowspec routes.

Type code 4 describes the forwarding behavior given to downstream allocated MPLS label, adveritsed in BGP route.

Usage of Type code 1 in a BGP route containing IP prefix gives similar result as advertising the route with nexthop contained in BGP path-attributes: Nexthop (code 3) or MP_REACH_NLRI (code 14).

Upstream allocation for MPLS routes is achieved by using mechanisms explained in [MPLS-NAMESPACES].

If an invalid Type Code (like 0) is received, the TLV is ignored gracefully handing the error.

If an unknown Type Code is received, it SHOULD be ignored but propagated further when the MNH attribute is propagated, because nexthop is not changed.

If the received Type Code is incompatible for the prefix in BGP NLRI, the TLV should be ignored.

5.2.1. Upstream signaled Primary forwarding path.

Type Code = 1 means the TLV describes forwarding state to be programmed at receiving speaker as primary path nexthop leg. This TLV is used with Upstream allocated or global scope prefixes carried in BGP NLRI. Value part of this TLV contains Nexthop Forwarding Information TLV.

A BGP speaker uses the nexthop forwarding information received in this TLV as a primary path nexthop leg when programming the route for the NLRI prefix in its Forwarding table.

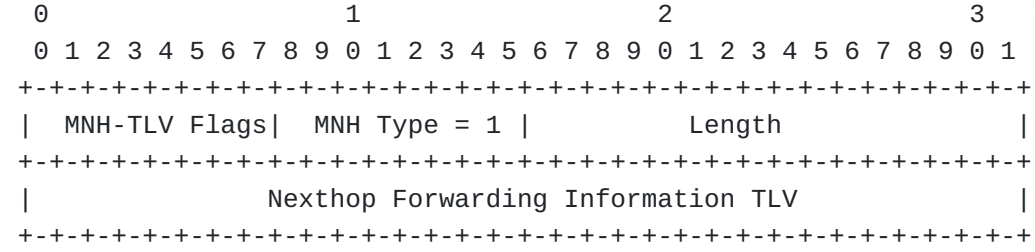


Fig 4: Upstream signaled Primary forwarding path TLV

5.2.2. Upstream signaled Backup forwarding path.

Type Code = 2 means the TLV describes forwarding state to be programmed at receiving speaker as backup-path nexthop leg. This TLV is used with Upstream allocated prefixes or global scoped prefixes. Value part contains Nexthop Forwarding Information TLV.

Signaling a different nexthop for use as backup path is desired in some labeled forwarding scenarios, where two multihomed edge devices use each other as backup path to protect traffic when primary path fails.

This is required to avoid label advertisement oscillation between the multihomed PEs when they implement per-nexthop label allocation mode.

The label advertised by a PE1 for primary path advertisement is allocated/forwarded using external paths as primary leg and backup-path label from other multihomed PE2 as backup-path label. Such that primary-path label allocation at PE1 is not a function of the primary-path label advertised by PE2. Thus the primary path label remains stable at a PE and does not change when a new primary path label is received from the other multihomed PE. This prevents the label oscillation problem.

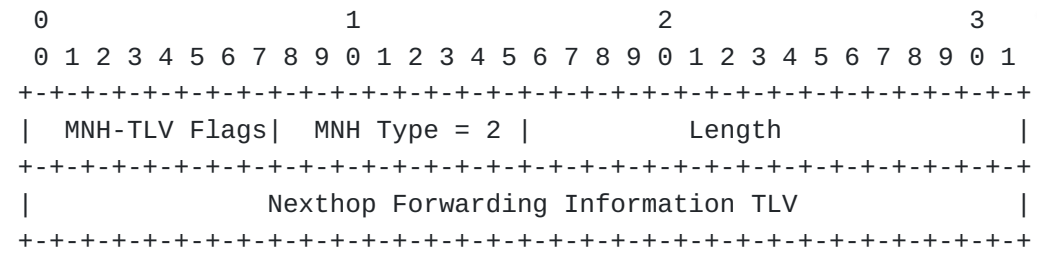


Fig 5: Upstream signaled Backup forwarding path TLV

The backup path label allocated and advertised by a PE is a function of only the primary path. E.g. path to the CE device. So this label value does not change when a new label is received from the other multihomed PE

5.2.3. Domain Local Preference (DOMAIN_LOCAL_PREF)

LOCAL_PREF defined in [RFC4271] is "AS Local" in scope, not allowed to propagate across EBGP boundaries. Only allowed to be sent over IBGP and Confed-EBGP sessions.

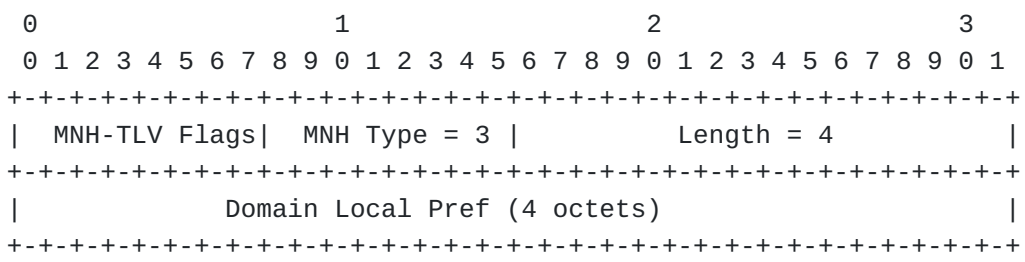
In some deployments where multiple AS are part of single administrative control (Inter-AS option C), it is desirable to use a similar construct across EBGP boundaries but within the administrative domain.

This document defines "Domain Local Preference (DOMAIN_LOCAL_PREF)" which is "Inter-AS option C Domain local" in scope.

When LOCAL_PREF is not available on a route, the DOMAIN_LOCAL_PREF if present can be used to tie-break in same position in the path selection as LOCAL_PREF.

The Propagation Scope Checker MUST ensure that MNH attribute containing DOMAIN_LOCAL_PREF is not advertised across EBGP boundary beyond the Inter-AS option C domain. This is done by Setting E bit, and including AS-numbers of Autonomous systems participating in the Option-C domain.

Information on AS-numbers participating in the Option-C domain is derived from device's local configuration or policy



- Domain Local Preference
  Local preference given to this nexthop-leg/route. Propagated across within Autonomous Systems under same administrative control.

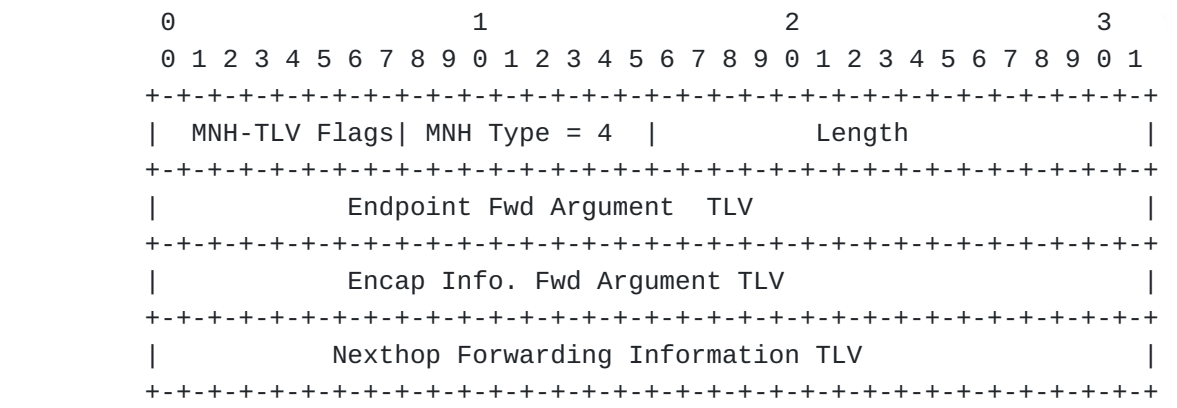
Fig 6: "Domain Local Preference" attribute sub-TLV

This TLV is used as input to path selection.

5.2.4. Downstream signaled Label Descriptor.

Type Code = 4 means the TLV describes forwarding state associated with downstream allocated MPLS label at the egress node identified in Endpoint FA TLV. Value part of this TLV contains Endpoint FA-TLV, Payload Info FA-TLV to identify the label being described, along with Nexthop Forwarding Information TLV that describes the forwarding state.

Signaling what a label advertised in BGP route signifies is helpful for debugging. The information provided by label descriptor can enable new usecases like network visualization and off box EPE decisions.



Endpoint Fwd Argument TLV:
Specifies the IP endpoint. Section 5.5.1.

Encap Info. Fwd Argument TLV:
Specifies the Label value being described. Section 5.5.3.1.

Nexthop Forwarding Information TLV:
Indicates the forwarding state. Described in next section.

Fig 6: Downstream signaled Label Descriptor TLV

TBD: pointer to sec

5.3. Nexthop Forwarding Information TLV

A Nexthop Forwarding Information TLV describes a MNH TLV. It contains one or more Forwarding Instruction TLVs. These Forwarding Instructions are the Forwarding Legs of the MNH.

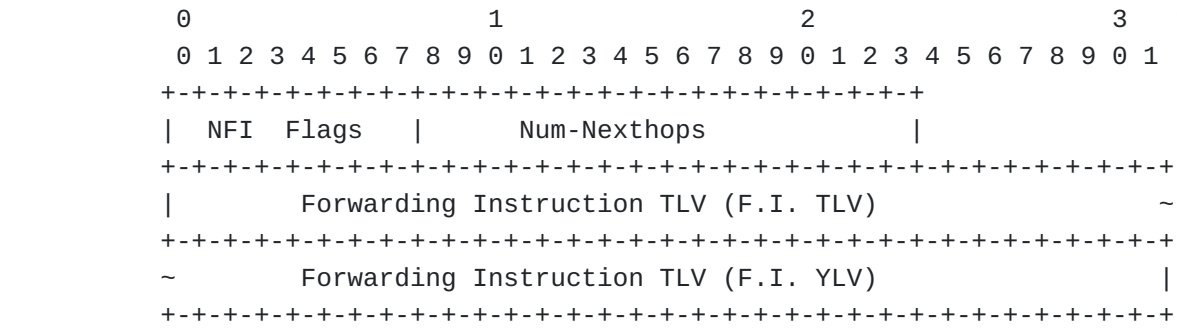


Fig 7: Nexthop Forwarding Information TLV

- NFI Flags (1 octet)

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|R R R R R R R R|
+-+-+-+-+-+-+-+-+
```

All bits are reserved.

R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Num-Nexthops
  Number of F.I. TLVs.

- Forwarding Instruction TLV
  Each F.I. TLV describes a Nexthop Leg.
  Layout of Forwarding Instruction TLV is described in next sectio

5.4. Forwarding Instruction TLV

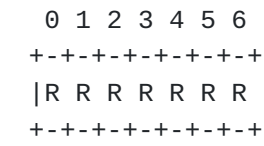
Each Forwarding Instruction TLV describes a Nexthop Leg. It expresses a "Forwarding Action" (FwdAction) along with arguments required to complete the action. The type of actions defined by this TLV are given below. The arguments are denoted by "Forwarding Argument TLVs". The Forwarding Argument TLVs takes appropriate values based on the FwdAction.

Each FwdAction should note the Arguments needed to complete the action. Any extranous arguments should be ignored. If the minimum set of arguments required to complete an action is not received, the Forwarding Instruction TLV should be ignored. Appropriate logging and diagnostic info MAY be provided by an implementation to help troubleshoot such scenarios.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  F.I. Flags   |          Relative Pref        |  FwdAction    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Fwd Argument TLV                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   Fwd Argument TLV                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fig 8: Forwarding Instruction TLV

- F.I. Flags (1 octet)



All bits are reserved.

R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Relative Pref (2 octets)

Unsigned 2 octet integer specifying relative order or preference, a the many forwarding instructions, to use in FIB. All usable nexthop with lowest relative-pref are installed in FIB as primary-path. Thu multiple legs exist with that lowest relative-pref, ECMP is formed.

FwdAction	Meaning
0	None
1	Forward
2	Pop-And-Forward
3	Swap
4	Push
5	Pop-And-Lookup
6	Replicate

Forwarding Instruction TLV with unknown FwdAction should be ignored, and rest of the attribute processed; gracefully handling the error. T may be appropriately logged for diagnosis.

- Length (2 octets)

Length in octets, of all Forwarding Argument TLVs.

Meaning of most of the above FwdAction semantics is well understood. FwdAction 1 is applicable for both IP and MPLS routes. FwdActions 2-5 are applicable for encapsulated payloads (like MPLS) only. FwdActions 1, 6 are applicable for Flowspec routes for Redirect and Mirror actions. FwdAction 6 can also be used to indicate multicast replication like functionality.

The "Forward" action means forward the IP/MPLS packet with the destination prefix (IP-dest-addr/MPLS-label) value unchanged. For IP routes, this is the forwarding-action given for next-hop addresses contained in BGP path-attributes: Nexthop (code 3) or MP_REACH_NLRI

(code 14). For MPLS routes, usage of this action is equivalent to SWAP with same label-value; one such usage is explained in [MPLS-NAMESPACES] when Upstream-label-allocation is in use.

The "Pop-And-Forward" action means Pop the payload header (e.g. MPLS-label) and forward the payload towards the Nexthop IP-address specified in the Endpoint Id TLV, using appropriate encapsulation to reach the Nexthop.

When applied to MPLS packet, the "Pop-And-Lookup" action may result in a MPLS-lookup or an upper-layer header (like IPv4, IPv6) lookup, depending on whether the label that was popped was the bottom of stack label.

If an incompatible FwdAction is received for a prefix-type, or an unsupported FwdAction is received, it is considered a semantic-error and MUST be dealt with as explained in "Error handling procedures" section.

5.5. Forwarding Argument TLV

The Forwarding Argument TLV describes various parameters required to execute a FwdAction.

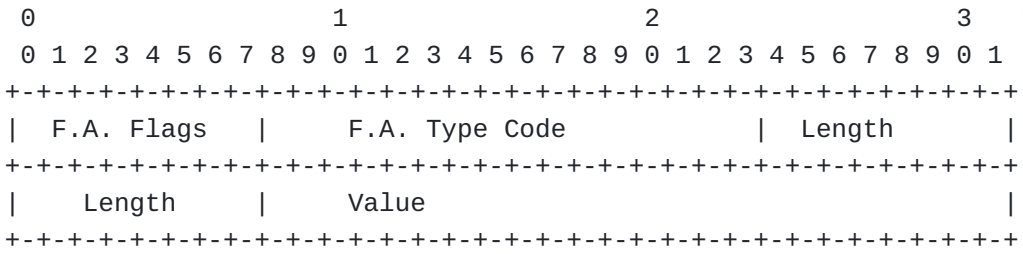


Fig 9: Forwarding Argument TLV

- F.A. Flags (1 octet)

```
          0 1 2 3 4 5 6 7
         +-+-+-+-+-+-+-+-+
         |R R R R R R R R|
         +-+-+-+-+-+-+-+-+
```

      All bits are reserved.

          R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

F.A. Type Code	Meaning
0	None
1	Endpoint Identifier
2	Path Constraints
3	Payload encapsulation info signaling
4	Endpoint attributes advertisement

- Length (2 octets)

   Length in bytes of Value field.

5.5.1. Endpoint Identifier

   F.A. Type Code = 1. This Forwarding Argument TLV identifies an
   Endpoint of different types.

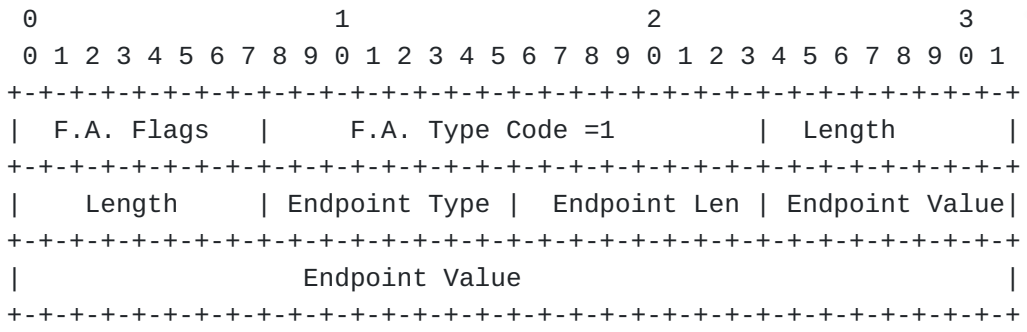


   Fig 10: Endpoint Identifier TLV

- F.A. Flags (1 octet)

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|R R R R R R R R|
+-+-+-+-+-+-+-+-+
```

R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

Endpoint Type	Value	Len (octets)
0	None	
1	IPv4 Address	4
2	IPv6 Address	16
3	MPLS Label (Upstream allocated or Global scope)	4
4	Fwd Context RD	8
5	Fwd Context RT	8

- Endpoint Len (1 octet)

  Length in bytes of Endpoint Value field.

5.5.2. Path Constraints

F.A. Type Code = 2. This Forwarding Argument TLV defines constraints for path to the Endpoint.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  F.A. Flags   |     F.A. Type Code = 2        |  Length       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Length     | ConstrainType | Constrain Len | ConstrainValue|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  ConstrainValue                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fig 11: Path Constraints TLV

- F.A. Flags (1 octet)

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|R R R R R R R R|
+-+-+-+-+-+-+-+-+
```

R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

ConstrainType	Value	Len (octets)
0	None	
1	Proximity check	2
2	Transport Class ID (Color)	4
3	Load balance factor	2

- Constrain Len (1 octet)

Length in bytes of Constrain Value field.

- Proximity check Flags (2 octets)
  Flags describing whether the nexthop endpoint is expected to be away, or multihop away. Format of flags is described in next sec

- Transport Class ID (Color):

This is a 32 bit identifier, associated with the Nexthop address. The Nexthop IP-address specified in "Endpoint Identifier" TLVs are resolved over tunnels of this color.
Defined in [BGP-CT] [draft-kaliraj-idr-bgp-classful-transport-planes

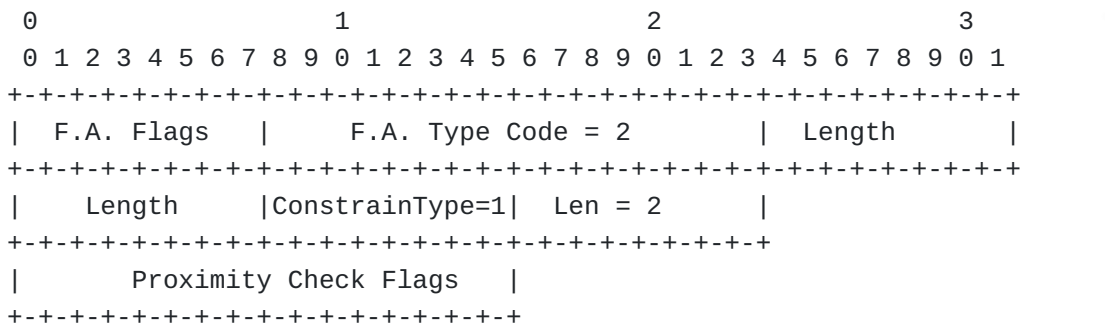
- Load balance factor (2 octets)
  Balance Percentage

5.5.2.1. Proximity check

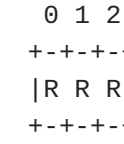
Usually EBGP singlehop received routes are expected to be one hop away, directly connected. And IBGP received routes are expected to be multihop away. Implementations today provide configuring exceptions to this rule.

The 'expected proximity' of the Nexthop can be signaled to the receiver using the Proximity check flags. Such that irrespective of whether the route is received from IBGP/EBGP peer, it can be treated as a single-hop away or multihop away nexthop.

The format of the Proximity check Sub-TLV is as follows:



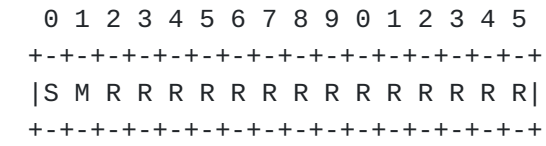
- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

- Proximity check Flags (2 octets)



S: Restrict to Singlehop path.
M: Expect Multihop path.
R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

Fig 12: "Proximity check sub-TLV" sub-TLV

This TLV would be valid with Forwarding Instructions TLV with FwdAction of Forward, Pop-And-Forward, Swap or Push.

When S bit is set, receiver considers the nexthop valid only if it is directly connected to the receiver.

When M bit is set, receiver assumes that the nexthop can be multiple hops away, and resolves the path to the nexthop via another route.

When both S and M bits are set, M bit behavior takes precedence. When both S and M bits are Clear, the current behavior of deriving

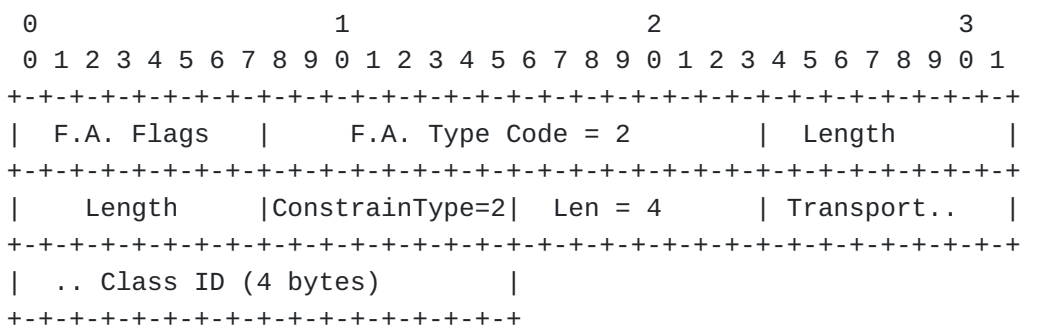
proximity from peer type (EBGP is singlehop, IBGP is multihop) is followed.

5.5.2.2. Transport Class ID (Color)

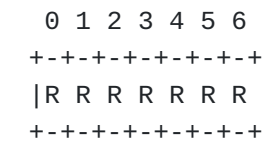
The Nexthop can be associated with a Transport Class, so as to resolve a path that satisfies required Transport tunnel characteristics. Transport Class is defined in [BGP-CT]

Transport Class is a per-nexthop scoped attribute. Without MNH, the Transport class is applied to the nexthop IP-address encoded in the BGP-Nexthop attribute (code 3), or inside the MP_REACH_NLRI attribute (code 14). With MNH, the Transport Class can be specified per Nexthop-Leg (Forwarding Instruction TLV). It is applied to the IP-address encoded in the Endpoint Identifier TLV of type "IPv4 Address", "IPv6 Address" , "MPLS Label (Upstream allocated or Global scope)".

The format of the Transport Class ID Sub-TLV is as follows:



- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

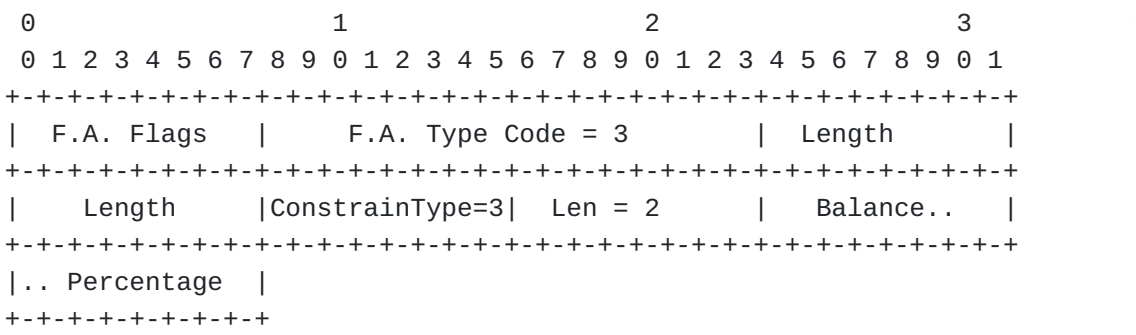
- Length (2 octets)
  Length in bytes of Value field.

- Transport Class ID (Color):
  This is a 32 bit identifier, associated with the Nexthop address.
  The Nexthop specified in Endpoint Identifier TLVs
  are resolved over tunnels of this color.
  Defined in [BGP-CT] [draft-kaliraj-idr-bgp-classful-transport-planes]

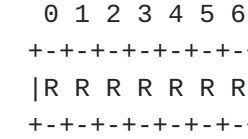
Fig 12: "Transport Class ID (Color)" sub-TLV

This TLV would be valid with Forwarding Instructions TLV with FwdAction of Forward, Swap or Push.

5.5.2.3. Load balance factor



- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

- Len (1 octet)
  Length of the Constrain Value field.

- Balance Percentage:
  This is the explicit "balance percentage" requested by the sender, for unequal load-balancing over these Nexthop-Descriptor-TLV legs. This balance percentage would override the implicit balance-percentage calculated using "Bandwidth" attribute sub-TLV.

Fig 13: "Load-Balance-Factor" sub-TLV

This sub-TLV would be valid with Forwarding Instructions TLV with FwdAction of Forward, Swap or Push.

This is the explicit "balance percentage" requested by the sender, for unequal load-balancing over these Nexthop-Descriptor-TLV legs. This balance percentage would override the implicit balance-percentage calculated using "Bandwidth" attribute sub-TLV

When the sum of "balance percentage" on the nexthop legs does not equal 100, it is scaled up or down to match 100. The individual balance percentages in each nexthop leg are also scaled up or down

proportionally to determine the effective balance percentage per nexthop leg.

5.5.3. Payload encapsulation info signaling

F.A. Type Code = 3. This Forwarding Argument TLV defines payload encapsulation information.

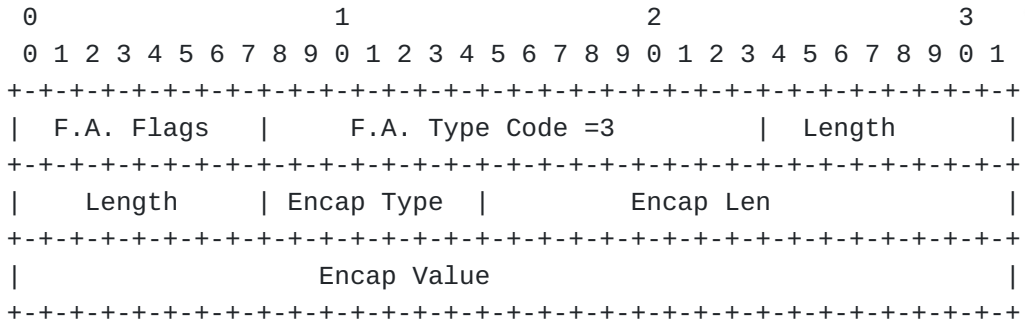
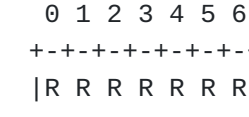


Fig 12: Payload encapsulation info signaling TLV

- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

Endcap Type	Value
0	None
1	MPLS Label Info
2	SR MPLS label Index Info
3	SRv6 SID info

- Encap Len (2 octets)

Length in octets of Encap Value field.

5.5.3.1. MPLS Label Info

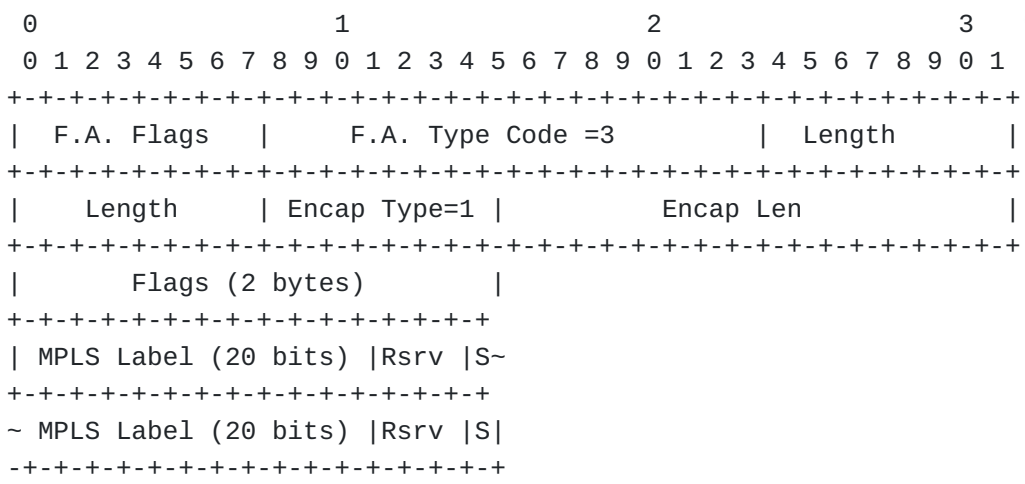
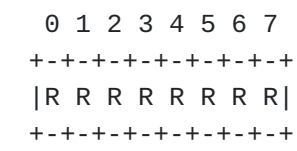


Fig 13: MPLS Label Info.

- F.A. Flags (1 octet)



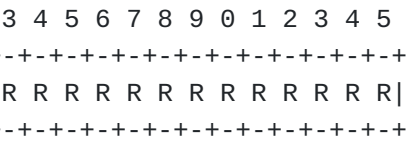
R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

- Encap Type
  = 1, to signify MPLS Label Info.

- Encap Len (2 octets)
  Length in bytes of following Encap Value field.

- Flags (2 octets):



E: ELC bit. Indicates if this egress NH is Entropy Label Capable.
  1 means the Entropy Label capable.
  0 means not capable to handle Entropy Label.

R: Reserved. MUST be set to zero, SHOULD be ignored by receiver.

- MPLS Label, Rsrv, S bit.
  20 bit MPLS Label stack encoded as in RFC 8277.
  S bit set on last label in label stack.

5.5.3.2. SR MPLS Label Index Info

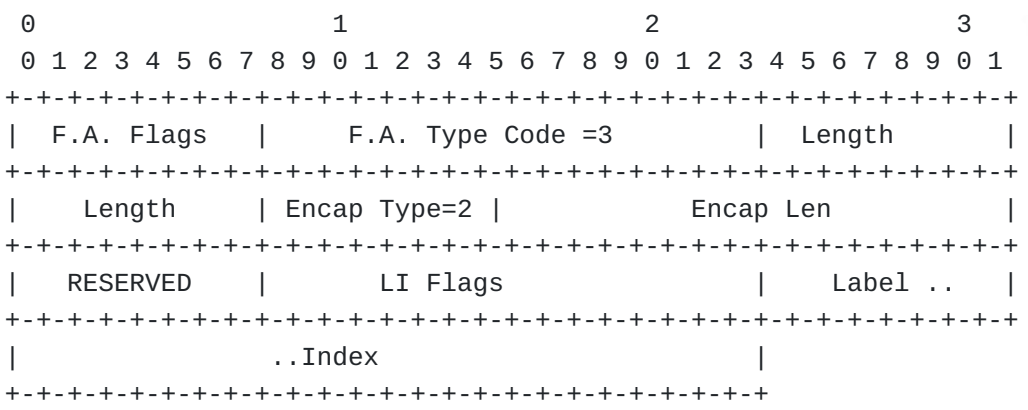
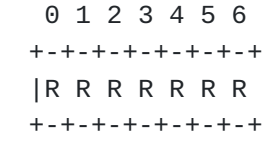


Fig 13: SR MPLS Label Index Info.

- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

- Encap Type
  = 1, to signify SR MPLS SID Info.

- Encap Len (2 octets)
  Length in bytes of following Encap Value field.

Rest of the value portion is encoded as specified in RFC-8669 sec 3.1.

- RESERVED: 8-bit field. MUST be set to zero, SHOULD be ignored by re

- LI Flags: 16 bits of flags. None defined. MUST be set to zero, SHOU

- Label Index:
  32-bit value representing the index value in the SRGB space.

5.5.3.3. SRv6 SID Info

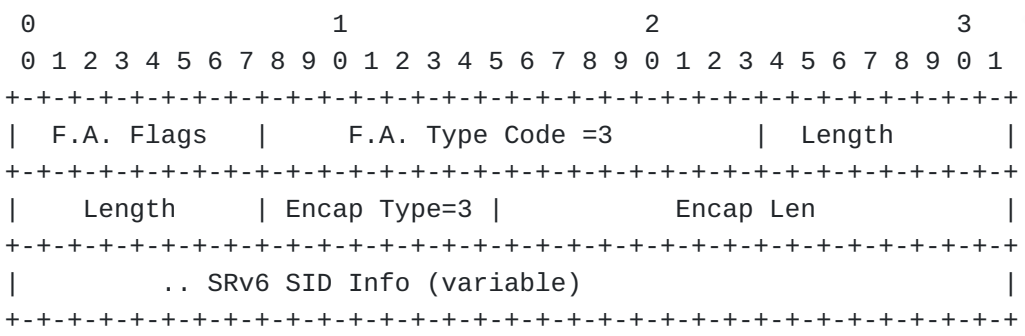
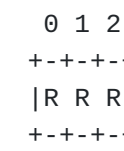


Fig 13: SRv6 SID Info.

- F.A. Flags (1 octet)



R: Reserved. MUST be set to zero, SHOULD be ignored by receiv

- Length (2 octets)
  Length in bytes of Value field.

- Encap Type
  = 1, to signify SR MPLS SID Info.

- Encap Len (2 octets)
  Length in bytes of following Encap Value field.

- SRv6 SID Info:
  One or more IPv6 Addresses (SRv6 SIDs), specified in RFC-8669 sec

5.5.4. Endpoint attributes advertisement

F.A. Type Code = 4. This Forwarding Argument TLV defines attributes of an endpoint.

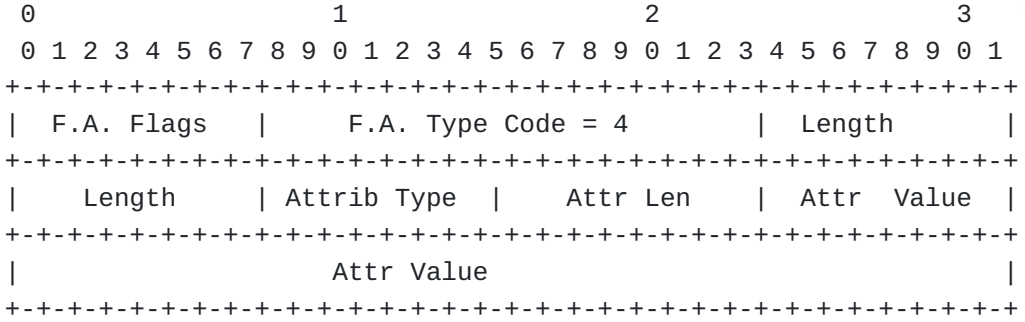
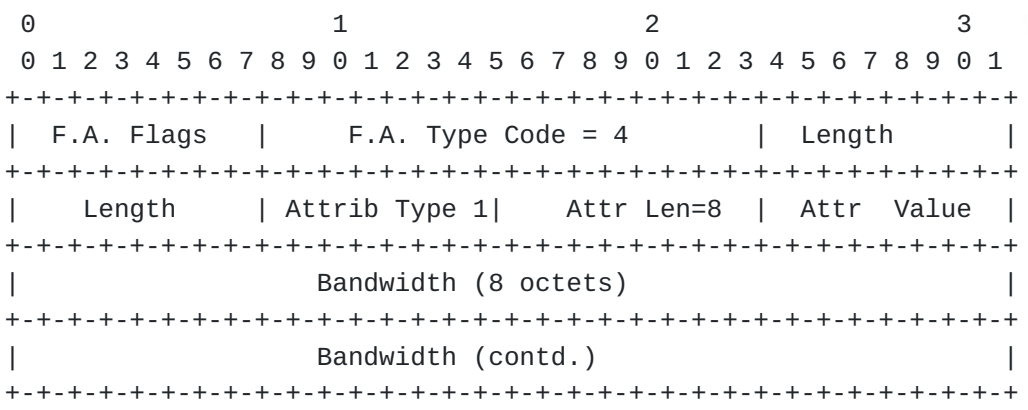


Fig 12: Endpoint attributes advertisement TLV

EP Attrib Type	Attrib Value	Attrib Len (octets)
0	None	
1	Available Bandwidth	8

5.5.4.1. Available Bandwidth



- Len (2 octets)
  Length in bytes of remaining portion of SubTLV.

- Bandwidth
  The bandwidth of the link expressed as 8 octets,
  units being bits per second.

Fig 6: "Available Bandwidth" attribute sub-TLV

This sub-TLV would be valid with Forwarding Instruction TLV with FwdAction of Forward, Swap or Push.

6. Error handling procedures

With MNH TLV Type = 4 (Downstream signaled Label Descriptor), this attribute is used to describe the label advertised by the BGP-peer. If the value in the attribute is syntactically parse-able, but not semantically valid, the receiving speaker should deal with the error gracefully and MUST NOT tear down the BGP session. In such cases the rest of the BGP-update can be consumed if possibe.

With other MNH TLV Types, this attribute is used to specify the forwarding action at the receiving BGP-peer. If the value in the attribute is syntactically parse-able, but not semantically valid, the receiving speaker SHOULD deal with the error gracefully by ignoring the MNH attribute, and continue processing the route. It MUST NOT tear down the BGP session.

If a MNH TLV Type = 4 is received for an IP-route (SAFI Unicast), the MNH attribute SHOULD be ignored. Because IP route prefixes are upstream allocated by nature.

If a MNH TLV Type = 4 is received for an [MPLS-NAMESPACES] route, the MNH attribute SHOULD be ignored. Because the label prefix in MPLS-NAMESPACE family routes is upstream allocated.

The receiving BGP speaker MAY consider the "Num-Nexthops" value in a Nexthop Forwarding Information TLV not acceptable, based on it's forwarding capabilities. In such cases, the MNH attribute SHOULD be considered Unusable, and not be used, ignored on receipt. The condition SHOULD be dealt gracefully and MUST NOT tear down the BGP session.

A TLV or sub-TLV of a certain Type in a MNH attribute can occur only once, unless specified otherwise by that type value. If multiple instances of such TLV or sub-TLV is received, the instances other than the first occurance are ignored.

If a TLV or sub-TLV of an unknown Type value is received, it is ignored and skipped. Remaining part of the MNH attribute if parseable is used

In case of length errors inside a TLV, such that the MNH attribute cannot be used, but the length value in MNH attribute itself is proper, the MNH attribute should be considered invalid and not used. But rest of the route update if parseable should be used. This follows the 'Attribute discard' approach described in [RFC7606] Section 2.

7. Scaling considerations

The MNH attribute allows receiving multiple nexthops on the same BGP session. This flexibility also opens up the possibility that a peer can send large number of multipath (ECMP/UCMP/FRR) nexthops that may overwhelm the local system's forwarding plane. Prefix-limit based checks will not avoid this situation.

To keep the scaling limits under check, a BGP speaker MAY keep account of number of unique multipath nexthops that are received from a BGP peer, and impose a configurable max-limit on that. This is especially useful for EBGP peers.

A good scaling property of conveying multipath nexthops using the MNH attribute with N nexthop legs on one BGP session, as against BGP routes on N BGP sessions is that, it limits the amount of transitionary multipath combinatorial state in the latter model. Because the final multipath state is conveyed by one route update in

deterministic manner, there is no transitionary multipath combinatorial explosion created during establishment of N sessions.

8. IANA Considerations

This document makes request to IANA to allocate the following codes in BGP attributes registry.

8.1. BGP Path Attribute Code: MULTI_NEXT_HOP

A new BGP attribute code TBD for "BGP MultiNexthop Attribute (MULTI_NEXT_HOP)", in "BGP Path Attributes" registry.

8.2. BGP Capability Code: MULTI_NEXT_HOP

This document makes request to IANA to allocate a BGP capability code TBD for "BGP MultiNexthop Attribute (MULTI_NEXT_HOP)".

8.3. Registries for MULTI_NEXT_HOP

This document creates the following sub registries for TLVs and Sub-TLVs within MULTI_NEXT_HOP attribute.

1. Registry of Type codes in "MULTI_NEXT_HOP TLV"

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

MNH Type Code	Meaning
0	None
1	Upstream signaled primary forwarding path.
2	Upstream signaled backup forwarding path.
3	Domain Local Preference (DOMAIN_LOCAL_PREF)
4	Downstream signaled Label Descriptor.

2. Registry of FwdAction values in MNH "Forwarding Instruction TLV"

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

FwdAction	Meaning
0	None
1	Forward
2	Pop-And-Forward
3	Swap
4	Push
5	Pop-And-Lookup
6	Replicate

3. Registry of Type codes in MNH "Forwarding Arguments TLV".

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

F.A. Type Code	Meaning
0	None
1	Endpoint Identifier
2	Path Constraints
3	Payload encapsulation info signaling
4	Endpoint attributes advertisement

4. Registry of Endpoint Types in MNH "Endpoint Identifier TLV" Forwarding Argument.

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

Endpoint Type	Value
0	None
1	IPv4 Address
2	IPv6 Address
3	MPLS Label
4	Fwd Context RD
5	Fwd Context RT

5. Registry of Constrain Types in MNH "Path Constrain TLV" Forwarding Argument.

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

ConstrainType	Value
0	None
1	Proximity check
2	Transport Class ID (Color)
3	Load balance factor

6. Registry of Encap Types in MNH "Payload Encapsulation Info Signaling TLV" Forwarding Argument.

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

Encap Type	Value
0	None
1	MPLS Label Info
2	SR MPLS label Index Info
3	SRv6 SID info

7. Registry of Endpoint Attribute Types in MNH "Endpoint attributes advertisement TLV" Forwarding Argument.

Registration Procedure(s)
Expert Review
Expert(s)
Kaliraj Vairavakkalai
Reference
draft-kaliraj-idr-multinexthop-attribute

EP Attrib Type	Attrib Value
0	None
1	Available Bandwidth

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

9. Security Considerations

The attribute is defined as optional non-transitive BGP attribute, such that it does not accidentally get propagated or leaked via BGP speakers that dont support this feature, especially does not unintentionally leak across EBGP boundaries.

10. Acknowledgements

Thanks to Jeff Haas, Natrajan Venkataraman, Reshma Das, Robert Raszuk, Ron Bonica for the review, discussions and input to the draft.

11. References

11.1. Normative References

[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>.

[RFC2545] Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing", RFC 2545, DOI 10.17487/RFC2545, March 1999, <https://www.rfc-editor.org/info/rfc2545>.

[RFC3392] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", RFC 3392, DOI 10.17487/RFC3392, November 2002, <https://www.rfc-editor.org/info/rfc3392>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC7311] Mohapatra, P., Fernando, R., Rosen, E., and J. Uttaro, "The Accumulated IGP Metric Attribute for BGP", RFC 7311, DOI 10.17487/RFC7311, August 2014, <https://www.rfc-editor.org/info/rfc7311>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[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>.

11.2. References

[ADDPATH-GUIDELINES] Uttaro, Ed., "BGP Flow-Spec Redirect to IP Action", 25 April 2016, <https://datatracker.ietf.org/

doc/html/draft-ietf-idr-add-paths-guidelines-08#section-2>.

[BGP-CT] Vairavakkalai, Ed., "BGP Classful Transport Planes", 25 August 2021, <https://datatracker.ietf.org/doc/draft-kaliraj-idr-bgp-classful-transport-planes/12/>.

[FLWSPC-REDIR-IP] Simpson, Ed., "BGP Flow-Spec Redirect to IP Action", 2 February 2015, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-redirect-ip#section-3>.

[MPLS-NAMESPACES] Vairavakkalai, Ed., "BGP signalled MPLS-namespaces", 28 December 2021, <https://datatracker.ietf.org/doc/html/draft-kaliraj-bess-bgp-sig-private-mpls-labels-04>.

[SRTE-COLOR-ONLY] Filsfils, Ed., "BGP Flow-Spec Redirect to IP Action", 21 February 2018, <https://tools.ietf.org/html/draft-filsfils-spring-segment-routing-policy-06#section-8.8.1>.

Authors' Addresses

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

Email: kaliraj@juniper.net

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

Email: minto@juniper.net

Gyan Mishra
Verizon Communications Inc.
13101 Columbia Pike
Silver Spring, MD 20904
United States of America

Email: gyan.s.mishra@verizon.com