



# Segment Routing Prefix SID extensions for BGP

*draft-keyupate-idr-bgp-prefix-sid-00*

Keyur Patel, Saikat Ray, Stefano Previdi, Clarence Filsfils

*IETF91, November 2013, Honolulu, US*

# Segment Routing Introduction

- Segment Routing (SR) allows a node to steer a packet flow through any topological path and/or service chain
  - Leverages Source Routing
  - Ingress node prepends SR header to packet containing set of segments
  - Each Segment represents a topological instruction (for Prefix P follow Shortest Path) or a service instruction (pass thru DPI)
- Segments are identified by Segment IDs (SIDs)
  - Leverage existing MPLS dataplane by using MPLS labels as SIDs

# Motivation

- Pure BGP based DataCenters (DCs) does not have any IGPs running
- Basic SR Transport within such BGP based DCs
  - Need a way to announce Prefix SIDs within BGP
  - A Prefix SID represents a SID of a given BGP Prefix
- Currently there is no mechanism to announce Prefix SIDs within BGP
  - BGP can only announce Labels as part of Label address family
  - Need a mechanism within BGP to facilitate announcement of Prefix SIDs (aka Global Segments for Prefixes)

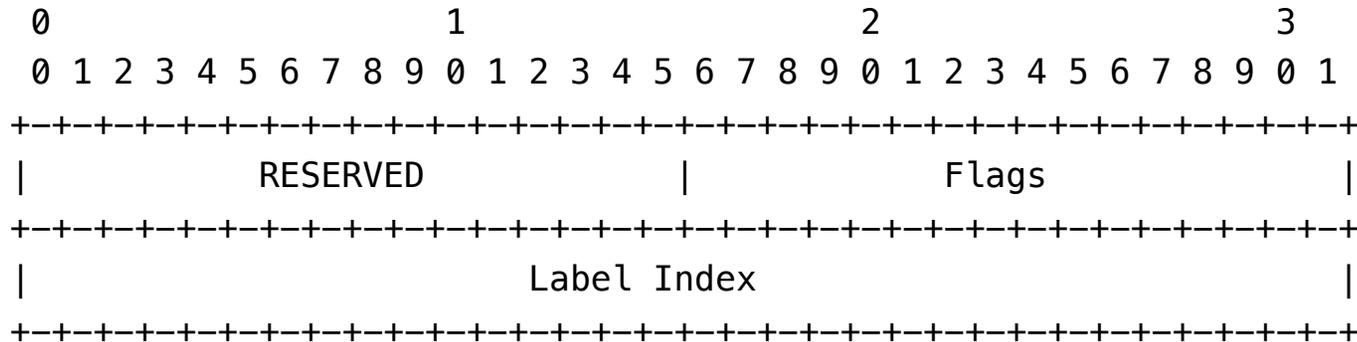
# BGP-Prefix-SID

- Segments associated with a BGP Prefix is known as a BGP-Prefix-SID
  - BGP-Prefix-SIDs are always global within a SR/BGP domain
  - Identifies an instruction to forward the packet over ECMP aware Bestpath computed by BGP for a given prefix
- SR requires BGP speaker to be configured with a Segment Routing Global block (SRGB)
  - Typically configured as a range of labels, SRGB = [SR\_S, SR\_E] SR\_S = Start of the range, SR\_E = End of the range
  - Each Prefix is assigned with its own unique label Index
  - Label Index is an offset value from SR\_S
  - SR label is typically configured as SR\_S + Label Index
- BGP Speaker assigns Label Index to prefixes it originates
  - Alternatively, BGP speaker MAY assign a label Index to a prefix received without an label Index attribute

# BGP-Prefix-SID

- BGP-Prefix-SID draft defines
  - New optionally transitive BGP attribute to announce BGP-Prefix-SID Label Index attribute
  - Rules for generating and processing the Label Index attribute
  - Error handling of the Label Index attribute
- BGP-Prefix-SID requires BGP to run Label address family
  - Labels are carried within NLRI
  - Label Index attribute carries the Label offset. This allows backward compatibility

# BGP-Prefix-SID Label Index Attribute



where:

- o RESERVED: 16 bit field. SHOULD be unset on transmission and MUST be ignored on reception.
- o Flags: 16 bits of flags. None are defined in this document. Flags SHOULD be unset on transmission and MUST be ignored at reception.
- o Label Index: 32 bit value representing the index value in the SRGB space.

# Rules for Assigning & Announcing Label Index Attribute

- BGP Speaker assigns Label Index value to a prefix it originates
  - Alternatively, BGP speaker MAY assign a label Index to a prefix received without an label Index attribute as part of the policy
- Label computed off a valid Label Index MUST be programmed in MPLS data plane for a given prefix
- BGP Speaker MUST re-announce the Label Index received in a Label Index attribute without modification regardless of whether it was used
- Label Index is announced in the BGP-Prefix-SID Label Index attribute
  - Computed Label is announced as part of BGP NLRI

# Rules for processing the received Label Index Attribute

- Received Label Index value is called “unacceptable” if the label computed off the Label Index falls outside the configured range of SRGB for a given BGP speaker
  - BGP speaker should locally assign a label from its private label range and record an error
- For acceptable Label Index value, a BGP speaker computes a Label value and programs the MPLS dataplane for a given prefix
- BGP speaker should ignore and discard BGP-Prefix-SID Label Index attribute from EBGP neighbors unless configured otherwise

# Error handling for the Label Index Attribute

- Discard the BGP-Prefix-SID Label Index attribute if
  - The received attribute is Malformed
  - The received attribute is from an EBGP peer (with a policy to ignore the attribute)

# Implementation Details

- Cisco has a working implementation on IOS-XR
- Happy to do interoperability testing if any vendor has an implementation



Questions?

Request WG to adopt the draft as a WG document.