IDR

Internet-Draft

Intended status: Standards Track

Expires: August 8, 2018

S. Previdi, Ed. C. Filsfils A. Lindem Cisco Systems A. Sreekantiah

H. Gredler RtBrick Inc. February 4, 2018

Segment Routing Prefix SID extensions for BGP draft-ietf-idr-bgp-prefix-sid-12

Abstract

Segment Routing (SR) architecture allows a node to steer a packet flow through any topological path and service chain by leveraging source routing. The ingress node prepends an SR header to a packet containing a set of segment identifiers (SID). Each SID represents a topological or a service-based instruction. Per-flow state is maintained only on the ingress node of the SR domain.

This document defines an optional, transitive BGP attribute for announcing BGP Prefix Segment Identifiers (BGP Prefix-SID) information.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of \underline{BCP} 78 and \underline{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 August 8, 2018.

Copyright Notice

Copyright (c) 2018 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. BGP-Prefix-SID 4 2.1. MPLS BGP Prefix SID 4 2.2. IPv6 Prefix Segment 5 3. BGP Prefix-SID Attribute 5 3.1. Label-Index TLV 6 3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 13 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16 Authors' Addresses 16					
2.1. MPLS BGP Prefix SID 4 2.2. IPv6 Prefix Segment 5 3. BGP Prefix-SID Attribute 5 3.1. Label-Index TLV 6 3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	$\underline{1}$. Introduction				<u>3</u>
2.2. IPv6 Prefix Segment 5 3. BGP Prefix-SID Attribute 5 3.1. Label-Index TLV 6 3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	<u>2</u> . BGP-Prefix-SID				4
2.2. IPv6 Prefix Segment 5 3. BGP Prefix-SID Attribute 5 3.1. Label-Index TLV 6 3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	<u>2.1</u> . MPLS BGP Prefix SID				4
3.1. Label-Index TLV 6 3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	<u>2.2</u> . IPv6 Prefix Segment				<u>5</u>
3.2. IPv6 SID 7 3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	3. BGP Prefix-SID Attribute				5
3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	3.1. Label-Index TLV				6
3.3. Originator SRGB TLV 7 4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	3.2. IPv6 SID				7
4. Receiving BGP Prefix-SID Attribute 9 4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
4.1. MPLS Dataplane: Labeled Unicast 9 4.2. IPv6 Dataplane 10 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
4.2. IPv6 Dataplane 16 5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
5. Advertising BGP Prefix-SID Attribute 11 5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	·				
5.1. MPLS Dataplane: Labeled Unicast 11 5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 16 12.2. Informative References 16					
5.2. IPv6 Dataplane 12 6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 16 12.2. Informative References 16					
6. Error Handling of BGP Prefix-SID Attribute 12 7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
7. IANA Considerations 12 8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
8. Manageability Considerations 13 9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16	•				
9. Security Considerations 14 10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
10. Contributors 14 11. Acknowledgements 14 12. References 15 12.1. Normative References 15 12.2. Informative References 16					
11. Acknowledgements					
12. References					
12.1. Normative References					
$\underline{12.2}$. Informative References $\underline{16}$					
Addition 5 Addition 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Authors' Addresses				16

1. Introduction

Segment Routing (SR) architecture leverages the source routing paradigm. A group of inter-connected nodes that use SR forms an SR domain. A segment represents either a topological instruction such as "go to prefix P following shortest path" or a service instruction (e.g., "pass through deep packet inspection"). Other types of segments may be defined in the future.

A segment is identified through a Segment Identifier (SID). Typically, the ingress node of the SR domain prepends an SR header containing segments identifiers (SIDs) to an incoming packet.

As described in [I-D.ietf-spring-segment-routing], when SR is applied to the MPLS dataplane ([I-D.ietf-spring-segment-routing-mpls]), the SID consists of a label while when SR is applied to the IPv6 dataplane the SID consists of an IPv6 address.

A BGP-Prefix Segment (and its BGP Prefix-SID) is a BGP segment attached to a BGP prefix. A BGP Prefix-SID is always a global SID ([I-D.ietf-spring-segment-routing]) within the SR/BGP domain (i.e., the set of Autonomous Systems under a common administration and control and where SR is used) and identifies an instruction to forward the packet over the ECMP-aware best-path computed by BGP to the related prefix. The BGP Prefix-SID is the identifier of the BGP prefix segment. In this document, we always refer to the BGP segment by the BGP Prefix-SID.

This document describes the BGP extension to signal the BGP Prefix-SID. Specifically, this document defines a BGP attribute known as the BGP Prefix-SID attribute and specifies the rules to originate, receive, and handle error conditions for the attribute.

The BGP Prefix-SID attribute defined in this document can be attached to prefixes from AFI/SAFI combinations:

Multiprotocol BGP labeled IPv4/IPv6 Unicast ([RFC8277]).

Multiprotocol BGP ([RFC4760]) unlabeled IPv6 Unicast.

Usage of the BGP Prefix-SID attribute for other AFI/SAFI combinations is not defined herein but may be specified in future specifications.

[I-D.ietf-spring-segment-routing-msdc] describes example use cases where the BGP Prefix-SID is used for the above AFI/SAFI combinations.

It should be noted that:

o A BGP Prefix-SID MAY be global between domains when the interconnected domains agree on the SID allocation scheme. Alternatively, when interconnecting domains, the ASBRs of each domain will have to handle the advertisement of unique SIDs. The mechanisms for such interconnection are outside the scope of the protocol extensions defined in this document.

o A BGP Prefix-SID MAY be attached to a prefix. In addition, each prefix will likely have a different AS_PATH attribute. This implies that each prefix is advertised individually, reducing the ability to pack BGP advertisements (when sharing common attributes).

2. BGP-Prefix-SID

The BGP Prefix-SID advertised for BGP prefix P indicates that the segment routed path should be used (as described below) if the BGP best path selects the corresponding NLRI.

2.1. MPLS BGP Prefix SID

The BGP Prefix-SID is realized on the MPLS dataplane ([I-D.ietf-spring-segment-routing-mpls]) in the following way:

The operator assigns a globally unique label index, L_I, to a locally sourced prefix of a BGP speaker N which is advertised to all other BGP speakers in the SR domain.

According to [I-D.ietf-spring-segment-routing], each BGP speaker is configured with a label block called the Segment Routing Global Block (SRGB). While [I-D.ietf-spring-segment-routing] recommends using the same SRGB across all the nodes within the SR domain, the SRGB of a node is a local property and could be different on different speakers. The drawbacks of the use case where BGP speakers have different SRGBs are documented in [I-D.ietf-spring-segment-routing] and [I-D.ietf-spring-segment-routing-msdc].

If traffic-engineering within the SR domain is required, each node may also be required to advertise topological information and Peering SIDs for each of its links and peers. This information is required to perform the explicit path computation and to express an explicit path as a list of SIDs. The advertisement of topological information and peer segments (Peer SIDs) is done through [I-D.ietf-idr-bgpls-segment-routing-epe].

If the BGP speakers are not all configured with the same SRGB, and if traffic-engineering within the SR domain is required, each node

may be required to advertise its local SRGB in addition to the topological information.

This documents assumes that BGP-LS is the preferred method for collecting both peer segments (Peer SIDs) and SRGB information through [RFC7752], [I-D.ietf-idr-bgpls-segment-routing-epe], and [I-D.ietf-idr-bgp-ls-segment-routing-ext]. However, as an optional alternative for the advertisement of the local SRGB without the topology nor the peer SIDs, hence without applicability for TE, the Originator SRGB TLV of the prefix-SID attribute is specified in Section 3.3 of this document.

As defined in [I-D.ietf-spring-segment-routing], the label index L_I is an offset into the SRGB. Each BGP speaker derives its local MPLS label, L, by adding L_I to the start value of its own SRGB, and programs L in its MPLS dataplane as its incoming/local label for the prefix. It should be noted that while SRGBs and SIDs are advertised using 32-bit values, the derived label is advertised in the 20 right-most bits. See Section 4.1 for more details.

The outgoing label for the prefix is found in the NLRI of the Multiprotocol BGP labeled IPv4/IPv6 Unicast prefix advertisement as defined in [RFC8277]. The label index L_I is only used as a hint to derive the local/incoming label.

<u>Section 3.1</u> of this document specifies the Label-Index TLV of the BGP Prefix-SID attribute; this TLV can be used to advertise the label index for a given prefix.

In order to advertise the label index of a given prefix P and, optionally, the SRGB, an extension to BGP is needed: the BGP Prefix-SID attribute. This extension is described in subsequent sections.

2.2. IPv6 Prefix Segment

When SR is used over an IPv6 dataplane, the BGP Prefix-SID consists of an IPv6 address assigned to the BGP speaker.

3. BGP Prefix-SID Attribute

The BGP Prefix-SID attribute is an optional, transitive BGP path attribute. The attribute type code 40 has been assigned by IANA (see Section 7).

The BGP Prefix-SID attribute is defined here to be a set of elements encoded as "Type/Length/Value" tuples (i.e., a set of TLVs). All BGP

Prefix-SID attribute TLVs will start with a 1-octet type and a 2-octet length. The following TLVs are defined in this document:

- o Label-Index TLV
- o IPv6 SID TLV
- o Originator SRGB TLV

The Label-Index and Originator SRGB TLVs are used only when SR is applied to the MPLS dataplane.

The IPv6 SID TLV is used only when SR is applied to the IPv6 dataplane.

For future extendibility, unknown TLVs MUST be ignored and propagated unmodified.

3.1. Label-Index TLV

The Label-Index TLV MUST be present in the BGP Prefix-SID attribute attached to Labeled IPv4/IPv6 unicast prefixes ([RFC8277]). It MUST be ignored when received for other BGP AFI/SAFI combinations. The Label-Index TLV has the following format:

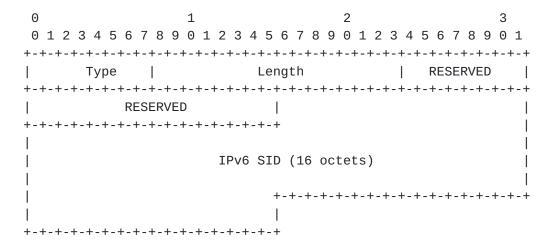
where:

- o Type is 1.
- o Length: is 7, the total length in octets of the value portion of the TLV.
- o RESERVED: 8-bit field. MUST be clear on transmission and MUST be ignored on reception.
- o Flags: 16 bits of flags. None are defined by this document. The flag field MUST be clear on transmission and MUST be ignored on reception.

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

3.2. IPv6 SID

The IPv6 SID TLV MAY be present in the BGP Prefix-SID attribute attached to MP-BGP unlabeled IPv6 unicast prefixes ([RFC4760]). It MUST be ignored for other BGP AFI/SAFI combinations. The IPv6 SID TLV has the following format:

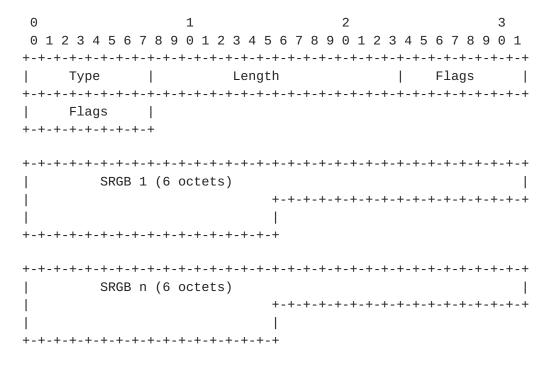


where:

- o Type is 2.
- o Length: is 19, the total length in octets of the value portion of the TLV.
- o RESERVED: 24-bit field for future use. MUST be clear on transmission and MUST be ignored on reception.
- o IPv6 SID: 16 octets.

3.3. Originator SRGB TLV

The Originator SRGB TLV is an optional TLV and has the following format:



where:

- o Type is 3.
- o Length is the total length in octets of the value portion of the TLV: 2 + (multiple of 6).
- o Flags: 16 bits of flags. None are defined in this document. Flags MUST be clear on transmission and MUST be ignored on reception.
- o SRGB: 3 octets of base followed by 3 octets of range. Note that the SRGB field MAY appear multiple times. If the SRGB field appears multiple times, the SRGB consists of multiple ranges that are concatenated.

The Originator SRGB TLV contains the SRGB of the node originating the prefix to which the BGP Prefix-SID is attached. The Originator SRGB TLV MUST NOT be changed during the propagation of the BGP update.

The originator SRGB describes the SRGB of the node where the BGP Prefix SID is attached. It is used to build segment routing policies when different SRGBs are used in the fabric, for example ([I-D.ietf-spring-segment-routing-msdc]).

The receiving routers concatenate the ranges and build the Segment Routing Global Block (SRGB) as follows:

```
SRGB = [100, 199]
[1000, 1099]
[500, 599]
```

The indexes span multiple ranges:

```
index=0 means label 100
...
index 99 means label 199
index 100 means label 1000
index 199 means label 1099
...
index 200 means label 500
```

The originator SRGB may only appear in a BGP Prefix-SID attribute attached to Labeled IPv4/IPv6 unicast prefixes ([RFC8277]). It MUST be ignored when received for other BGP AFI/SAFI combinations. Since the Label-Index TLV is required for IPv4/IPv6 prefix applicability, the originator SRGB will be ignored if it is not specified consistent with $\underline{Section~6}$.

4. Receiving BGP Prefix-SID Attribute

A BGP speaker receiving a BGP Prefix-SID attribute from an EBGP neighbor residing outside the boundaries of the SR domain MUST discard the attribute unless it is configured to accept the attribute from the EBGP neighbor. A BGP speaker MAY log an error for further analysis when discarding an attribute.

4.1. MPLS Dataplane: Labeled Unicast

A Multiprotocol BGP labeled IPv4/IPv6 Unicast ([RFC8277]) session type is required.

A BGP speaker may be locally configured with an SRGB=[SRGB_Start, SRGB_End]. The preferred method for deriving the SRGB is a matter of local node configuration.

Given a label index L_I, we refer to (L = L_I + SRGB_Start) as the derived label. A BGP Prefix-SID attribute is designated "unacceptable" for a speaker M if the derived label value L lies outside the SRGB configured on M. Otherwise the Label-Index TLV is designated "acceptable" to speaker M.

The mechanisms through which a given label index value is assigned to a given prefix are outside the scope of this document.

The BGP Prefix-SID attribute MUST contain the Label-Index TLV and MAY contain the Originator SRGB TLV. A BGP Prefix-SID attribute received without a Label-Index TLV MUST be considered as "unacceptable" by the receiving speaker.

If multiple prefixes are received with the same label index value, all these prefixes MUST have their BGP Prefix-SID attribute considered as "unacceptable" by the receiving speaker.

When a BGP speaker receives a path from a neighbor with an acceptable BGP Prefix-SID attribute, it MUST program the derived label as the local label for the prefix in its MPLS dataplane. In case of an error, a BGP speaker MUST follow to the error handling rules specified in Section 6. A BGP speaker MAY log an error for further analysis.

When a BGP speaker receives a path from a neighbor with an unacceptable BGP Prefix-SID attribute or when a BGP speaker receives a path from a neighbor with a BGP Prefix-SID attribute but is unable to process it (it does not have the capability or local policy disables the capability), it MUST treat the path as if it came without a BGP Prefix-SID attribute. For the purposes of local label allocation, a BGP speaker MUST assign a local (also called dynamic) label (non-SRGB) for such a prefix as per classic Multiprotocol BGP labeled IPv4/IPv6 Unicast ([RFC8277]) operation. A BGP speaker MAY log an error for further analysis.

The outgoing label is always programmed as per classic Multiprotocol BGP labeled IPv4/IPv6 Unicast ([RFC8277]) operation. Specifically, a BGP speaker receiving a prefix with a BGP Prefix-SID attribute and a label NLRI field of Implicit NULL from a neighbor MUST adhere to standard behavior and program its MPLS dataplane to pop the top label when forwarding traffic to the prefix. The label NLRI defines the outbound label that MUST be used by the receiving node.

The label index provides the receiving BGP speaker with guidance as to the incoming label that SHOULD be assigned by that BGP speaker.

4.2. IPv6 Dataplane

When an SR IPv6 BGP speaker receives an IPv6 Unicast BGP Update with a prefix having the BGP Prefix-SID attribute attached, it checks whether the IPv6 SID TLV is present. If present and chosen as the best path, the prefix is installed into the Segment Routing IPv6 dataplane as described in [I-D.ietf-spring-segment-routing].

The Label-Index and Originator SRGB TLVs MUST be ignored on reception. For future extendibility, no TLVs are required for the

BGP IPv6 unicast address family. However, a BGP Prefix-SID attribute corresponding to the BGP IPv6 address family without an IPv6 SID TLV MUST be ignored.

5. Advertising BGP Prefix-SID Attribute

The BGP Prefix-SID attribute MAY be attached to labeled BGP prefixes (IPv4/IPv6) [RFC8277] or to IPv6 unicast prefixes [RFC4760]. In order to prevent distribution of the BGP Prefix-SID attribute beyond its intended scope of applicability, attribute filtering SHOULD be deployed.

A BGP speaker that advertises a path received from one of its neighbors SHOULD advertise the BGP Prefix-SID received with the path without modification, as long as the BGP Prefix-SID was acceptable. If the path did not come with a BGP Prefix-SID attribute, the speaker MAY attach a BGP Prefix-SID to the path if configured to do so. The content of the TLVs present in the BGP Prefix-SID is determined by the configuration.

5.1. MPLS Dataplane: Labeled Unicast

A BGP speaker that originates a prefix attaches the BGP Prefix-SID attribute when it advertises the prefix to its neighbors via Multiprotocol BGP labeled IPv4/IPv6 Unicast ([RFC8277]). The value of the label index in the Label-Index TLV is determined by configuration.

A BGP speaker that originates a BGP Prefix-SID attribute MAY optionally announce the Originator SRGB TLV along with the mandatory Label-Index TLV. The content of the Originator SRGB TLV is determined by configuration.

Since the label index value must be unique within an SR domain, by default an implementation SHOULD NOT advertise the BGP Prefix-SID attribute outside an Autonomous System unless it is explicitly configured to do so.

In all cases, the label field of the advertised NLRI ([RFC8277], [RFC4364]) MUST be set to the local/incoming label programmed in the MPLS dataplane for the given advertised prefix. If the prefix is associated with one of the BGP speaker's interfaces, this is the usual MPLS label (such as the Implicit or Explicit NULL label).

5.2. IPv6 Dataplane

A BGP speaker that originates an IPv6 prefix with the BGP Prefix-SID attribute MAY include the IPv6 SID TLV.

6. Error Handling of BGP Prefix-SID Attribute

When a BGP Speaker receives a BGP Update message containing a malformed or unacceptable BGP Prefix-SID attribute attached to a Labeled IPv4/IPv6 unicast prefix [RFC8277], it MUST ignore the received BGP Prefix-SID attributes and not advertise it to other BGP peers. This is equivalent to the "Attribute discard" action specified in [RFC7606]. When discarding an attribute, a BGP speaker SHOULD log an error for further analysis.

When a BGP Speaker receives a BGP Update message containing a malformed or unacceptable BGP Prefix-SID attribute attached to an unlabeled IPv6 unicast prefix [RFC4760], it MUST treat the advertisement as a withdrawal. This is equivalent to the "Treat-as-withdraw" action specified in [RFC7606]. This action is required since simply ignoring the BGP Prefix-SID attribute would modify the installed path and the "Attribute discard" option is not applicable in this case [RFC7606]. When withdrawing the prefix, a BGP speaker SHOULD log an error for further analysis.

Consistent with [RFC7606], only the first occurrence of the BGP Prefix-SID attribute will be considered and subsequent occurrences will be discarded. Similarly, only the first occurrence of a BGP Prefix-SID attribute TLV of a given TLV type will be considered unless the specification of that TLV type allows for multiple occurrences.

For future extendibility, unknown TLVs MUST be ignored and propagated unmodified.

7. IANA Considerations

This document defines a BGP path attribute known as the BGP Prefix-SID attribute. This document requests IANA to assign an attribute code type (suggested value: 40) to the BGP Prefix-SID attribute from the BGP Path Attributes registry.

Currently, IANA temporarily assigned the following:

40 BGP Prefix-SID (TEMPORARY - registered 2015-09-30, expires 2016-09-30) [draft-ietf-idr-bgp-prefix-sid]

This document defines 3 TLVs for the BGP Prefix-SID attribute. These TLVs need to be registered with IANA. We request IANA to create a registry for BGP Prefix-SID Attribute TLVs as follows:

Under "Border Gateway Protocol (BGP) Parameters" registry, "BGP Prefix-SID TLV Types" Reference: draft-ietf-idr-bgp-prefix-sid Registration Procedure(s): Values 1-254 First Come First Served (FCFS), Value 0 and 255 reserved

Value	Туре	Reference
0	Reserved	this document
1	Label-Index	this document
2	IPv6 SID	this document
3	Originator SRGB	this document
4-254	Unassigned	
255	Reserved	this document

This document also creates a registry for the 16 bits of flags in the Label-Index TLV. Initially, the registry will be empty. Flag bits will be allocated First Come First Served (FCFS) consistent with the BGP-SID TLV Types registry.

Finally, this document creates a registry for the 16 bits of flags in the SRGB Originator TLV. Initially, the registry will be empty. Flag bits will be allocated First Come First Served (FCFS) consistent with the BGP-SID TLV Types registry.

8. Manageability Considerations

This document defines a BGP attribute to address use cases such as the one described in [I-D.ietf-spring-segment-routing-msdc]. It is assumed that advertisement of the BGP Prefix-SID attribute is controlled by the operator in order to:

- o Prevent undesired origination/advertisement of the BGP Prefix-SID attribute. By default, a BGP Prefix-SID attribute SHOULD NOT be attached to a prefix and advertised. Hence, BGP Prefix-SID advertisement SHOULD require explicit enablement.
- o Prevent any undesired propagation of the BGP Prefix-SID attribute. By default, the BGP Prefix-SID is not advertised outside the boundary of a single SR/administrative domain which may include one or more ASes. The propagation to other ASes MUST be explicitly configured.

The deployment model described in [<u>I-D.ietf-spring-segment-routing-msdc</u>] assumes multiple Autonomous Systems (ASes) under a common administrative domain. For this use

case, the BGP Prefix-SID advertisement is applicable to the inter-AS context, i.e., EBGP, while it is confined to a single administrative domain.

9. Security Considerations

This document introduces a BGP attribute (BGP Prefix-SID) which inherits the security considerations expressed in: [RFC4271], [RFC8277], and [I-D.ietf-spring-segment-routing].

It should be noted that, as described in <u>Section 8</u>, this document refers to a deployment model where all nodes are under the single administrative domain. In this context, we assume that the operator doesn't want to leak any information related to internal prefixes and topology outside of the administrative domain. The internal information includes the BGP Prefix-SID. In order to prevent such leaking, the standard BGP mechanisms (filters) are applied at the boundary of the SR/administrative domain.

10. Contributors

Keyur Patel Arrcus, Inc. US

Email: Keyur@arrcus.com

Saikat Ray Unaffiliated US

Email: raysaikat@gmail.com

11. Acknowledgements

The authors would like to thank Satya Mohanty for his contribution to this document.

The authors would like to thank Alvaro Retana for substantive comments as part of the Routing AD review.

The authors would like to thank Shyam Sethuram for comments and discussion of TLV processing and validation.

The authors would like to thank Peter Yee and Tony Przygienda for IETF last call directorate reviews.

12. References

12.1. Normative References

- [I-D.ietf-spring-segment-routing]
 Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B.,
 Litkowski, S., and R. Shakir, "Segment Routing
 Architecture", draft-ietf-spring-segment-routing-15 (work
 in progress), January 2018.
- [I-D.ietf-spring-segment-routing-mpls]
 Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
 Litkowski, S., and R. Shakir, "Segment Routing with MPLS
 data plane", draft-ietf-spring-segment-routing-mpls-11
 (work in progress), October 2017.
- [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.
- [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.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", <u>RFC 4364</u>, 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.
- [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>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, https://www.rfc-editor.org/info/rfc8174.
- [RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", <u>RFC 8277</u>, DOI 10.17487/RFC8277, October 2017, https://www.rfc-editor.org/info/rfc8277.

12.2. Informative References

[I-D.ietf-idr-bgp-ls-segment-routing-ext]

Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H.,

and M. Chen, "BGP Link-State extensions for Segment

Routing", draft-ietf-idr-bgp-ls-segment-routing-ext-04

(work in progress), January 2018.

[I-D.ietf-idr-bgpls-segment-routing-epe]
Previdi, S., Filsfils, C., Patel, K., Ray, S., and J.
Dong, "BGP-LS extensions for Segment Routing BGP Egress
Peer Engineering", draft-ietf-idr-bgpls-segment-routingepe-14 (work in progress), December 2017.

[I-D.ietf-spring-segment-routing-msdc]
Filsfils, C., Previdi, S., Mitchell, J., Aries, E., and P.
Lapukhov, "BGP-Prefix Segment in large-scale data
centers", <u>draft-ietf-spring-segment-routing-msdc-08</u> (work
in progress), December 2017.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and
S. Ray, "North-Bound Distribution of Link-State and
Traffic Engineering (TE) Information Using BGP", RFC 7752,
DOI 10.17487/RFC7752, March 2016, https://www.rfc-editor.org/info/rfc7752.

Authors' Addresses

Stefano Previdi (editor) Cisco Systems IT

Email: stefano@previdi.net

Clarence Filsfils Cisco Systems Brussels Belgium

Email: cfilsfils@cisco.com

Acee Lindem Cisco Systems 301 Midenhall Way Cary, NC 27513 USA

Email: acee@cisco.com

Arjun Sreekantiah

Email: arjunhrs@gmail.com

Hannes Gredler RtBrick Inc.

Email: hannes@rtbrick.com