The bridge to possible

SRv6 uSID Introduction

ver 3k

Jakub Horn
Principal Technical Marketing Engineer

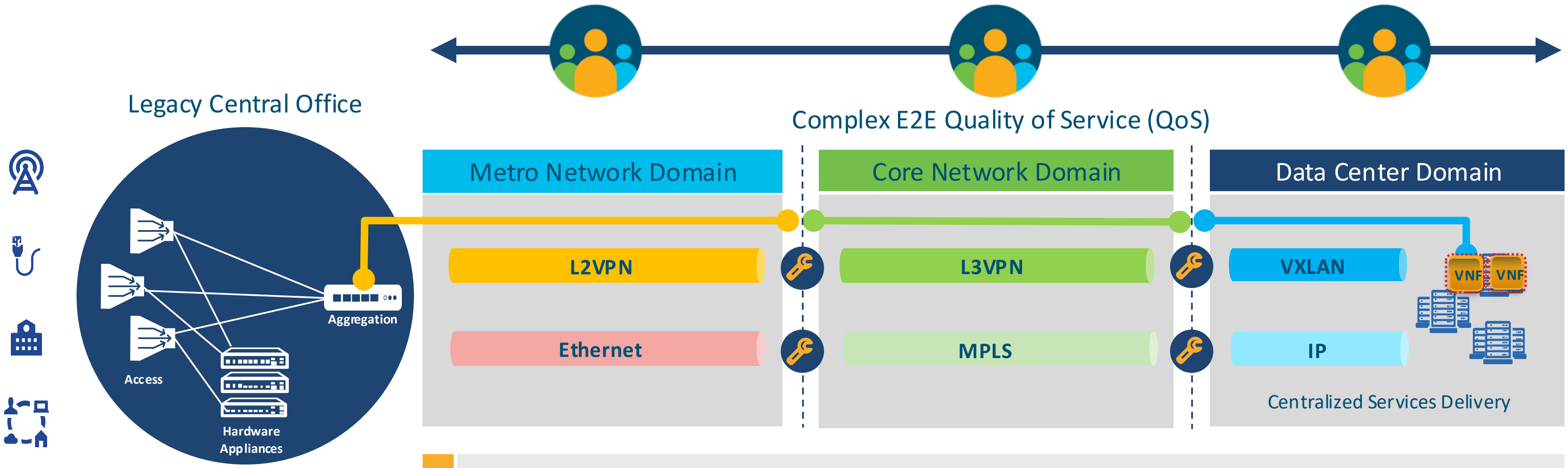
Dec 2024

Content:

- **uSID Technology**
  - **SRv6 DataPlane**
  - **SRv6 Network Programming**
  - **SRv6 ISIS**
  - **SRv6 BGP**
  - **SRv6 Flexible Algorithm**
- **SRv6 Design**
  - **Addressing Plan**
  - **SRv6 Migration**

Understanding Today’s Service Creation

Limited Cross-domain Automation, Cumbersome Service Assurance



Legacy Central Office

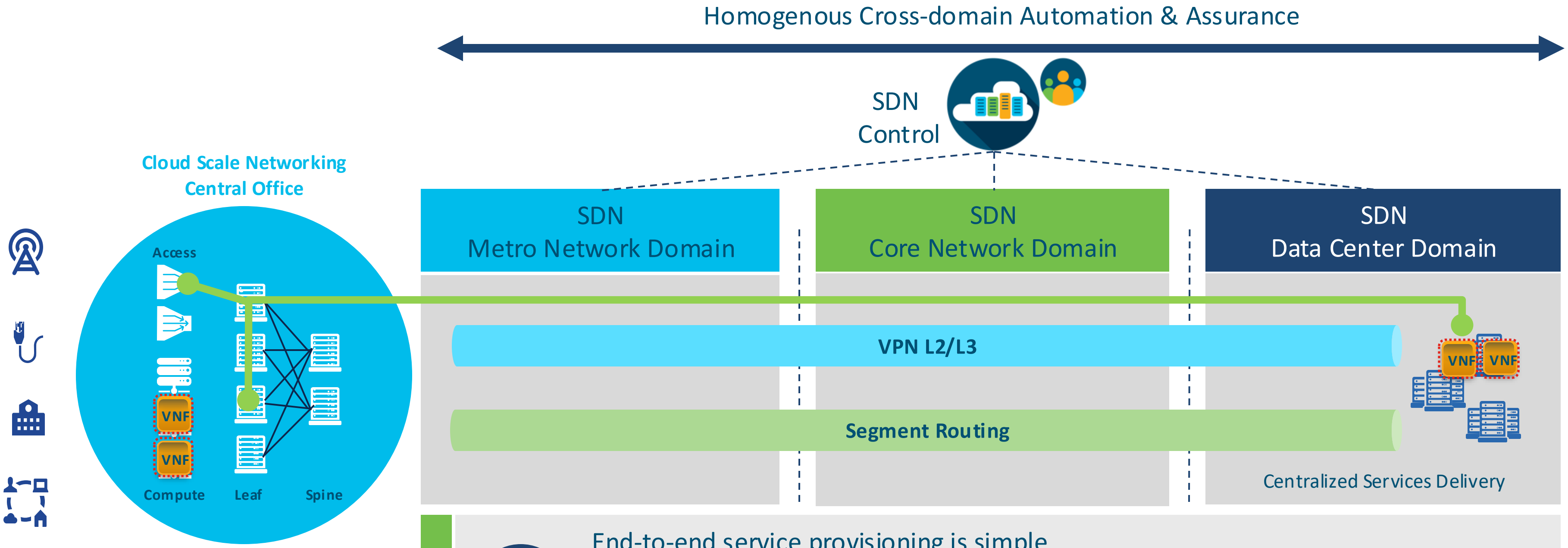
Complex E2E Quality of Service (QoS)

Centralized Services Delivery

End-to-end service provisioning is lengthy and complex

- Multiple network domains under different management teams
- Manual operations
- Heterogeneous underlay and overlay networks

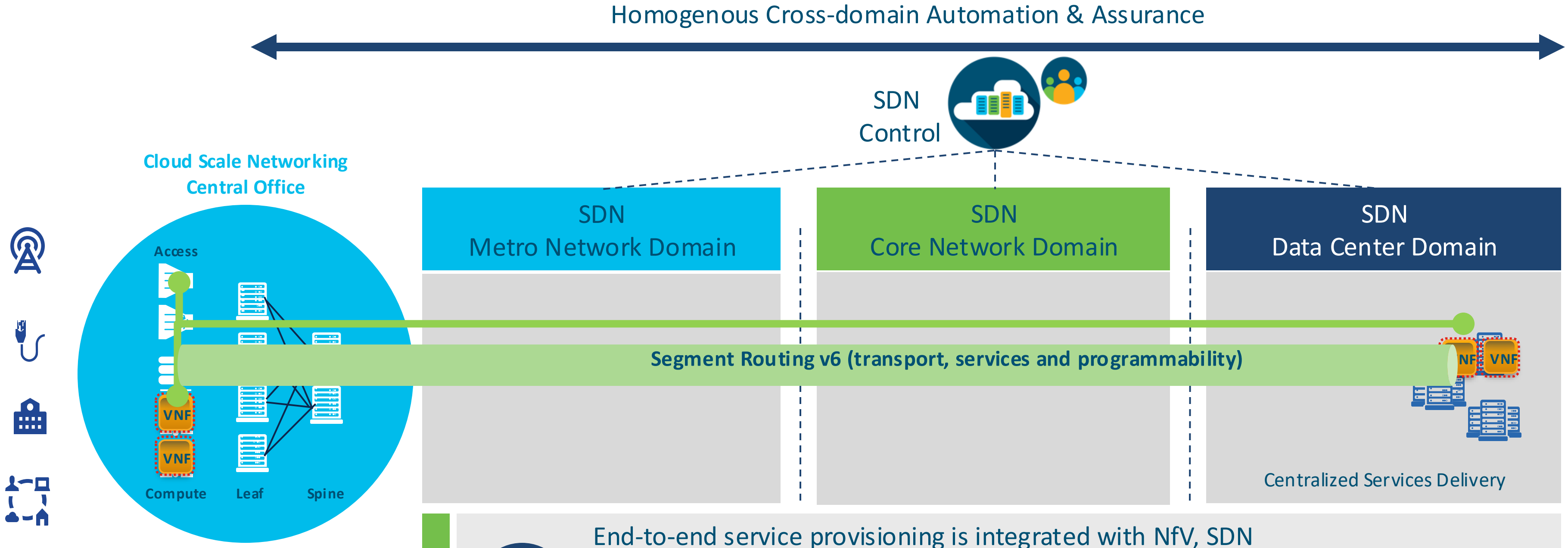
SR-MPLS: SDN ready
“Network as a Fabric” for Service Creation



End-to-end service provisioning is simple

- Multiple network domains under same management teams
- Automated operations
- Homogenous underlay and overlay networks

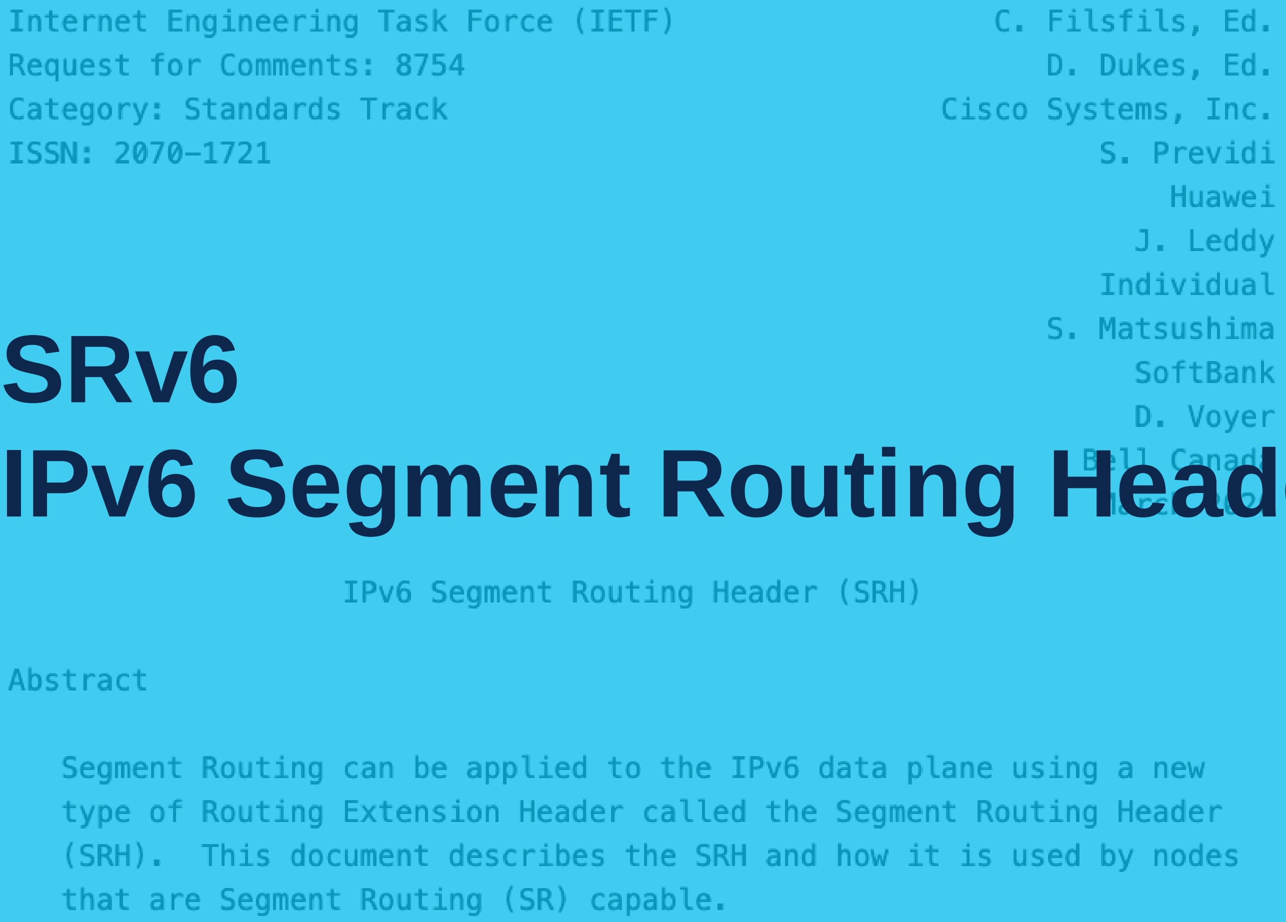
SRv6: SDN, NfV, 5G ready
"Network as an API" for Service Creation


Homogenous Cross-domain Automation & Assurance
SDN Control
Cloud Scale Networking Central Office
Access
VNF
VNF
Compute
Leaf
Spine
SDN Metro Network Domain
SDN Core Network Domain
SDN Data Center Domain
Segment Routing v6 (transport, services and programmability)
VNF
VNF
Centralized Services Delivery

End-to-end service provisioning is integrated with NfV, SDN

- Multiple network domains under same management teams
- Automated operations
- Integrated underlay and overlay networks (NfV)
- Network as API (NfV)
- Hyper Scale (5G)

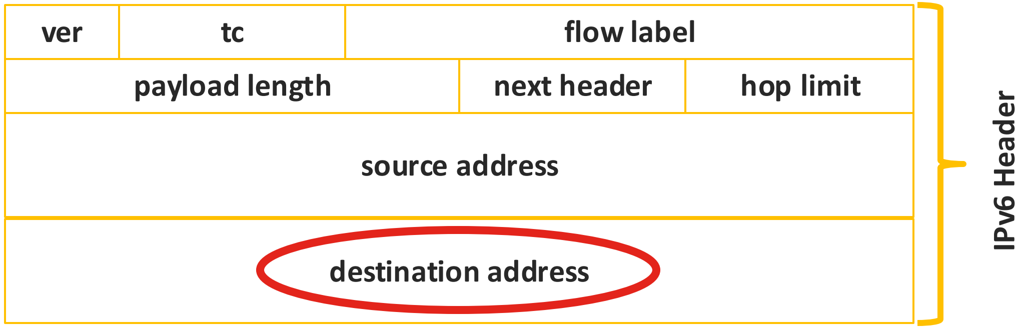
SRv6
IPv6 Segment Routing Header (SRH)



```
Internet Engineering Task Force (IETF)                      C. Filsfils, Ed.
Request for Comments: 8754                                     D. Dukes, Ed.
Category: Standards Track                                Cisco Systems, Inc.
ISSN: 2070-1721                                                    S. Previdi
                                                                       Huawei
                                                                     J. Leddy
                                                                   Individual
                                                                 S. Matsushima
                                                                     SoftBank
                                                                     D. Voyer
                                                                  Bell Canada
                                                                   March 2020


                    IPv6 Segment Routing Header (SRH)

Abstract

   Segment Routing can be applied to the IPv6 data plane using a new
   type of Routing Extension Header called the Segment Routing Header
   (SRH).  This document describes the SRH and how it is used by nodes
   that are Segment Routing (SR) capable.
```

SRv6

- IPv6 Header
- Destination IP address



ver

tc

flow label

payload length

next header

hop limit

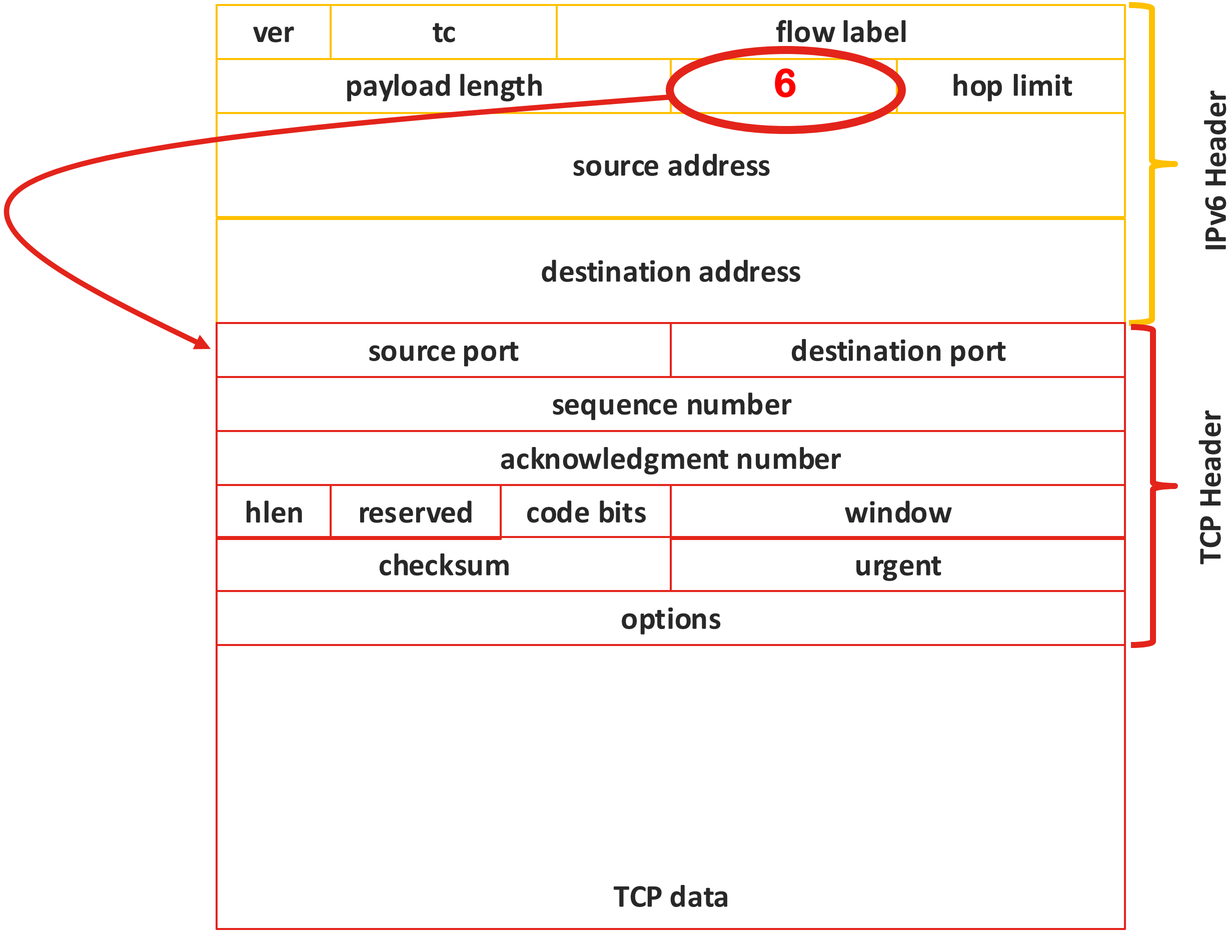
source address

destination address

IPv6 Header

SRv6

- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….



ver | tc | flow label

payload length | 6 | hop limit

source address

destination address

source port | destination port

sequence number

acknowledgment number

hlen | reserved | code bits | window

checksum | urgent

options

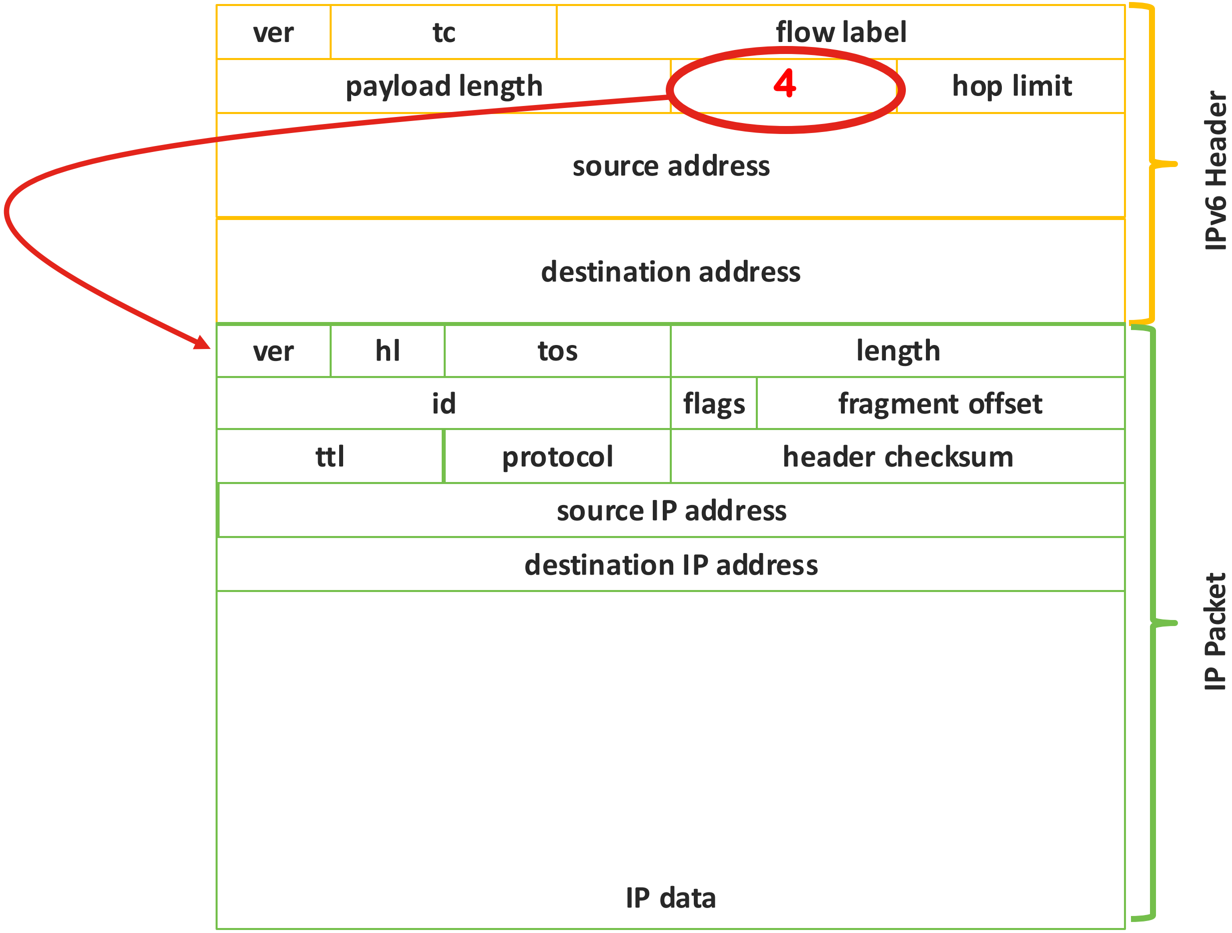
TCP data

IPv6 Header

TCP Header

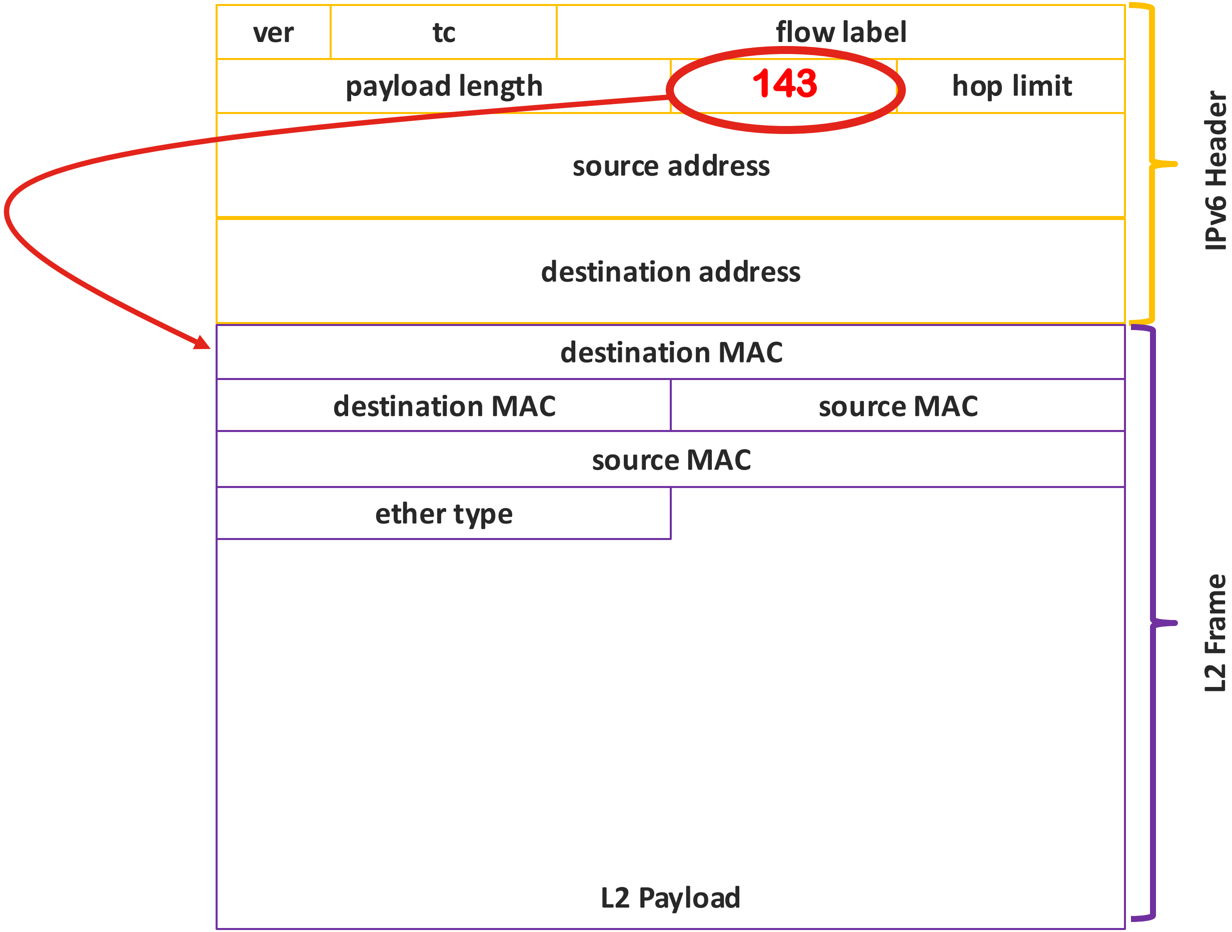
SRv6

- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….
  - IPv4, IPv6



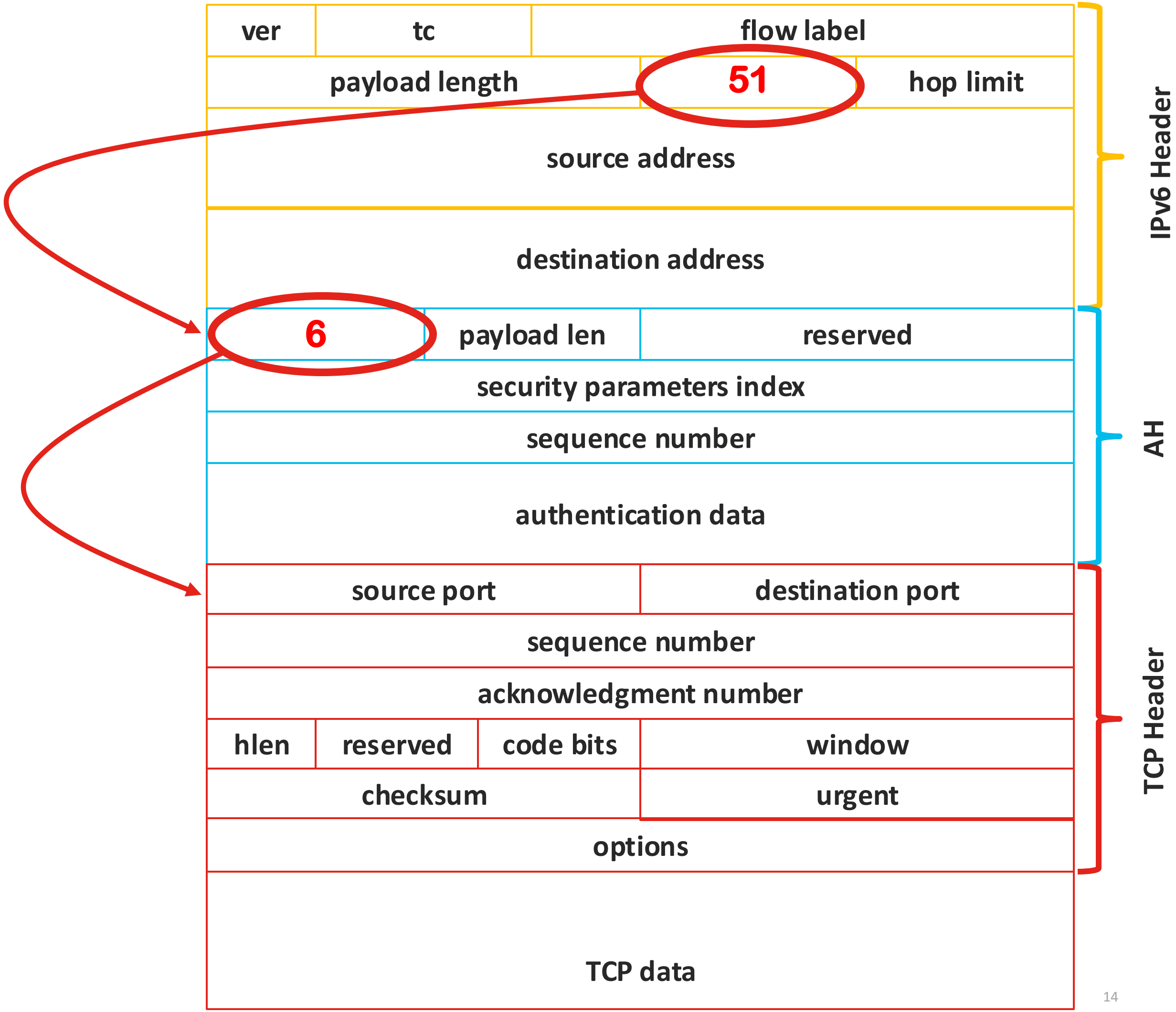
SRv6

- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….
  - IPv4, IPv6, L2



SRv6

- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….
  - IPv4, IPv6, L2
  - Hop by Hop, Dest. Options, Fragmentation, Authentication Header …



ver | tc | flow label

payload length | 51 | hop limit

source address

destination address

6 | payload len | reserved

security parameters index

sequence number

authentication data

source port | destination port

sequence number

acknowledgment number

hlen | reserved | code bits | window

checksum | urgent

options

TCP data

SRv6

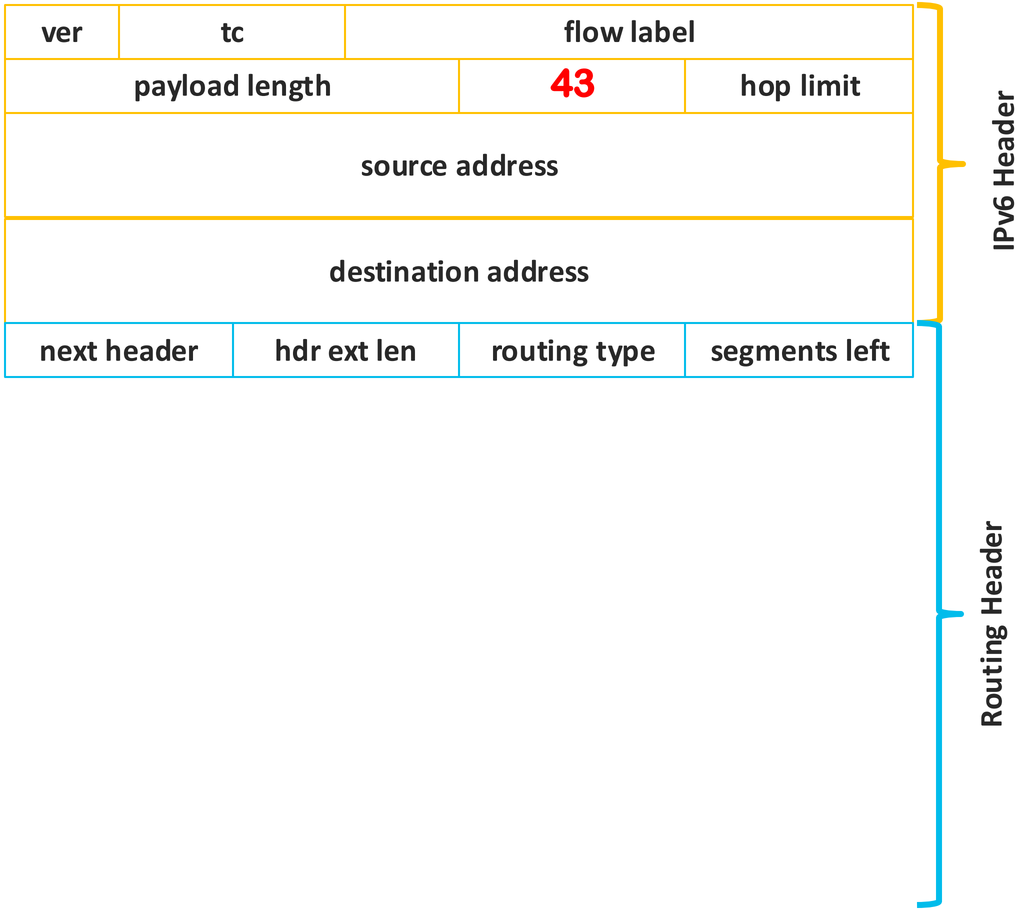
- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….
  - IPv4, IPv6, L2
  - Hop by Hop, Dest. Options, Fragmentation, Authentication Header …
  - Routing Header

0 Source Route (deprecated)

1 Nimrod (deprecated)

2 Type 2 (RFC 6275)

3 RPL (RFC 6554)



ver | tc | flow label

payload length | 43 | hop limit

source address

destination address

next header | hdr ext len | routing type | segments left

IPv6 Header

Routing Header

SRv6

- IPv6 Header
- Destination IP address
- Next header field:
  - TCP, UDP, ICMP….
  - IPv4, IPv6, L2
  - Hop by Hop, Dest. Options, Fragmentation, Authentication Header …
  - Routing Header

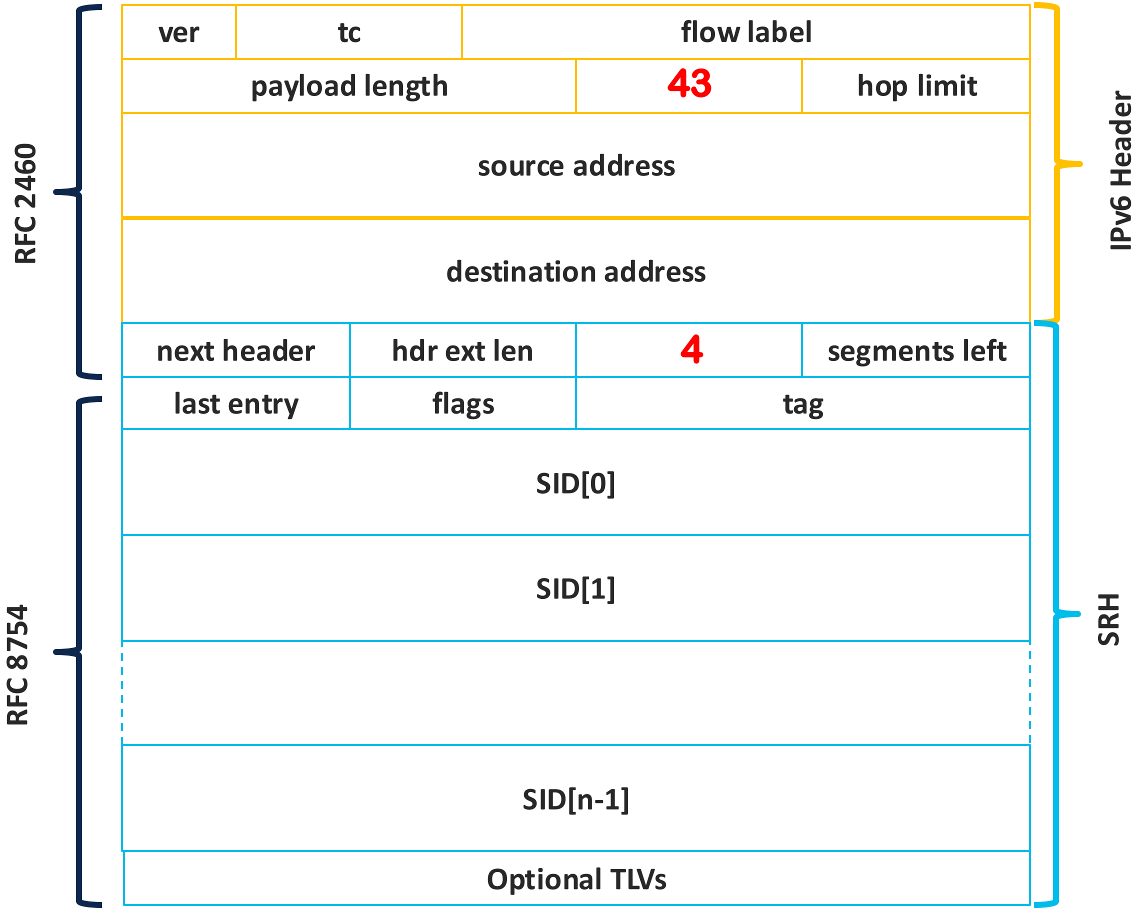
    0 Source Route (deprecated)

    1 Nimrod (deprecated)

    2 Type 2 (RFC 6275)

    3 RPL (RFC 6554)

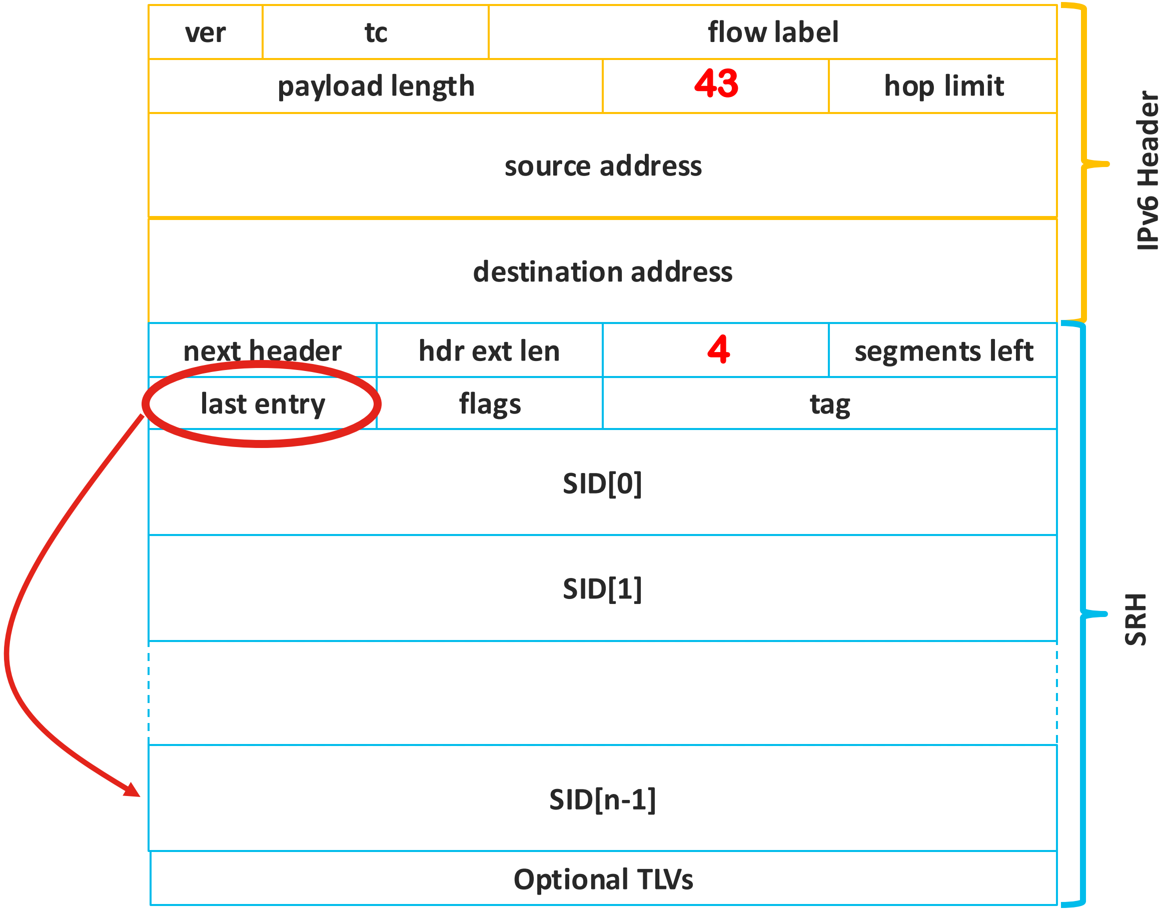
    4 SRH (RFC 8754)



ver	tc	flow label	
payload length		43	hop limit
source address			
destination address			
next header	hdr ext len	4	segments left
last entry	flags	tag	
SID[0]			
SID[1]			
SID[n-1]			
Optional TLVs			

SRH

- Segment Routing Header
- First Segment
  - Pointer to very first SID



ver | tc | flow label

payload length | 43 | hop limit

source address

destination address

next header | hdr ext len | 4 | segments left

last entry | flags | tag

SID[0]

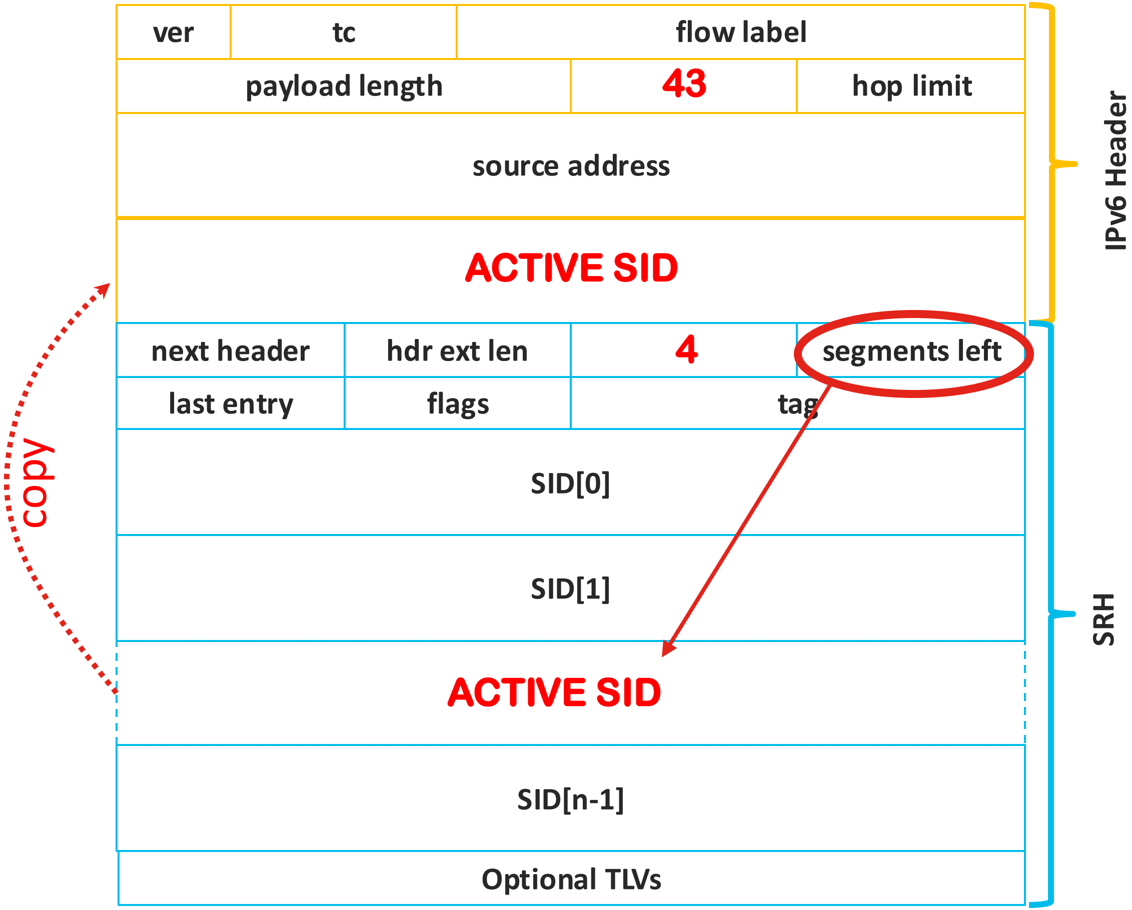
SID[1]

SID[n-1]

Optional TLVs

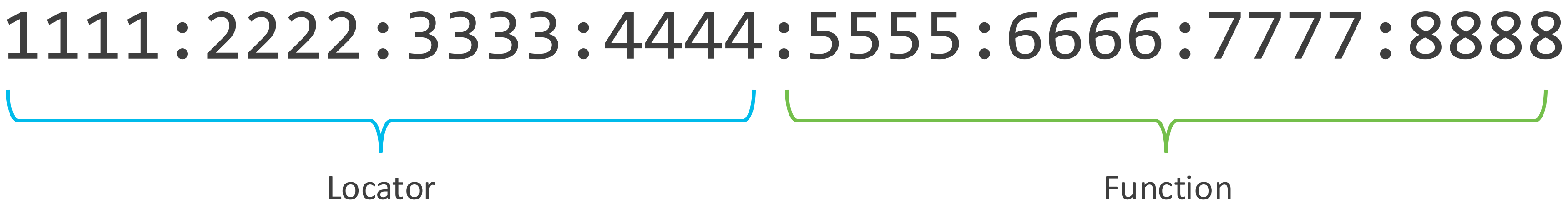
SRH

- Segment Routing Header
- First Segment
  - Pointer to very first SID
- Segments left
  - Pointer to Active SID
  - Active SID always in destination addr



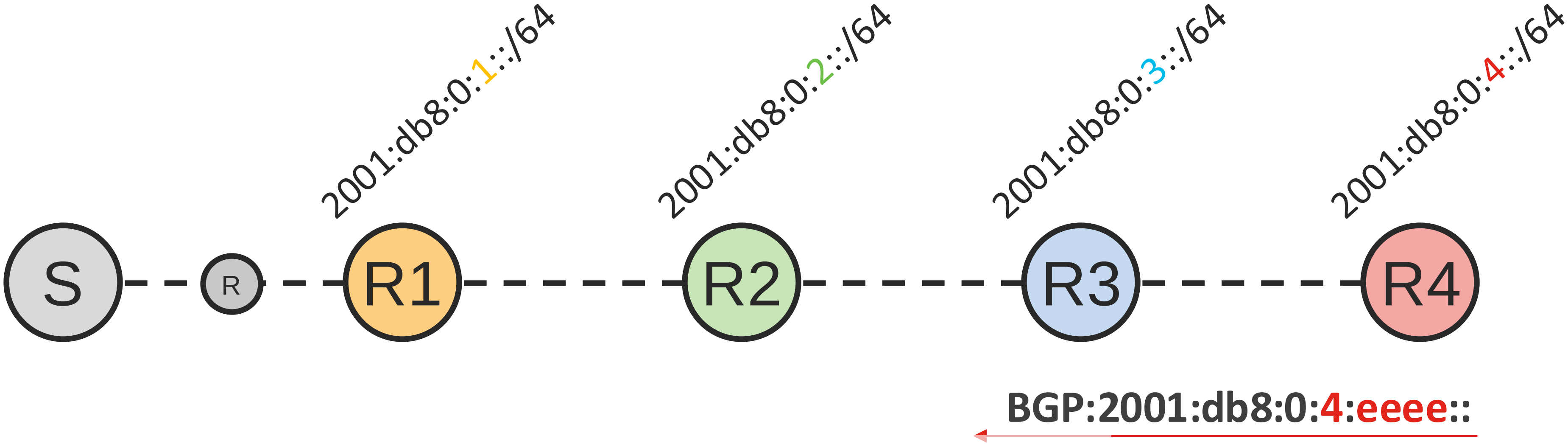
SID Structure -Locator

128 Bits Like IPv6 address but different semantics



1111:2222:3333:4444:5555:6666:7777:8888

SRv6 Full SID



```
SA:2001::1
DA:2001:db8:0:1:1::
NH:RH
Type:4(SRH)
NH:IPv4|SL:3
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:48::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

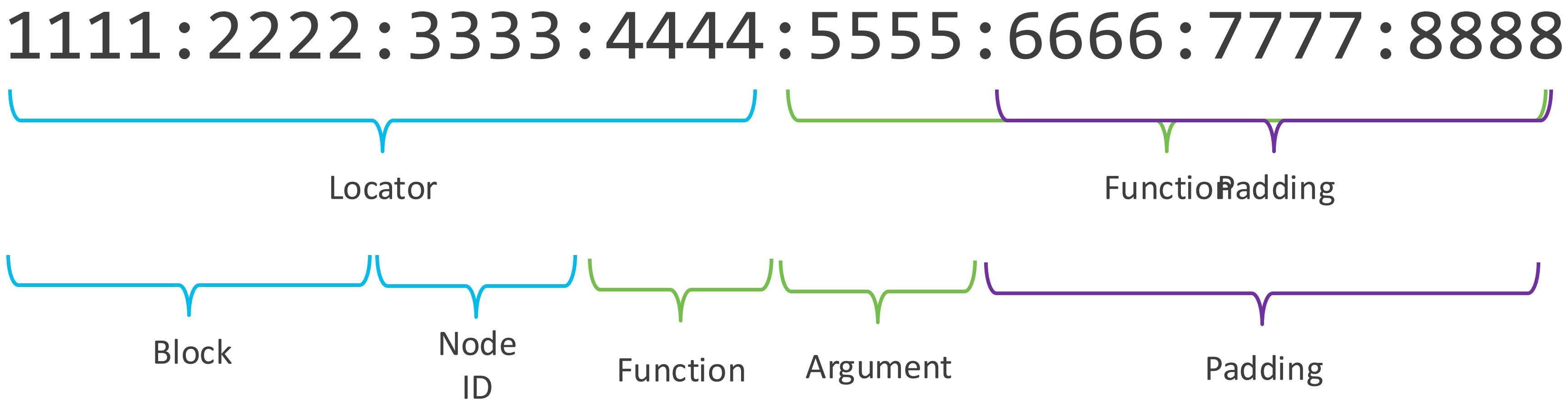
```
SA:2001::1
DA:2001:db8:0:2:1::
NH:RH
Type:4(SRH)
NH:IPv4|SL:2
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:48::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

```
SA:2001::1
DA:2001:db8:0:3:48::
NH:RH
Type:4(SRH)
NH:IPv4|SL:1
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:48::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

```
SA:2001::1
DA:2001:db8:0:4:eeee::
NH:IPv4
Type:4(SRH)
NH:IPv4|SL:0
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:48::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

SID Structure

128 Bits Like IPv6 address but different semantics



1111:2222:3333:4444:5555:6666:7777:8888

Locator

Function

Padding

Block

Node ID

Function

Argument

Padding

SRv6 uSID

```
SPRING                                                W. Cheng, Ed.
Internet-Draft                                         China Mobile
Intended status: Standards Track                        C. Filsfils
Expires: 15 July 2023                           Cisco Systems, Inc.
                                                              Z. Li
                                               Huawei Technologies
                                                        B. Decraene
                                                             Orange
                                                       F. Clad, Ed.
                                                Cisco Systems, Inc.
                                                    11 January 2023
```

Compressed SRv6 Segment List Encoding in SRH
draft-ietf-spring-srv6-srh-compression-03

Abstract

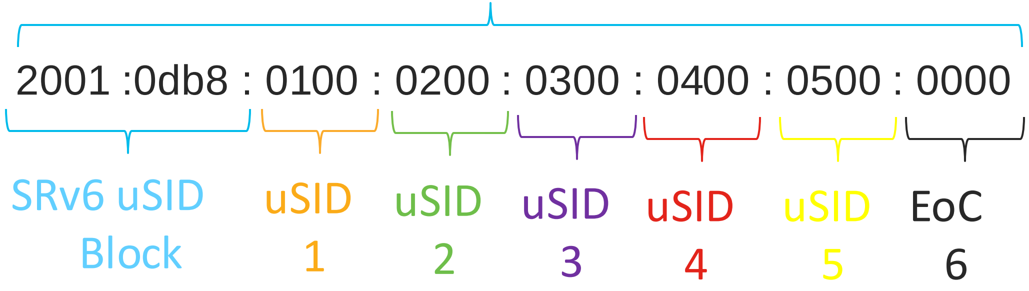
This document specifies new flavors for the SR endpoint behaviors defined in RFC 8986, which enable a compressed SRv6 Segment-List encoding in the Segment Routing Header (SRH).

SRv6 uSID format

: 0100 : =SRV6 uSID

16 bits here, but can be anything

SRV6 uSID Container



2001 :0db8 : 0100 : 0200 : 0300 : 0400 : 0500 : 0000

SRv6 uSID Block

uSID 1

uSID 2

uSID 3

uSID 4

uSID 5

EoC 6

32 bits here,
but can be anything

SRV6 Encapsulation

```
SA:2001::1
DA:2001:db8:0:4:1:0:0:0
NH:RH
```

```
Type:4(SRH)
NH:IPv4|SL:1
Segment List:
[0]: 2001:db8:0:5:45:0:0:0
[1]: 2001:db8:0:4:1:0:0:0
[2]: 2001:db8:0:3:48:0:0:0
[3]: 2001:db8:0:2:1:0:0:0
[4]: 2001:db8:0:1:42:0:0:0
```

```
SA:7.5.4.3
DA:11.6.19.71
Port:UDP
```

```
UDP Header/Data
```

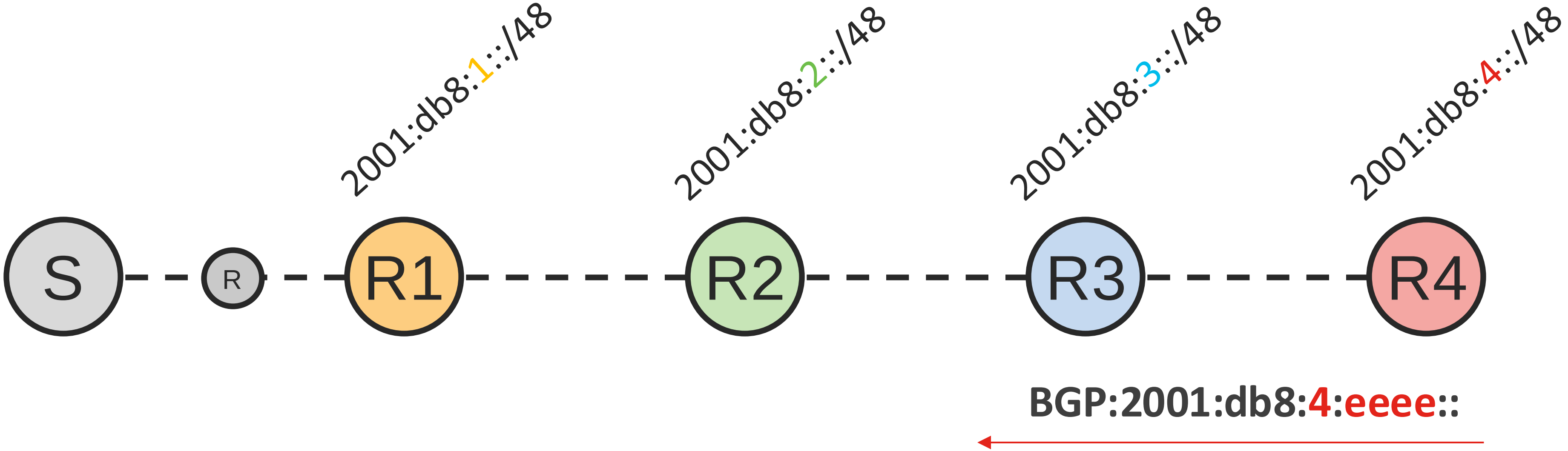
SRV6 uSID Encapsulation

```
SA:2001::1
DA:2001:db8:100:200:300:400:500::
NH:IPv4
```

```
SA:7.5.4.3
DA:11.6.19.71
Port:UDP
```

```
UDP Header/Data
```

SRv6 uSID F3216



BGP:2001:db8:4:eeee::

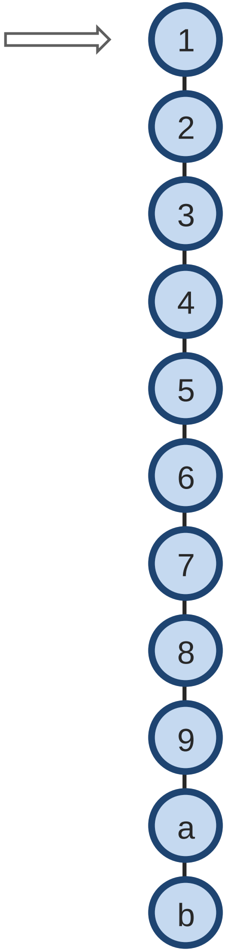
```
SA:2001::1
DA:2001:db8:1:2:3:e000:4:eeee
NH:IPV4
```

```
SA:2001::1
DA:2001:db8:2:3:e000:4:eeee::
NH:IPV4
```

```
SA:2001::1
DA:2001:db8:3:e000:4:eeee::
NH:IPV4
```

```
SA:2001::1
DA:2001:db8:4:eeee::
NH:IPV4
```

SRv6 uSID More Than 6 SIDs?



100->200->300->400->500->600->700->800->900->a00->b00

Container 1 2001 : 0db8 : 0100 : 0200 : 0300 : 0400 : 0500 : 0600

Container 2 2001 : 0db8 : 0700 : 0800 : 0900 : 0a00 : 0b00 : 0000

```
SA:2001::1
DA:2001:db8:100:200:300:400:500:600
NH: RH
```

```
Type:4(SRH)
NH:IPv4|SL:1
Segment List:
[0]: 2001:db8:700:800:900:a00:b00::
```

```
SA:7.5.4.3
DA:11.6.19.71
Port:UDP
```

```
UDP Header/Data
```

Shift & Forward

SRv6 uSID More Than 6 SIDs?

100->200->300->400->500->600->700->800->900->a00->b00

Container 1 2001 : 0db8 : 0100 : 0200 : 0300 : 0400 : 0500 : 0600

Container 2 2001 : 0db8 : 0700 : 0800 : 0900 : 0a00 : 0b00 : 0000

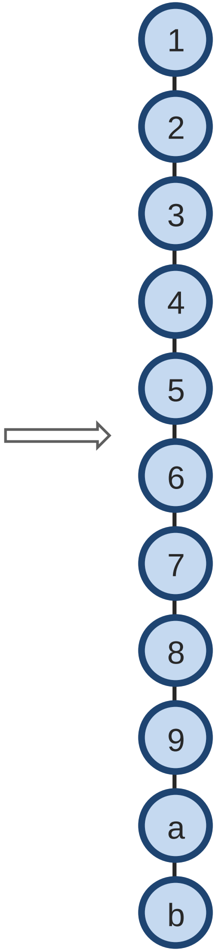
SA:2001::1
DA:2001:db8:600:0::
NH: RH

Type:4(SRH)
NH:IPv4|SL:1
Segment List:
[0]: 2001:db8:700:800:900:a00:b00::

SA:7.5.4.3
DA:11.6.19.71
Port:UDP

UDP Header/Data

Shift & Forward
END of Container
-> is there SRH?
Decrement SL
Copy New SID (Container)
PSP



SRv6 uSID More Than 6 SIDs?

100->200->300->400->500->600->700->800->900->a00->b00

Container 1	2001 : 0db8 : 0100 : 0200 : 0300 : 0400 : 0500 : 0600
Container 2	2001 : 0db8 : 0700 : 0800 : 0900 : 0a00 : 0b00 : 0000

```
SA:2001::1
DA:2001:db8:700:800:900:a00:b00::
NH:IPv4
```

```
SA:7.5.4.3
DA:11.6.19.71
Port:UDP
```

```
UDP Header/Data
```

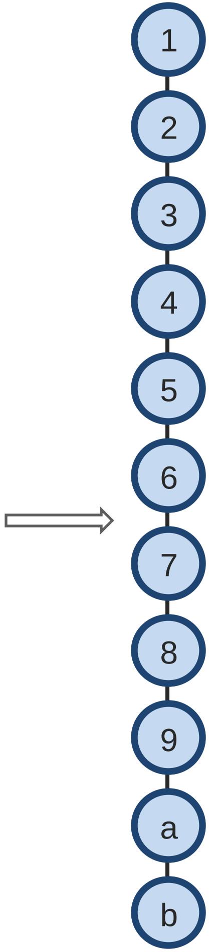
END of Container

-> is there SRH?

Decrement SL

Copy New SID (Container)

PSP



SRv6 uSID More Than 6 SIDs?

100->200->300->400->500->600->700->800->900->a00->b00

Container 1 2001 : 0db8 : 0100 : 0200 : 0300 : 0400 : 0500 : 0600

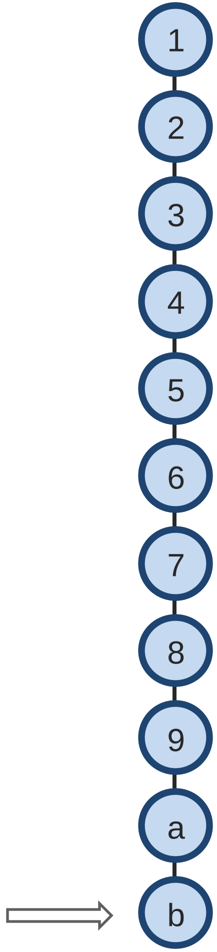
Container 2 2001 : 0db8 : 0700 : 0800 : 0900 : 0a00 : 0b00 : 0000

SA:2001::1
DA:2001:db8:b00::
NH:IPv4

SA:7.5.4.3
DA:11.6.19.71
Port:UDP

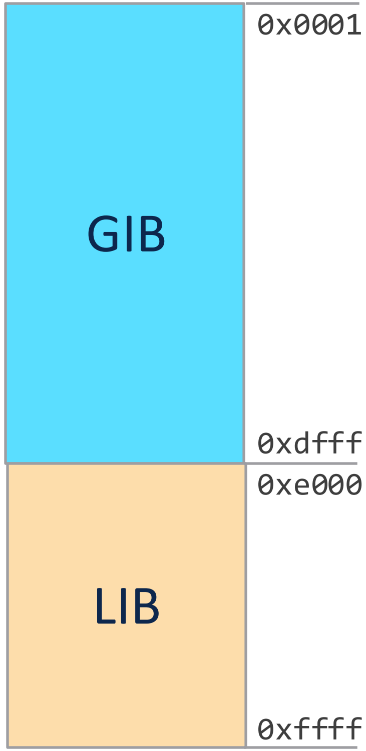
UDP Header/Data

Shift & Forward



Sets, Global ID Block (GIB), Local ID Block (LIB)

- Within a Block, SIDs are allocated: `FCBB:BB00:XXXX::/48`
- SID can be:
  - Global: shortest path to a node – globally unique
  - Local: a local function – not globally unique



SRv6 uSID Configuration

```
segment-routing
 srv6
  locators
   locator MAIN
    micro-segment behavior unode psp-usd
    prefix fcbb:bb00:1::/48
```

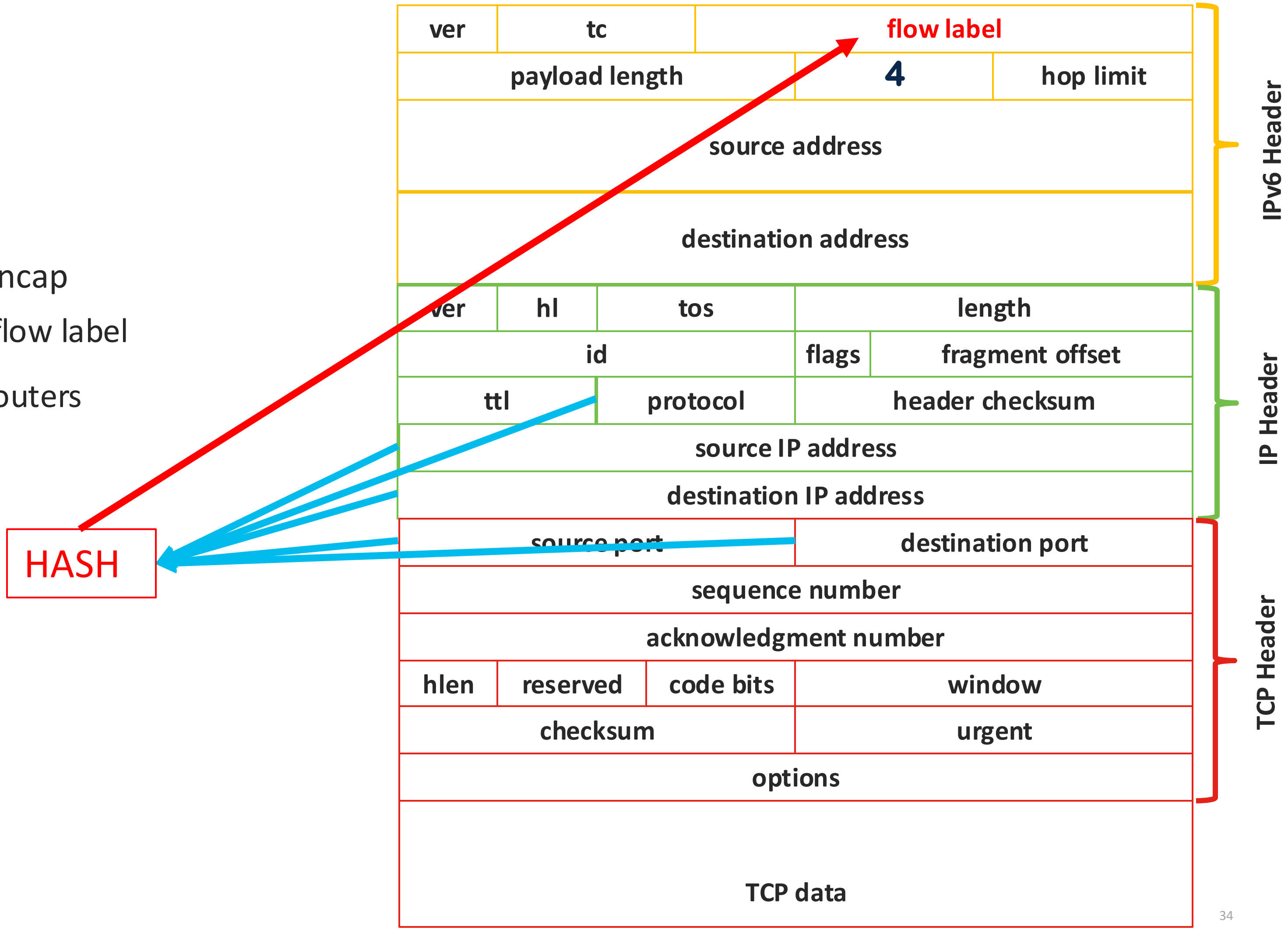
Name to reference

uSID

Locator Prefix

Flow Label

- What for?
- Entropy encoding encap
  - 5 Tuple Hash into flow label
- Used for Hash on P routers



HASH

ver | tc | flow label

payload length | 4 | hop limit

source address

destination address

IPv6 Header

ver | hl | tos | length

id | flags | fragment offset

ttl | protocol | header checksum

source IP address

destination IP address

IP Header

source port | destination port

sequence number

acknowledgment number

hlen | reserved | code bits | window

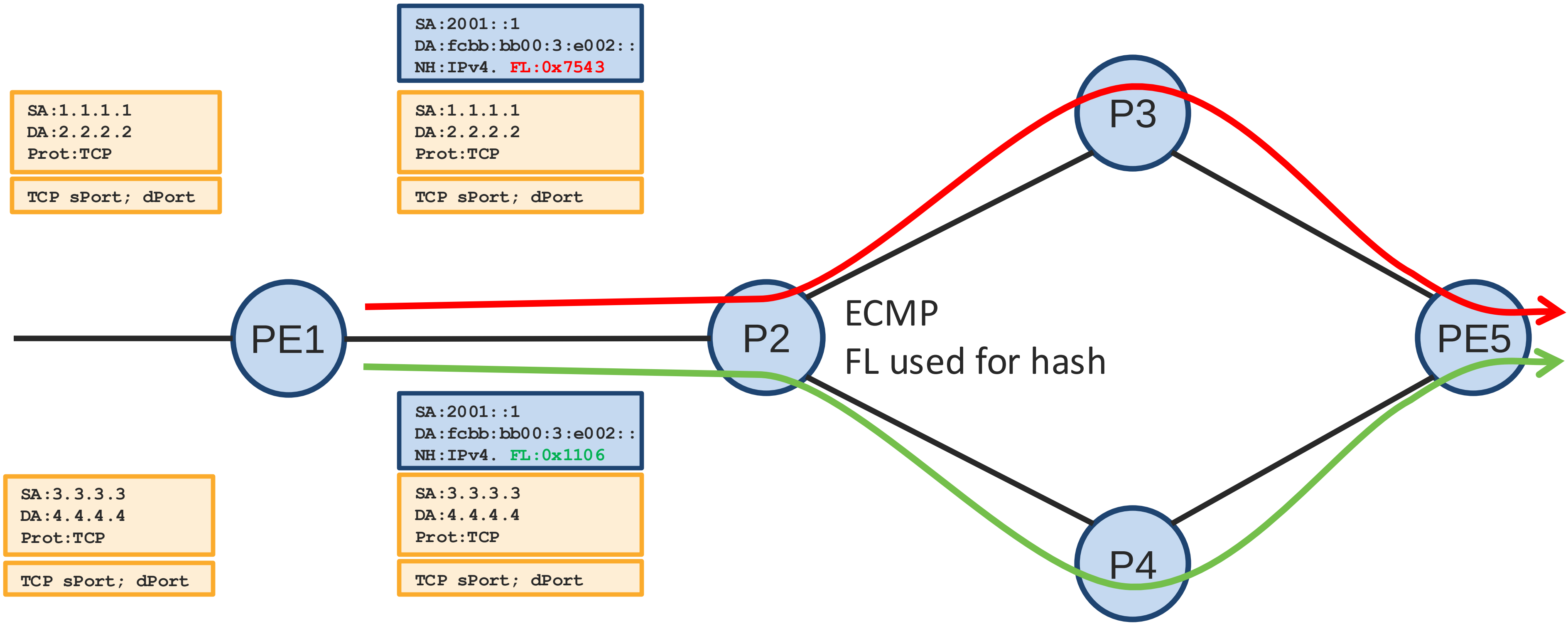
checksum | urgent

options

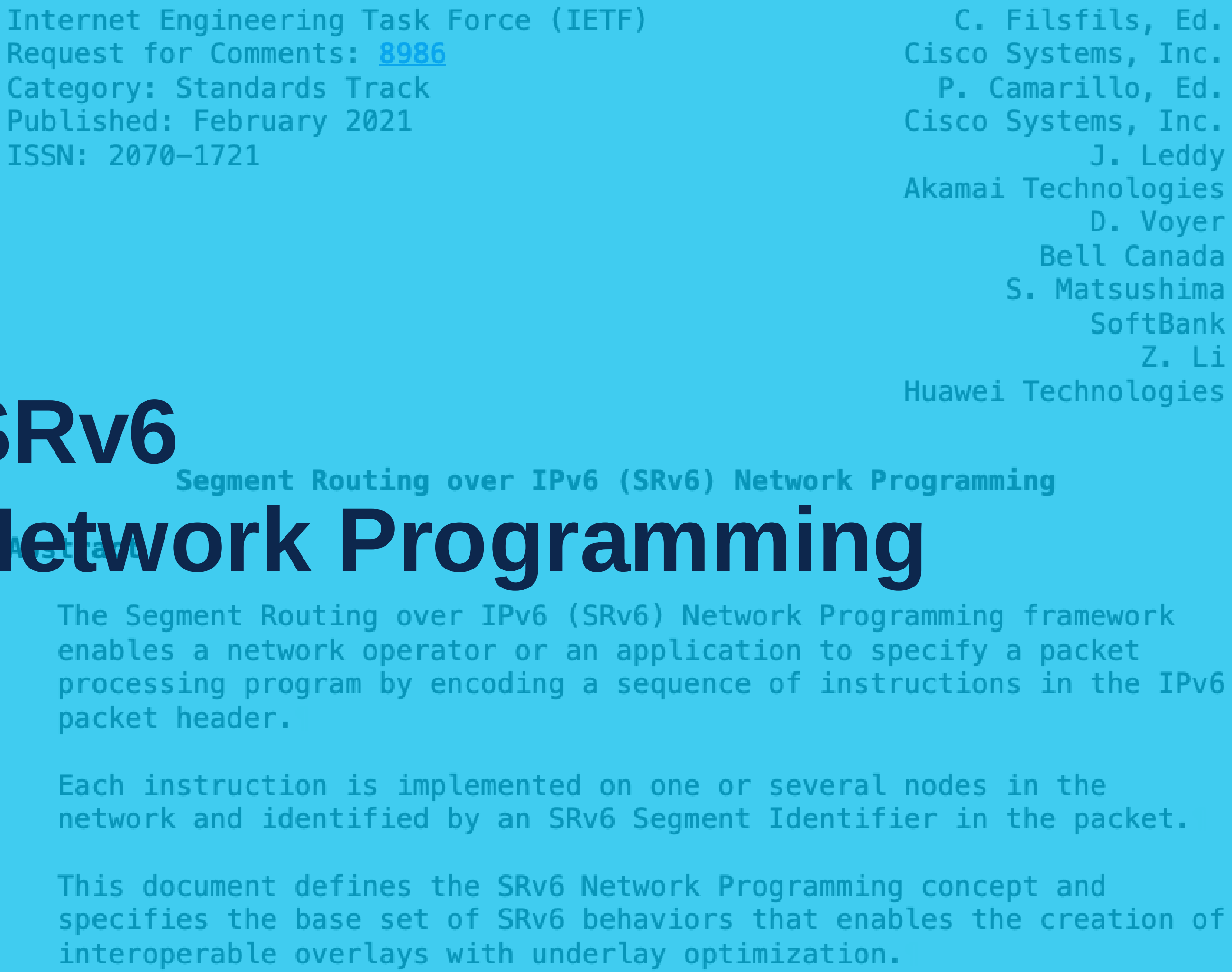
TCP data

TCP Header

Flow Label


SA:1.1.1.1
DA:2.2.2.2
Prot:TCP
TCP sPort; dPort
SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4. FL:0x7543
SA:1.1.1.1
DA:2.2.2.2
Prot:TCP
TCP sPort; dPort
P3
PE1
P2
PE5
ECMP
FL used for hash
SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4. FL:0x1106
SA:3.3.3.3
DA:4.4.4.4
Prot:TCP
TCP sPort; dPort
SA:3.3.3.3
DA:4.4.4.4
Prot:TCP
TCP sPort; dPort
P4

SRv6 Network Programming



Internet Engineering Task Force (IETF)
Request for Comments: 8986
Category: Standards Track
Published: February 2021
ISSN: 2070-1721

C. Filsfils, Ed.
Cisco Systems, Inc.
P. Camarillo, Ed.
Cisco Systems, Inc.
J. Leddy
Akamai Technologies
D. Voyer
Bell Canada
S. Matsushima
SoftBank
Z. Li
Huawei Technologies

Segment Routing over IPv6 (SRv6) Network Programming

Abstract

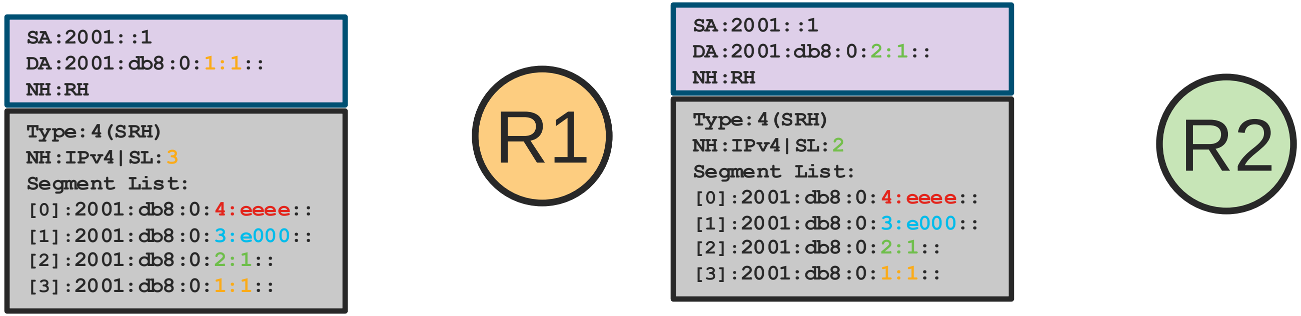
The Segment Routing over IPv6 (SRv6) Network Programming framework enables a network operator or an application to specify a packet processing program by encoding a sequence of instructions in the IPv6 packet header.

Each instruction is implemented on one or several nodes in the network and identified by an SRv6 Segment Identifier in the packet.

This document defines the SRv6 Network Programming concept and specifies the base set of SRv6 behaviors that enables the creation of interoperable overlays with underlay optimization.

END– Default endpoint (Node SID)

- *Decrement SL*
- *Copy Active SID*
- *Forward*



```
SA:2001::1
DA:2001:db8:0:1:1::
NH:RH

Type:4(SRH)
NH:IPv4|SL:3
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:e000::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

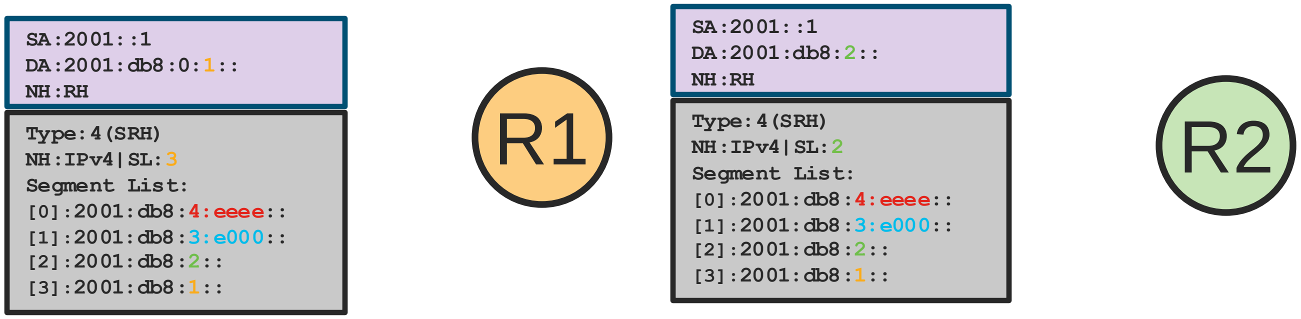
```
SA:2001::1
DA:2001:db8:0:2:1::
NH:RH

Type:4(SRH)
NH:IPv4|SL:2
Segment List:
[0]:2001:db8:0:4:eeee::
[1]:2001:db8:0:3:e000::
[2]:2001:db8:0:2:1::
[3]:2001:db8:0:1:1::
```

- Different Flavors:
  - End
  - End with PSP
  - End with USP
  - End with PSP & USP
  - End with USD
  - End with PSP & USD
  - End with USP & USD
  - End with PSP, USP & USD

- End with **NEXT**-ONLY-CSID
- End with **NEXT**-CSID
- End with **NEXT**-CSID & PSP
- End with **NEXT**-CSID & USP
- End with **NEXT**-CSID, PSP & USP
- End with **NEXT**-CSID & USD
- End with **NEXT**-CSID, PSP & USD
- End with **NEXT**-CSID, USP & USD
- End with **NEXT**-CSID, PSP, USP & USD

uN=END with Next – Default endpoint (Node SID)

- *Decrement SL*
- *Copy Active SID*
- *Forward*



```
SA:2001::1
DA:2001:db8:0:1::
NH:RH
```

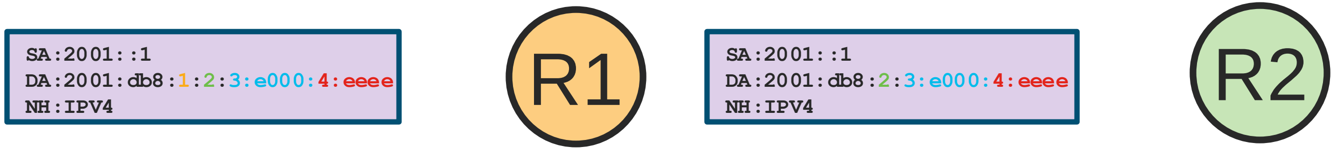
```
Type:4(SRH)
NH:IPv4|SL:3
Segment List:
[0]:2001:db8:4:eeee::
[1]:2001:db8:3:e000::
[2]:2001:db8:2::
[3]:2001:db8:1::
```

```
SA:2001::1
DA:2001:db8:2::
NH:RH
```

```
Type:4(SRH)
NH:IPv4|SL:2
Segment List:
[0]:2001:db8:4:eeee::
[1]:2001:db8:3:e000::
[2]:2001:db8:2::
[3]:2001:db8:1::
```

Better way:

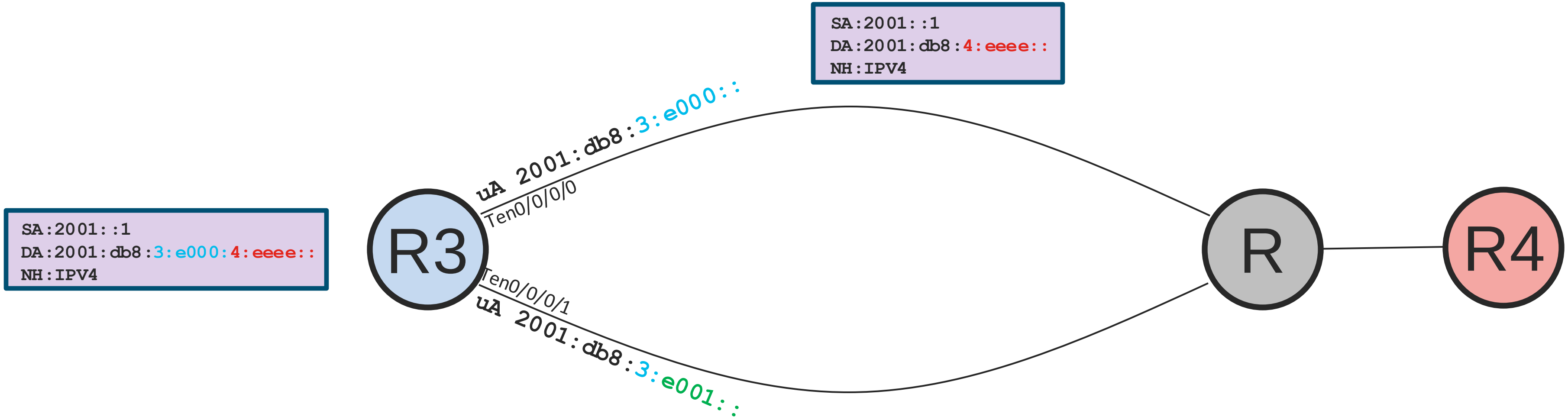
- *Shift & Forward*



```
SA:2001::1
DA:2001:db8:1:2:3:e000:4:eeee
NH:IPV4
```

```
SA:2001::1
DA:2001:db8:2:3:e000:4:eeee
NH:IPV4
```

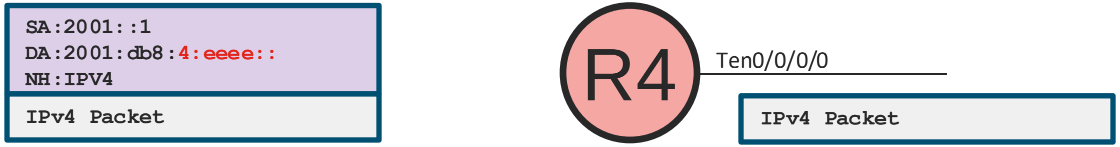
uA=END.X with Next – (Adjacency SID)



- *Shift & Forward to* ***SPECIFIC INTERFACE***

uDX4=END.DX4, uDX6=END.DX6, uDX2=END.DX2 Endpoint with **D**ecapsulation and **X**connect

- *Decapsulate and Forward to **SPECIFIC INTERFACE***
- ***Same as Per CE Label Allocation***
- *Must be last function in SID list*



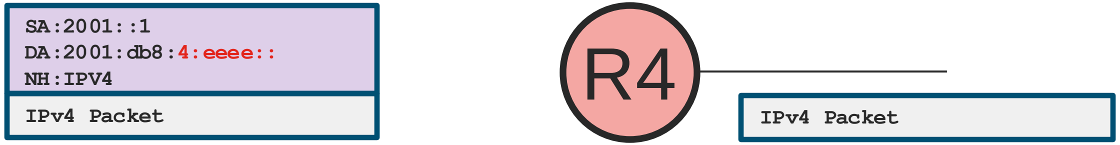
```
SA:2001::1
DA:2001:db8:4:eeee::
NH:IPV4
```

```
IPv4 Packet
```

```
IPv4 Packet
```

uDT4=END.DT4, uDT6=END.DT6
Endpoint with **D**ecapsulation and **T**able Lookup

- *Decapsulate and Table Lookup* ***(VRF)***
- ***Same as Per VRF Label Allocation (aggregate label)***
- *Must be last function in SID list*



SRv6 functions: Steering and Services

Codename		Behavior	
End	uN	Endpoint	[Node SID]
End.X	uA	Endpoint with Layer-3 cross-connect	[Adj SID]
End.B6.Insert	uB6.Insert	Endpoint bound to an SRv6 policy	[BSID]
End.B6.Encap	uB6.Encaps	Endpoint bound to an SRv6 encapsulation policy	[BSID]
End.DX6	uDX6	Endpoint with decapsulation and IPv6 cross-connect	[L3VPN Per-CE]
End.DX4	uDX4	Endpoint with decapsulation and IPv4 cross-connect	[L3VPN Per-CE]
End.DT6	uDT6	Endpoint with decapsulation and specific IPv6 table lookup	[L3VPN Per-VRF]
End.DT4	uDT4	Endpoint with decapsulation and specific IPv4 table lookup	[L3VPN Per-VRF]
End.DT46	uDT46	Endpoint with decapsulation and specific IPv4&v6 table lookup	[L3VPN Per-VRF]
End.DX2	uDX2	Endpoint with decapsulation and L2 cross-connect	[E-LINE]
End.DT2U/M	uDT2U/M	Endpoint with decapsulation and L2 unicast lookup / flooding	[E-LAN]
End.DTM	uDTM	Endpoint with decapsulation and MPLS table lookup	[Interworking]
H.Insert / H.Encaps		Headend with Insertion / Encapsulation of / into an SRv6 policy	[TiLFA]
H. Encaps.L2		H.Encaps Applied to Received L2 Frames	[L2 Port Mode]
H.Encaps.M		H.Encaps Applied to MPLS Label Stack	[Interworking]

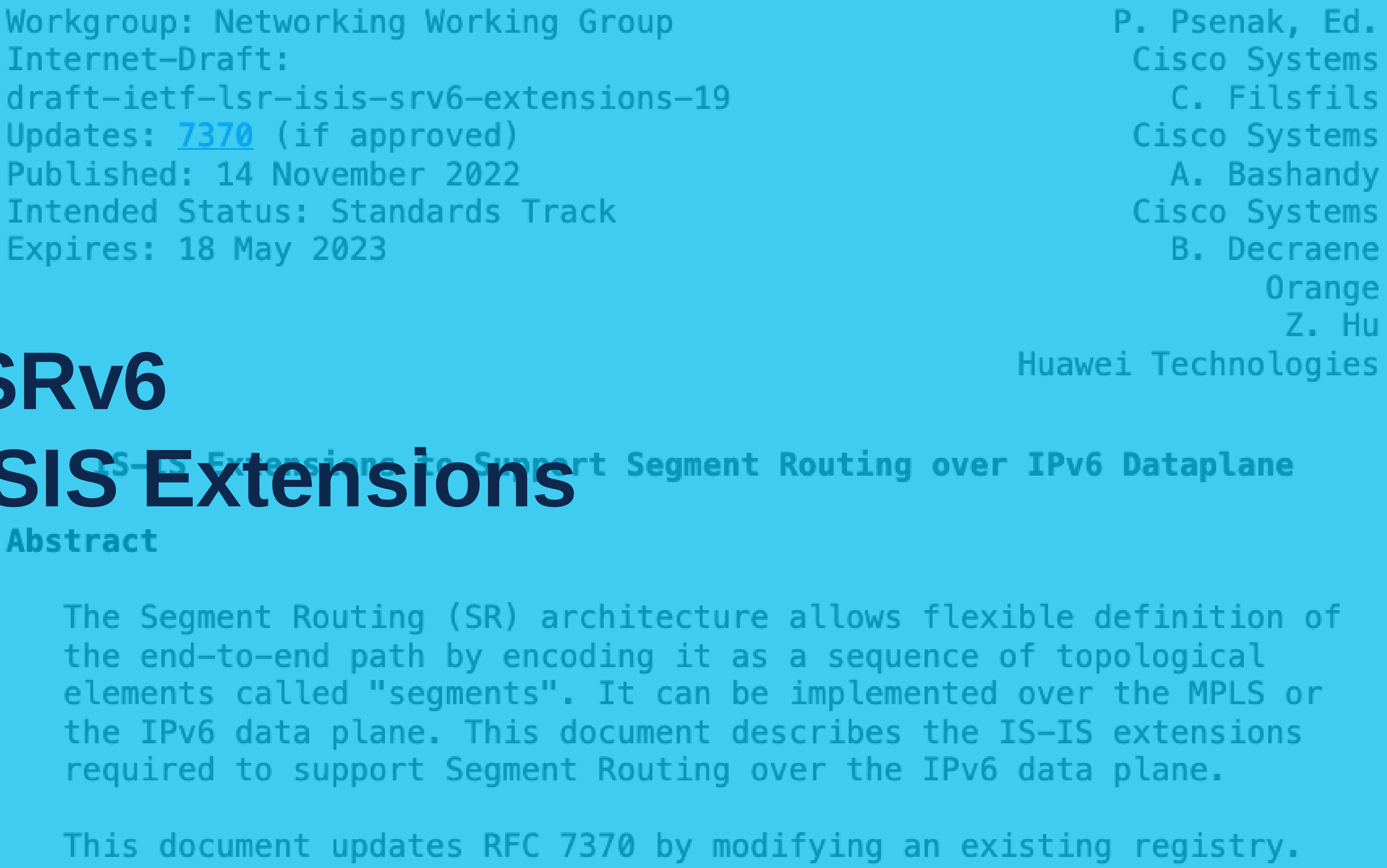
Functions might be signaled differently

Signalling	IGP	BGP-LS	BGP-IP/VPN
End, uN	Yes	Yes	
End.X, uA	Yes	Yes	
End.T	Yes	Yes	
End.DX4,uDX4		Yes	Yes
End.DX6,uDX6	Yes	Yes	Yes
End.DX2,uDX2		Yes	Yes
END.DT4,uDT4		Yes	Yes
End.DT6,uDT6	Yes	Yes	Yes
End.B		Yes	

Signalling	IGP	BGP-LS	BGP-IP/VPN
H.insert		Yes	
H.Encap		Yes	

Locator – routing table

SRv6
ISIS Extensions



Workgroup: Networking Working Group
Internet-Draft:
draft-ietf-lsr-isis-srv6-extensions-19
Updates: 7370 (if approved)
Published: 14 November 2022
Intended Status: Standards Track
Expires: 18 May 2023

P. Psenak, Ed.
Cisco Systems
C. Filsfils
Cisco Systems
A. Bashandy
Cisco Systems
B. Decraene
Orange
Z. Hu
Huawei Technologies

IS-IS Extensions to Support Segment Routing over IPv6 Dataplane

Abstract

The Segment Routing (SR) architecture allows flexible definition of the end-to-end path by encoding it as a sequence of topological elements called "segments". It can be implemented over the MPLS or the IPv6 data plane. This document describes the IS-IS extensions required to support Segment Routing over the IPv6 data plane.

This document updates RFC 7370 by modifying an existing registry.

ISIS for SRv6

LSP (Link State Packet):

TLVs:

Hostname: r2

Interfaces: Hu0/0/0/0 uA:fcbb:0:2:e001::

Structure: BL=32;NL=16;FL=16;AL=0

Hu0/0/0/1 uA:fcbb:0:2:e002::

Structure: BL=32;NL=16;FL=16;AL=0

Lo0

Neighbors： r1

r3

IP addresses: fcbb:0:2::1/128

2001:12::2/64

Locator: fcbb:0:2::/48

uN:fcbb:0:2::

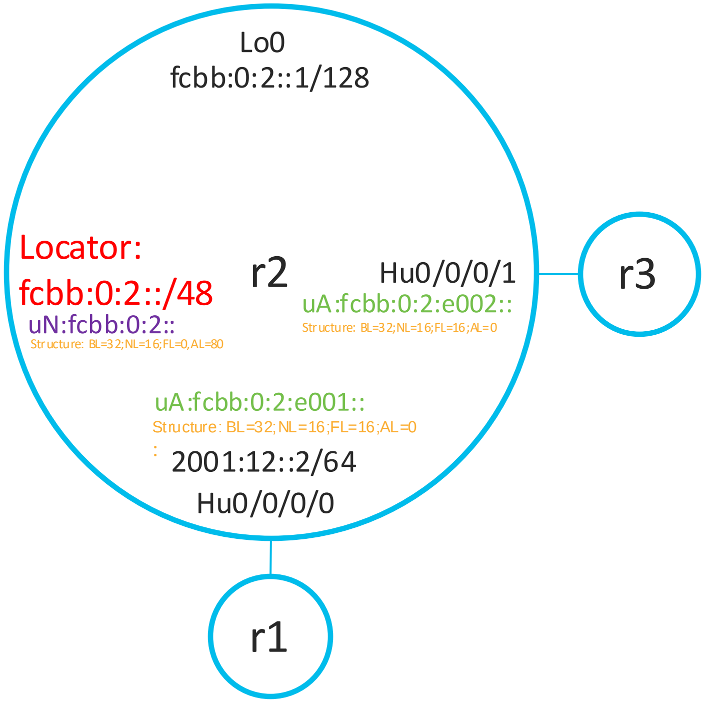
Structure: BL=32;NL=16;FL=0,AL=80

Capabilities: Algorithms ……

SIDs can insert

SIDs can decap

……….



ISIS LSP Example

```
IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
r2.00-00              0x00000009   0x4f06        1145 /1200         0/0/0
  Area Address:   49
  NLPID:          0x8e
  Hostname:       r1
  IPv6 Address:   2001::2
  Metric: 10         MT (IPv6 Unicast) IPv6 fcbb:bb00:2::1/128
    Prefix Attribute Flags: X:0 R:0 N:1 E:0 A:0
  Metric: 1          MT (IPv6 Unicast) IPv6 fcbb:bb00:2::/48
    Prefix Attribute Flags: X:0 R:0 N:0 E:0 A:0
  MT:             IPv6 Unicast                             0/0/0
  SRv6 Locator:   MT (IPv6 Unicast) fcbb:bb00:2::/48 D:0 Metric: 0 Algorithm: 0
    Prefix Attribute Flags: X:0 R:0 N:0 E:0 A:0
    END SID: fcbb:bb00:2:: uN (PSP/USD)
      SID Structure:
        Block Length: 32, Node-ID Length: 16, Func-Length: 0, Args-Length: 0
  Router Cap:     0.0.0.0 D:0 S:0
    IPv6 Router ID: 2001::2
    SR Algorithm:
      Algorithm: 0
      Algorithm: 1
    SRv6: O:0
    Node Maximum SID Depth:
      SRH Max SL:        3
      SRH Max End Pop:   3
      SRH Max T.insert:  3
      SRH Max T.encaps:  4
      SRH Max End D:     4
  Metric: 10         MT (IPv6 Unicast) IS-Extended r1.00
    Local Interface ID: 6, Remote Interface ID: 6
    Interface IPv6 Address: 2001:12::2
    Neighbor IPv6 Address: 2001:12::1
    END.X SID: fcbb:bb00:2:e001:: B:0 S:0 P:0 uA (PSP/USD) Alg:0
      SID Structure:
        Block Length: 32, Node-ID Length: 16, Func-Length: 16, Args-Length: 0
Total Level-2 LSP count: 1     Local Level-2 LSP count: 0
```

Locator

Capabilities

END

END.X

SID Structure

SRv6 ISIS Configuration

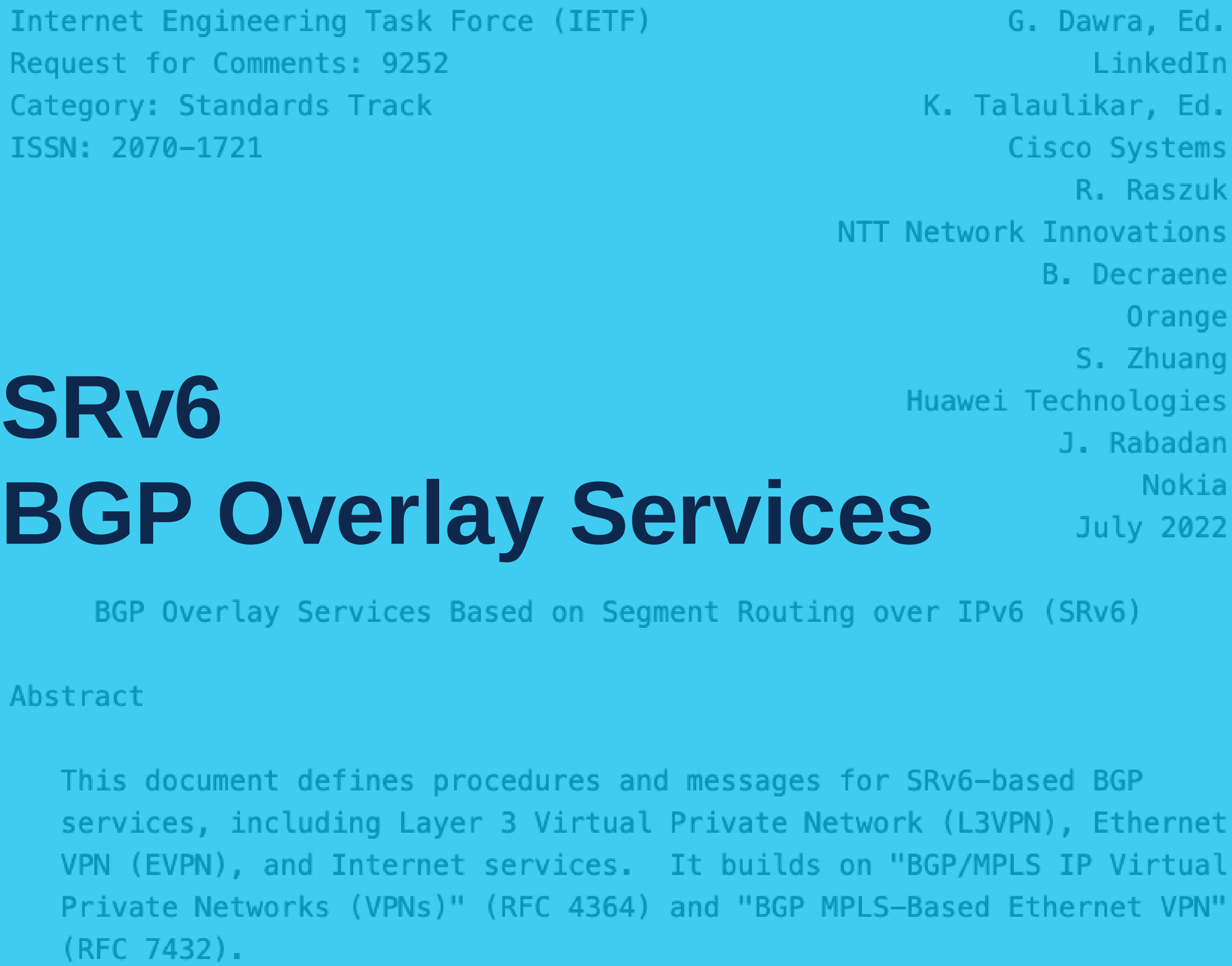
```
router isis 1
 address-family ipv6 unicast
  segment-routing srv6
   locator MAIN
```

Name of the Locator

This will result in:

- Locator is advertised
- uN function is advertised
- uA for each ISIS interface is allocated and advertised

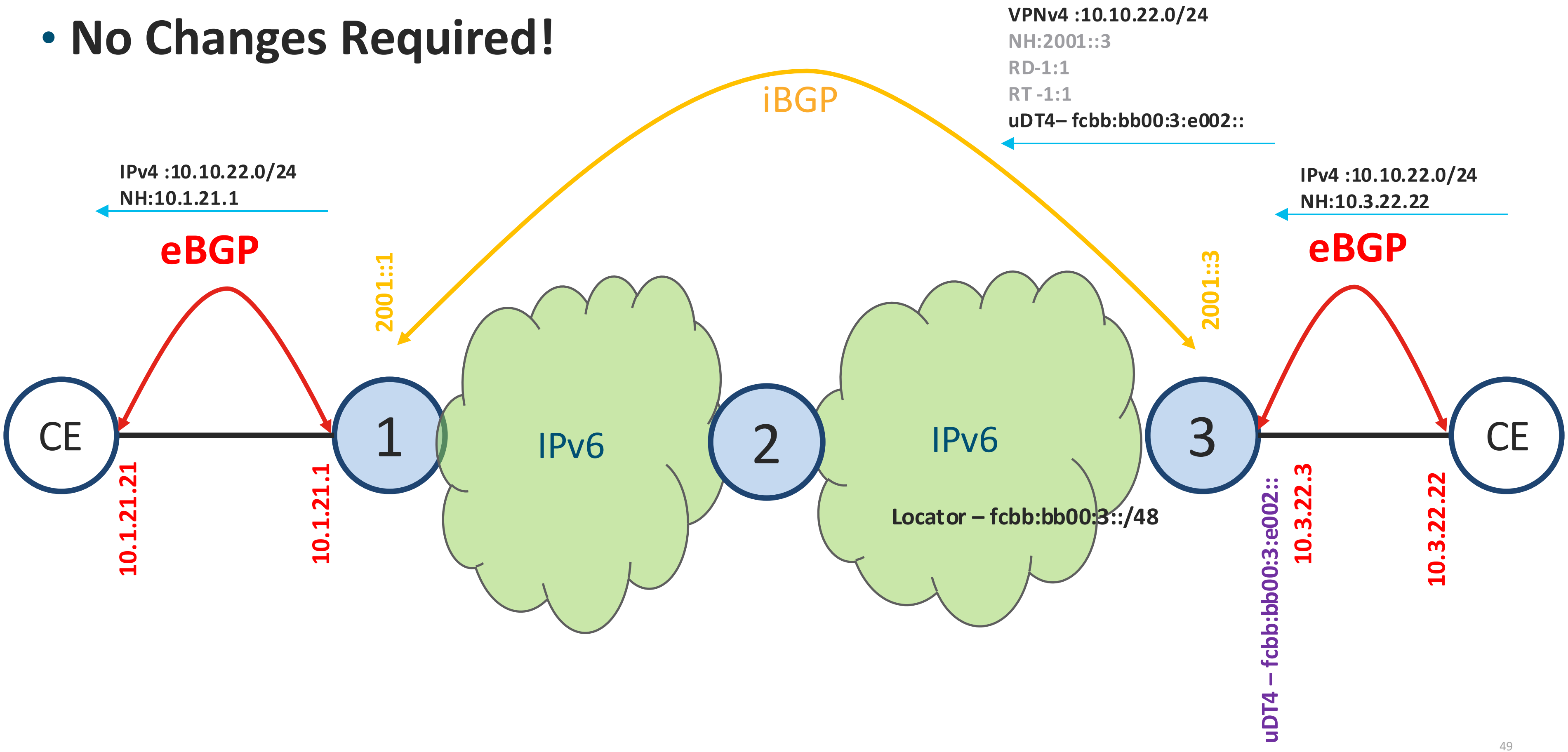
SRv6
BGP Overlay Services



```
Internet Engineering Task Force (IETF)                    G. Dawra, Ed.
Request for Comments: 9252                                     LinkedIn
Category: Standards Track                           K. Talaulikar, Ed.
ISSN: 2070-1721                                          Cisco Systems
                                                              R. Raszuk
                                               NTT Network Innovations
                                                            B. Decraene
                                                                 Orange
                                                              S. Zhuang
                                                    Huawei Technologies
                                                             J. Rabadan
                                                                  Nokia
                                                              July 2022

     BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)

Abstract

   This document defines procedures and messages for SRv6-based BGP
   services, including Layer 3 Virtual Private Network (L3VPN), Ethernet
   VPN (EVPN), and Internet services.  It builds on "BGP/MPLS IP Virtual
   Private Networks (VPNs)" (RFC 4364) and "BGP MPLS-Based Ethernet VPN"
   (RFC 7432).
```

BGP

- **No Changes Required!**

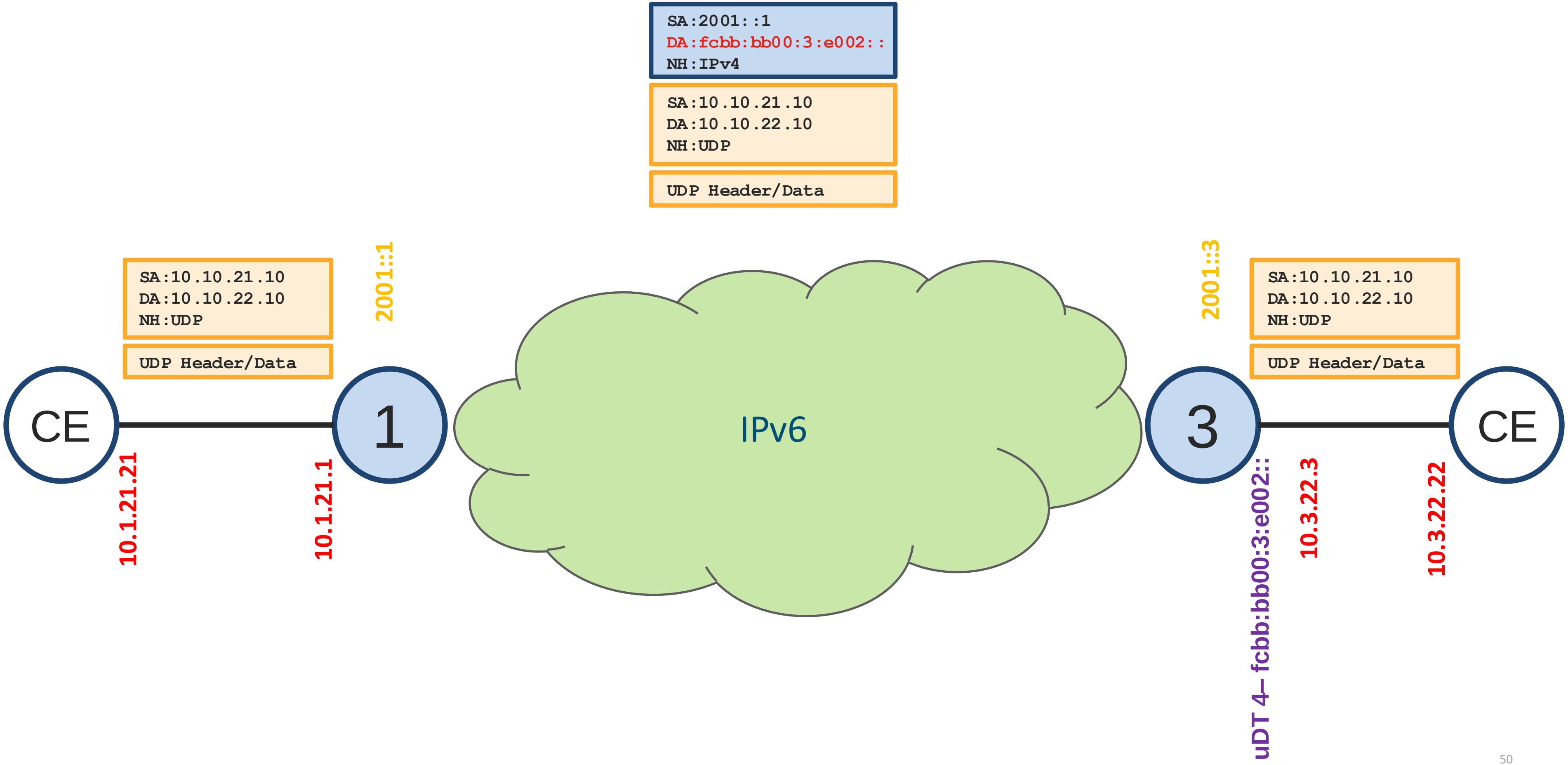


VPNv4 :10.10.22.0/24
NH:2001::3
RD-1:1
RT -1:1
uDT4– fcbb:bb00:3:e002::

IPv4 :10.10.22.0/24
NH:10.1.21.1

IPv4 :10.10.22.0/24
NH:10.3.22.22

L3 VPN Dataplane



SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4

SA:10.10.21.10
DA:10.10.22.10
NH:UDP

UDP Header/Data

SA:10.10.21.10
DA:10.10.22.10
NH:UDP

UDP Header/Data

SA:10.10.21.10
DA:10.10.22.10
NH:UDP

UDP Header/Data

SRv6 L3 VPN Configuration

```
router bgp 1
 address-family vpnv4 unicast
 vrf BestEffort
  rd 1:1
  address-family ipv4 unicast
   segment-routing srv6
    locator MAIN
    alloc mode per-vrf
```

Name of the Locator

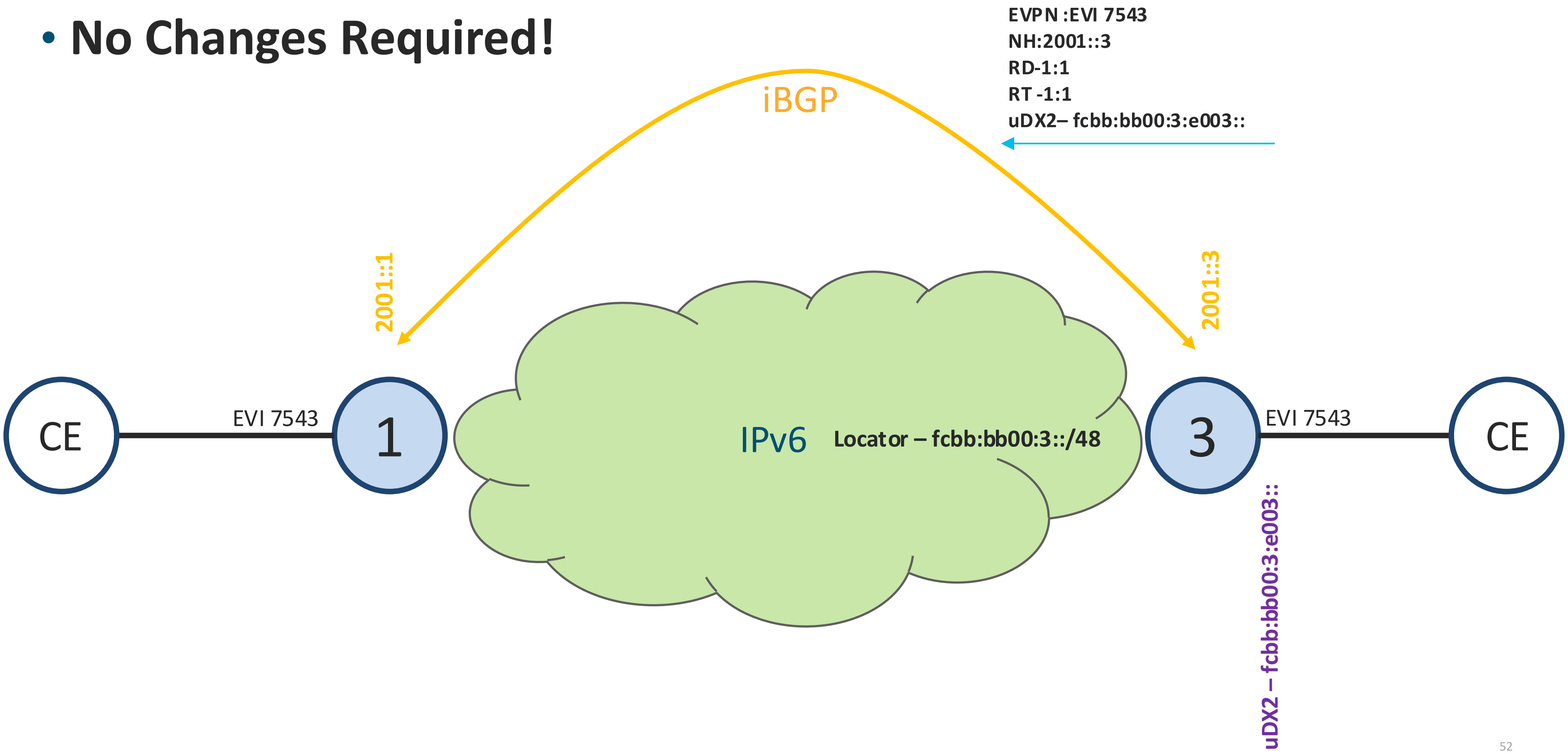
Single DT function is allocated per VRF and AF

This will result in:

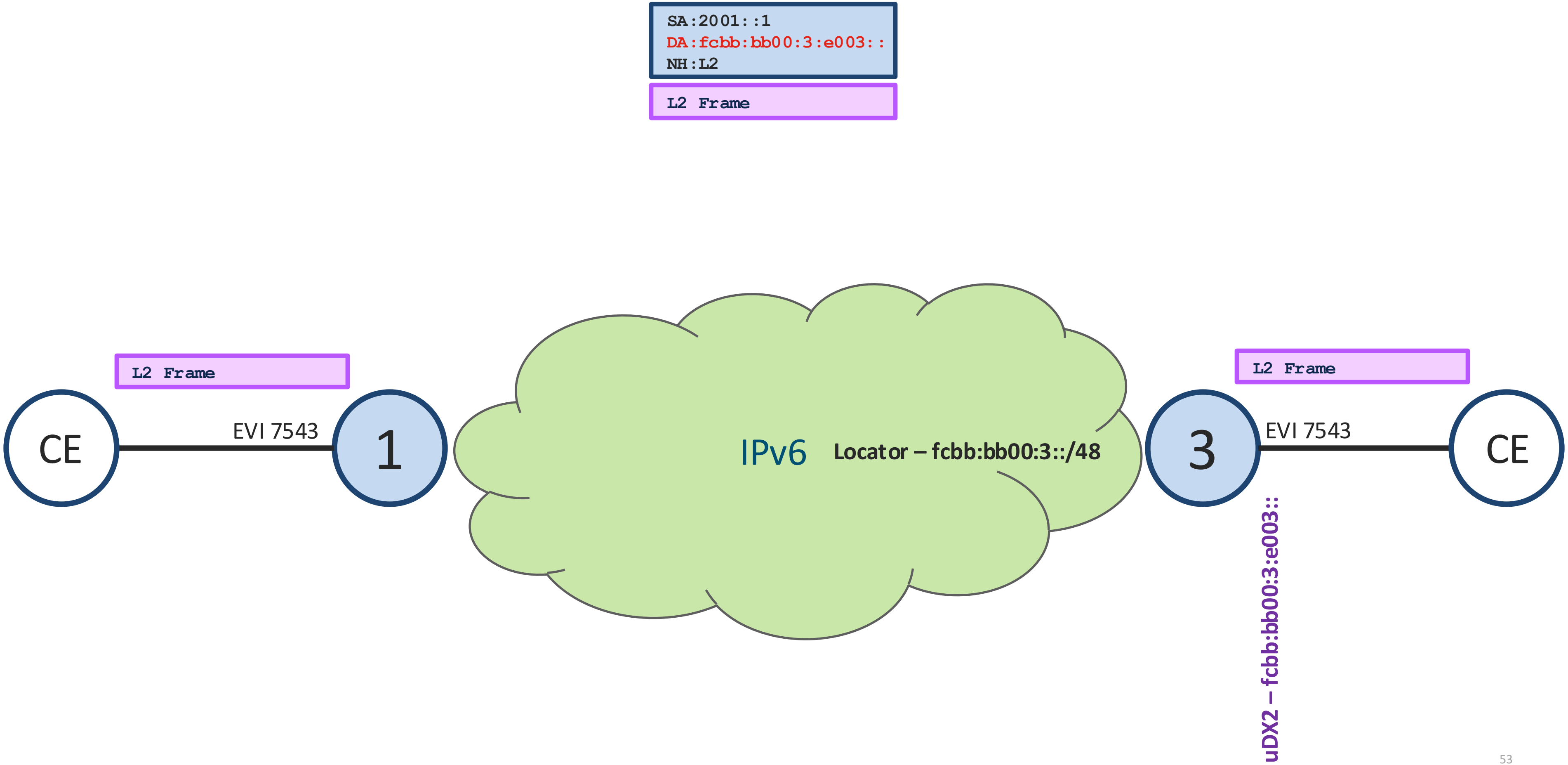
- uDT4 function is allocated
- All prefixes are advertised with uDT4 function

EVPN

- **No Changes Required!**



EVPN Dataplane



SA:2001::1
DA:fcbb:bb00:3:e003::
NH:L2

L2 Frame

SRv6 L2 VPWS

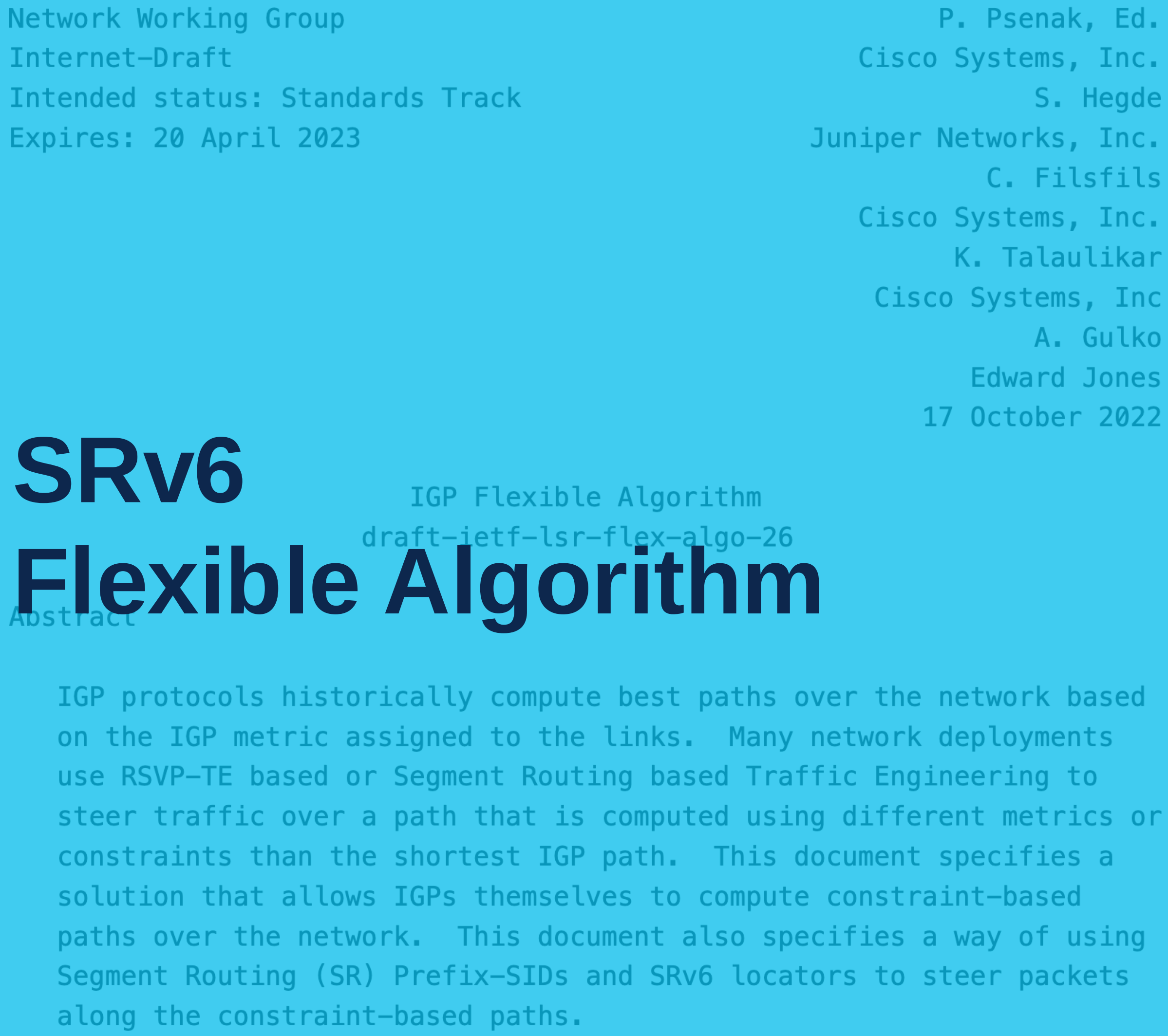
```
interface TenGigE0/0/0/0.7543 l2transport
 encapsulation dot1q 7543
 rewrite ingress tag pop 1 symmetric

l2vpn
 xconnect group P2P
  p2p 13-14
   interface TenGigE0/0/0/0.7543
   neighbor evpn evi 7543 service 7543
segment-routing srv6

evpn
 evi 7543 segment-routing srv6
  locator MAIN
 segment-routing srv6
 !
```

This will result in:

- uDX2 function is allocated per EVI
- EVI is advertised with uDX2 function

Name of the Locator

SRv6
Flexible Algorithm



Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 20 April 2023

P. Psenak, Ed.
Cisco Systems, Inc.
S. Hegde
Juniper Networks, Inc.
C. Filsfils
Cisco Systems, Inc.
K. Talaulikar
Cisco Systems, Inc
A. Gulko
Edward Jones

17 October 2022

IGP Flexible Algorithm
draft-ietf-lsr-flex-algo-26

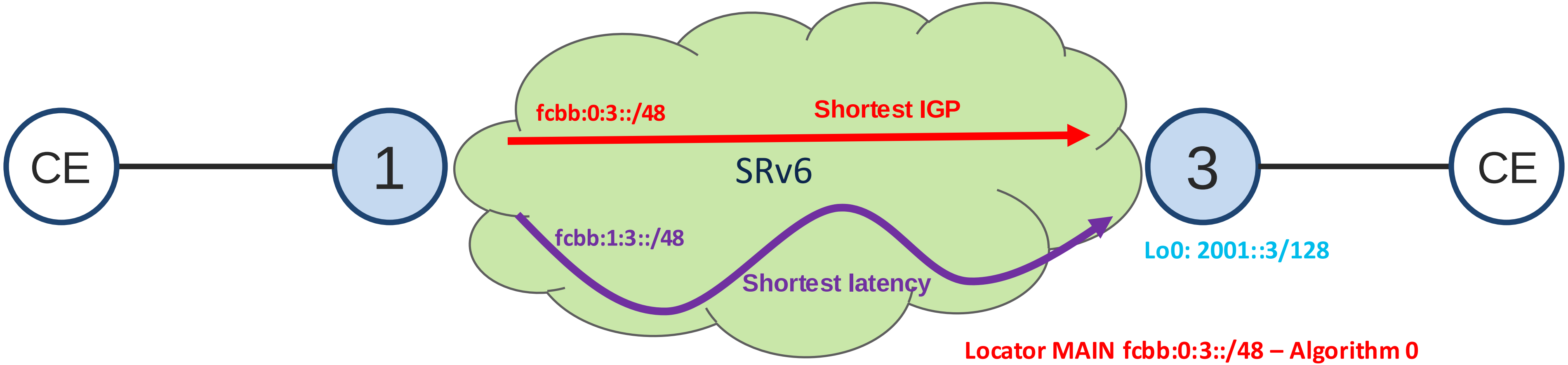
Abstract

IGP protocols historically compute best paths over the network based on the IGP metric assigned to the links. Many network deployments use RSVP-TE based or Segment Routing based Traffic Engineering to steer traffic over a path that is computed using different metrics or constraints than the shortest IGP path. This document specifies a solution that allows IGPs themselves to compute constraint-based paths over the network. This document also specifies a way of using Segment Routing (SR) Prefix-SIDs and SRv6 locators to steer packets along the constraint-based paths.

Flexible Algorithm

- We call “Flex-Algo”
  - The algorithm is defined by the operator, on a per-deployment basis
- Flex-Algo K is defined as
  - The minimization of a specified metric: IGP, delay, …
  - The exclusion of certain link properties: link-affinity, SRLG, …
- Example
  - Operator1 defines Flex-Algo 128 as “minimize IGP metric and avoid link-affinity “green”
  - Operator2 defines Flex-Algo 128 as “minimize delay metric and avoid link-affinity “blue”

SRv6 Flex Algo IGP

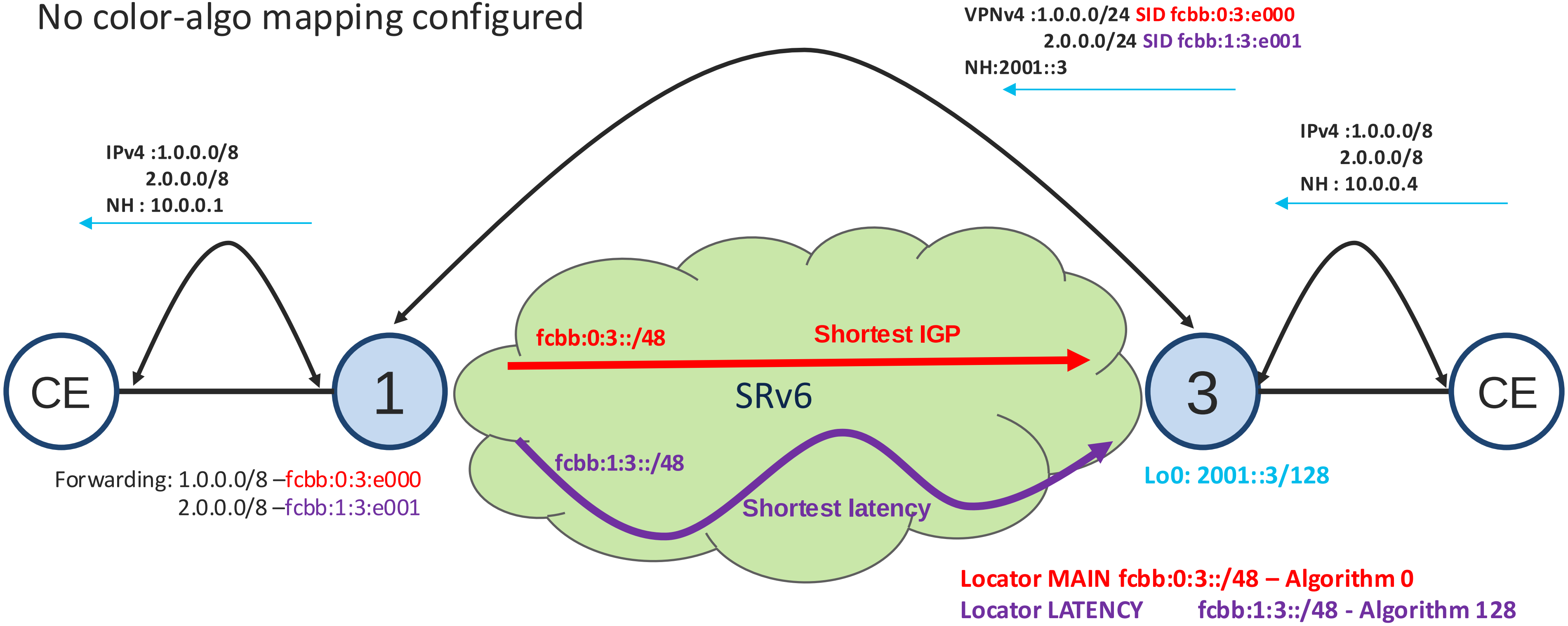


Locator MAIN fcbb:0:3::/48 – Algorithm 0
Locator LATENCY fcbb:1:3::/48 - Algorithm 128

BGP – SRv6

No color advertised

No color-algo mapping configured



VPNv4 :1.0.0.0/24 SID fcbb:0:3:e000
2.0.0.0/24 SID fcbb:1:3:e001
NH:2001::3

IPv4 :1.0.0.0/8
2.0.0.0/8
NH : 10.0.0.1

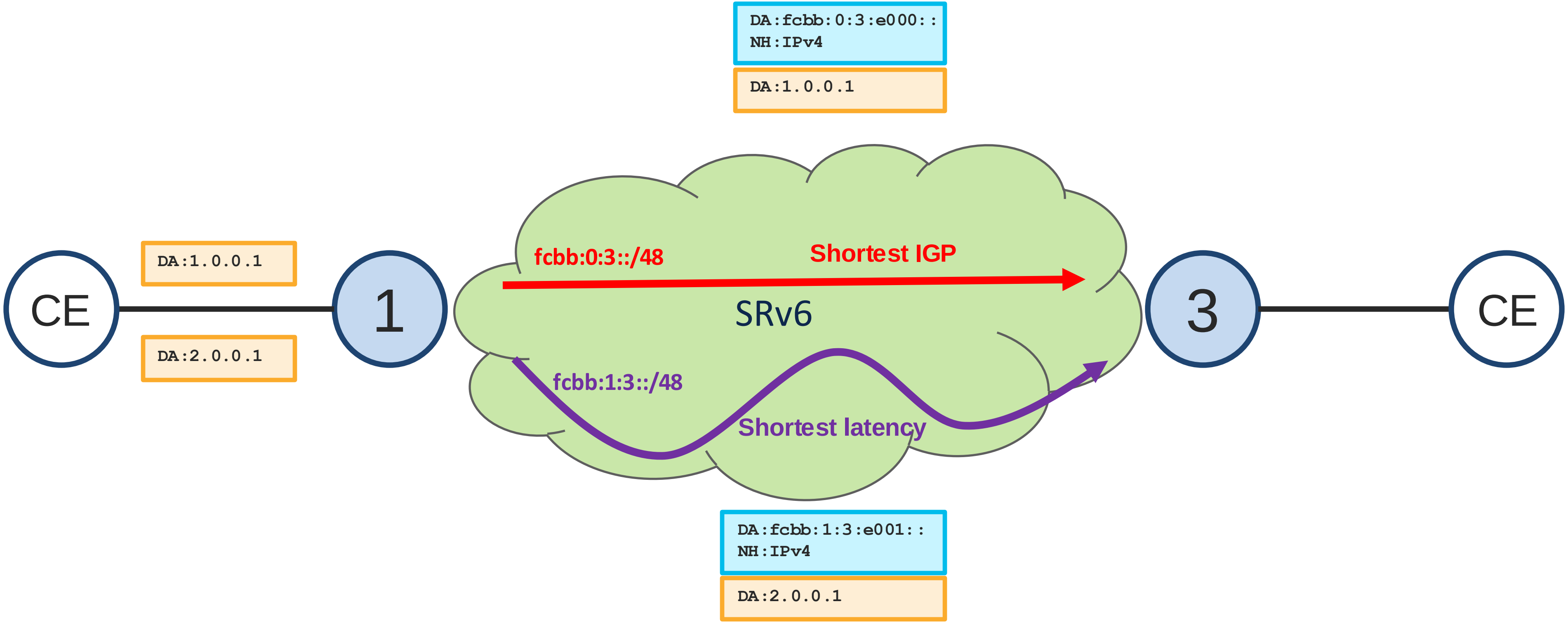
IPv4 :1.0.0.0/8
2.0.0.0/8
NH : 10.0.0.4

Forwarding: 1.0.0.0/8 –fcbb:0:3:e000
2.0.0.0/8 –fcbb:1:3:e001

Lo0: 2001::3/128

Locator MAIN fcbb:0:3::/48 – Algorithm 0
Locator LATENCY fcbb:1:3::/48 - Algorithm 128

SRv6 DATAPLANE



SRv6 Flex Algo -IGP

```
segment-routing
 srv6
  locators
   locator LATENCY
    micro-segment behavior unode psp-usd
    prefix fcbb:bb01:1::/48
    algorithm 128
router isis 1
 flex-algo 128
  metric-type delay
  advertise-definition
 address-family ipv6 unicast
  segment-routing srv6
   locator LATENCY
```

New Locator Name

Locator Prefix (Different)

Flex Algo number 128-255

Definition of specific Flex Algo
Latency metric for 128

This Router will advertise
FA definition within the domain

This will result in:

- Locator is advertised +FA definition
- uN function is advertised - for FA
- uA for each ISIS interface is allocated and advertised for FA

SRv6 L3 VPN Flex Algo

```
router bgp 1
 address-family vpnv4 unicast
 vrf LowLatency
  rd 1:2
  address-family ipv4 unicast
   segment-routing srv6
    locator LATENCY
    alloc mode per-vrf
```

Name of the Locator

Single DT function is allocated per VRF and AF

This will result in:

- uDT4 function is allocated from LATENCY locator
- All prefixes in VRF are advertised with uDT4 function

SRv6 L3 VPN Multiple Algorithms in VRF

```
 route-policy MIX
   if destination in (1.1.1.1/32) then
     set srv6-alloc-mode per-vrf locator LATENCY
   else
     set srv6-alloc-mode per-vrf locator MAIN
   endif
 end-policy
router bgp 1
 vrf Both
  address-family ipv4 unicast
    segment-routing srv6
      alloc mode route-policy MIX
```

For prefix 1.1.1.1 we will allocate uDT from LATENCY

For all others uDT from MAIN

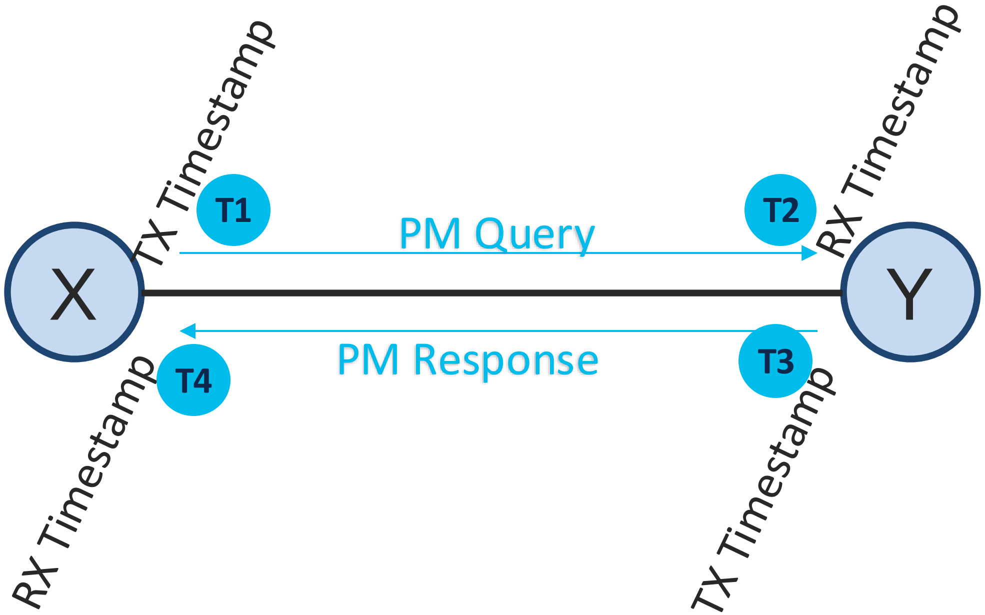
Route-Policy application

Latency Configuration

```
performance-measurement
 interface Gig0/0/0/0
  delay-measurement
   advertise-delay 7543
```

This will set latency of the link to 7543 **microseconds**

Performance Measurement



- TWAMP Ligth Protocol
- HW Level Timestamping
- ns precision!
- Link Latency Calculation:
  - One WAY Measurement

$$Latency = T2 - T1$$

  - Two Way Measurement

$$Latency = \frac{(T4 - T1) - (T3 - T2)}{2}$$

PM Configuration

```
performance-measurement
 interface Gig0/0/0/0
  delay-measurement
```

This will:

- Start PM probes on interface
- Provide Dynamic measurement values to IGP
- Both ends must be PM capable (provide HW based timestamping)

SRv6 Addressing

Separation between SIDs and addresses

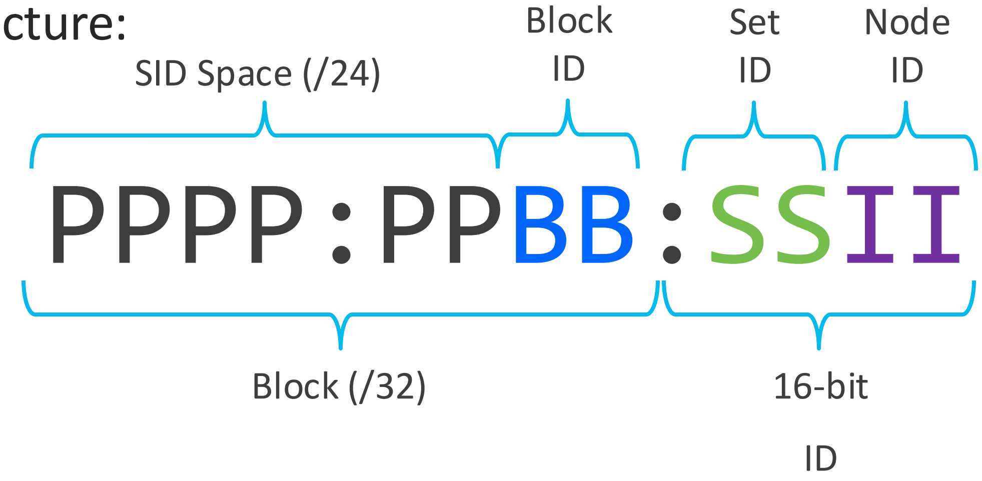
- Infrastructure addressing and SRv6 SID allocation belong to two different planes and are different
  - Infrastructure IP addresses (e.g., link interfaces, loopbacks) are allocated on the management plane
  - SRv6 SIDs are allocated on the service plane
- SRv6 SIDs are assigned to a node independently from the IP addressing of that node
- Even if they are both represented as IPv6 addresses, infrastructure addresses and SIDs cannot be merged and should be allocated off different blocks.

An existing IPv6 address plan is not a constraint for a future SRv6 SID allocation plan.

Terminology – uSID F3216

- uSID F3216: uSID format with
  - uSID Block size: 32 bits
  - ID size: 16 bits

- uSID F3216 structure:



SRv6 Space allocation recommendation

- Private range allocation ✓
  - **Recommended allocation**
  - Use /24 sub-range from ULA `FC00::/8` space
  - `FCBB:BB00::/24`, with B indicating a nibble value picked by operator
- Public range allocation
  - **Supported, not advised** ⚠
  - From allocated public GUA range
- SRv6 range allocation
  - Use /24 sub-range from `5F00::/16` ✓

uSID Block per slice (Flex Algo) if possible

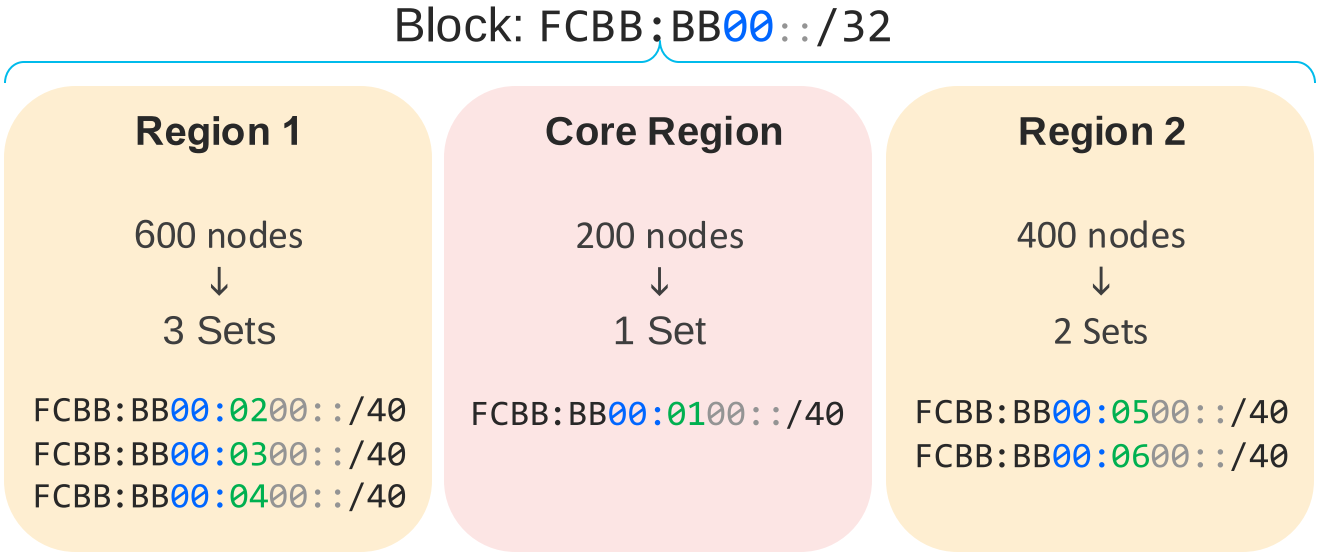
- 256 Blocks are available in the SRv6 Space:

  `FCBB:BBTT::/32`, with TT = slice ID

  - Multiple Blocks can be concurrently used on a node
  - 63 Blocks available on DNX1 platforms (TT = 00 to 3E)

- We assume 2 slices (Blocks), e.g.:
  - `FCBB:BB00::/32` Low-cost slice (algo 0) ← focus, other Blocks are similar
  - `FCBB:BB01::/32` Low-delay slice (algo 128)

Set Allocation Example



Block: FCBB:BB00::/32

Region 1

600 nodes
↓
3 Sets

```
FCBB:BB00:0200::/40
FCBB:BB00:0300::/40
FCBB:BB00:0400::/40
```

Core Region

200 nodes
↓
1 Set

```
FCBB:BB00:0100::/40
```

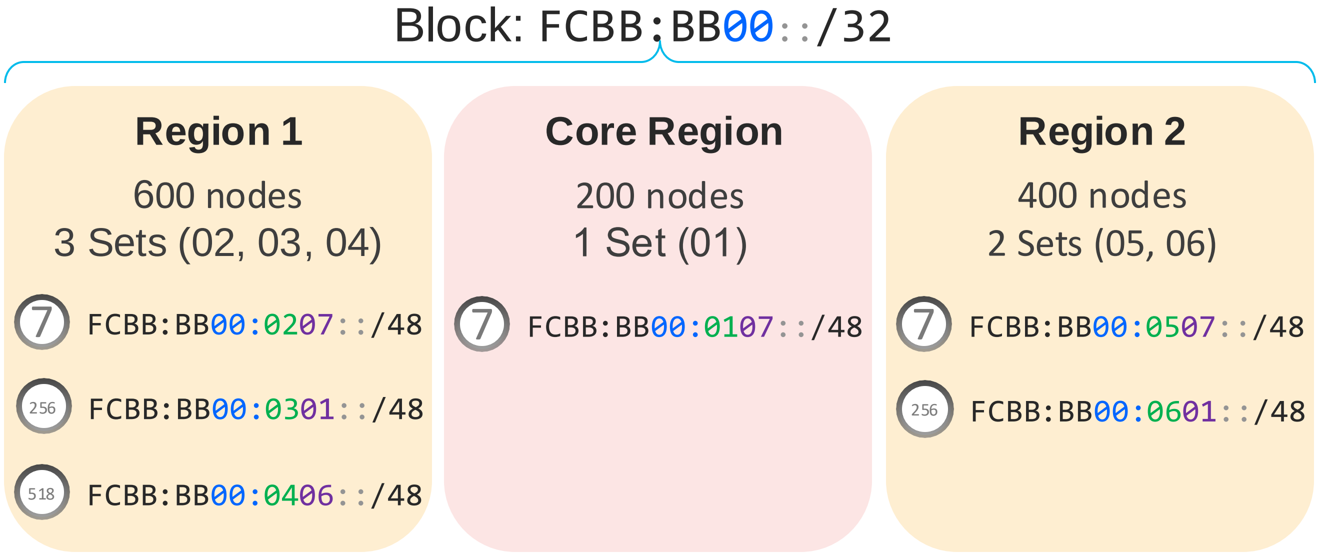
Region 2

400 nodes
↓
2 Sets

```
FCBB:BB00:0500::/40
FCBB:BB00:0600::/40
```

- If a region outgrows its allocated Sets, then allocate more Sets to this region

uSID Allocation Example



Block: FCBB:BB00::/32

Region 1

600 nodes
3 Sets (02, 03, 04)

7 FCBB:BB00:0207::/48

256 FCBB:BB00:0301::/48

518 FCBB:BB00:0406::/48

Core Region

200 nodes
1 Set (01)

7 FCBB:BB00:0107::/48

Region 2

400 nodes
2 Sets (05, 06)

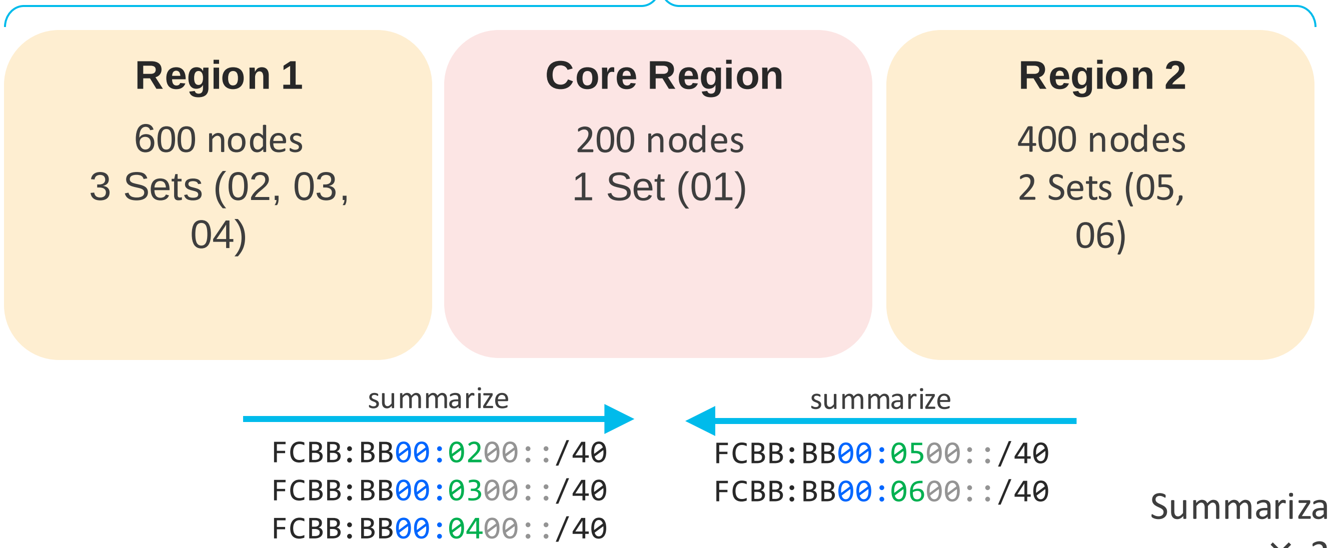
7 FCBB:BB00:0507::/48

256 FCBB:BB00:0601::/48

- Remaining unallocated uSIDs in Sets are for future growth

Summarization

Block: FCBB:BB00::/32



Region 1

600 nodes
3 Sets (02, 03, 04)

Core Region

200 nodes
1 Set (01)

Region 2

400 nodes
2 Sets (05, 06)

summarize

FCBB:BB00:0200::/40
FCBB:BB00:0300::/40
FCBB:BB00:0400::/40

summarize

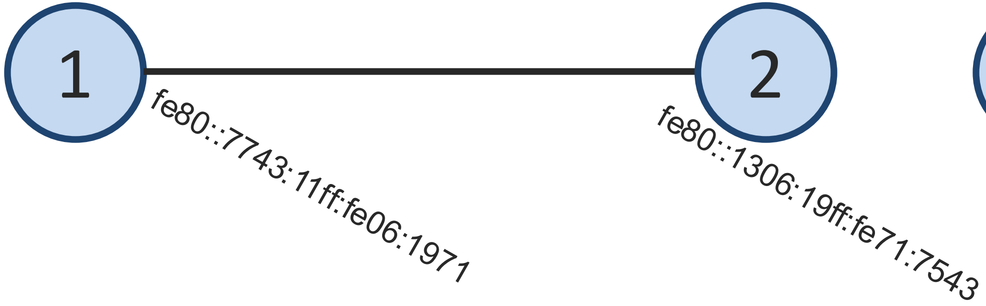
FCBB:BB00:0500::/40
FCBB:BB00:0600::/40

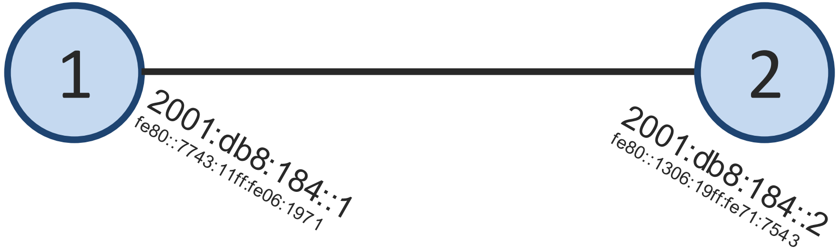
Summarization gain:
× 256

Interface Addressing

- Link Local Only
  - no config
  - no address planning
  - new device insertion
  - smaller LSP
  - Interfaces not remotely reachable
  - uA can be used instead for ping

- Global (ULA/GUA)
  - Must be configured
  - Must be planned, maintained
  - Interfaces remotely reachable





Loopback Addressing

- Independent

3

Locator:	fcbb:bb00:3::/48
Lo0:	2001:db8::1/128

- From Locator

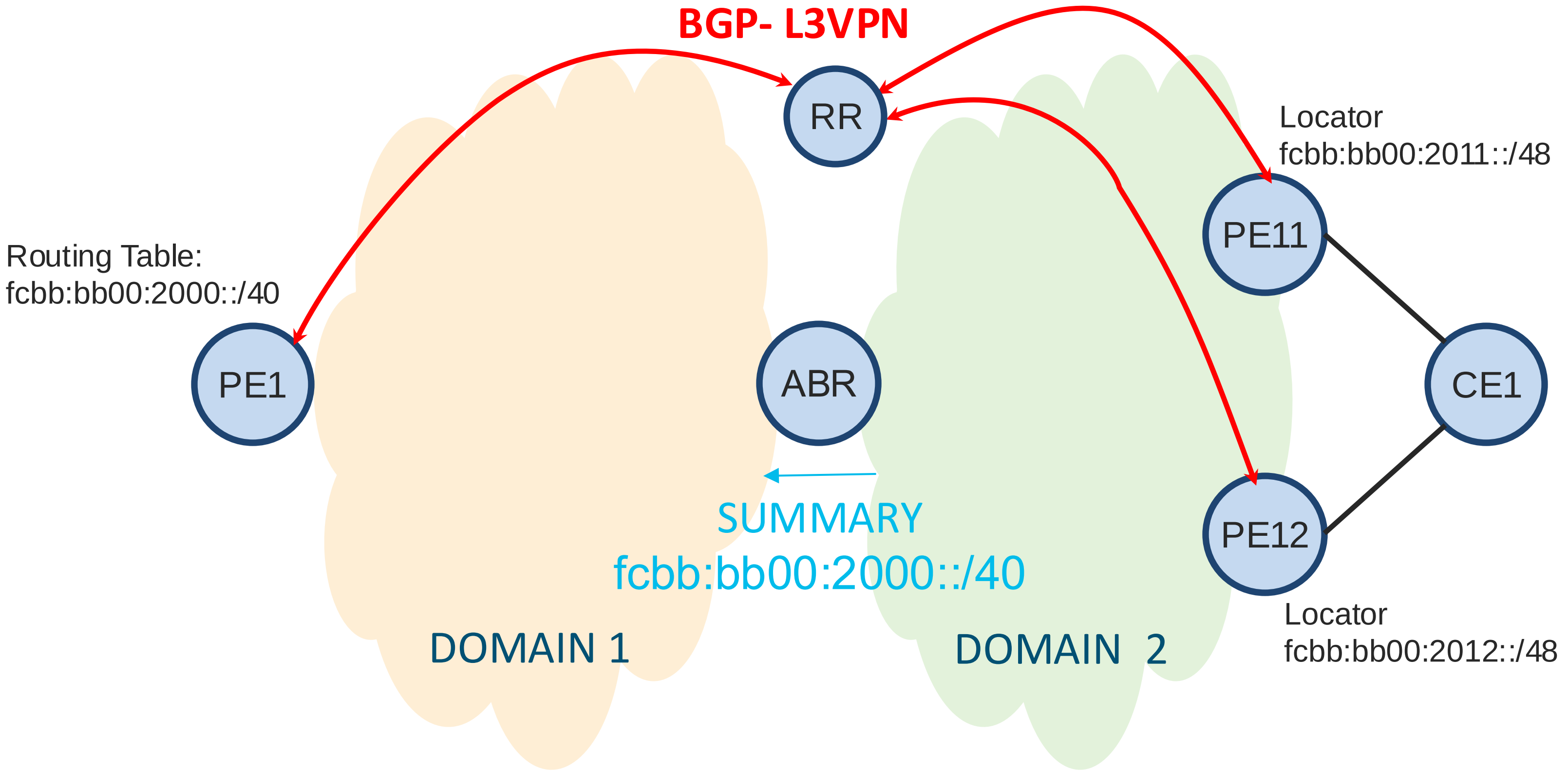
3

Locator:	fcbb:bb00:3::/48
Lo0:	fcbb:bb00:3::1/128

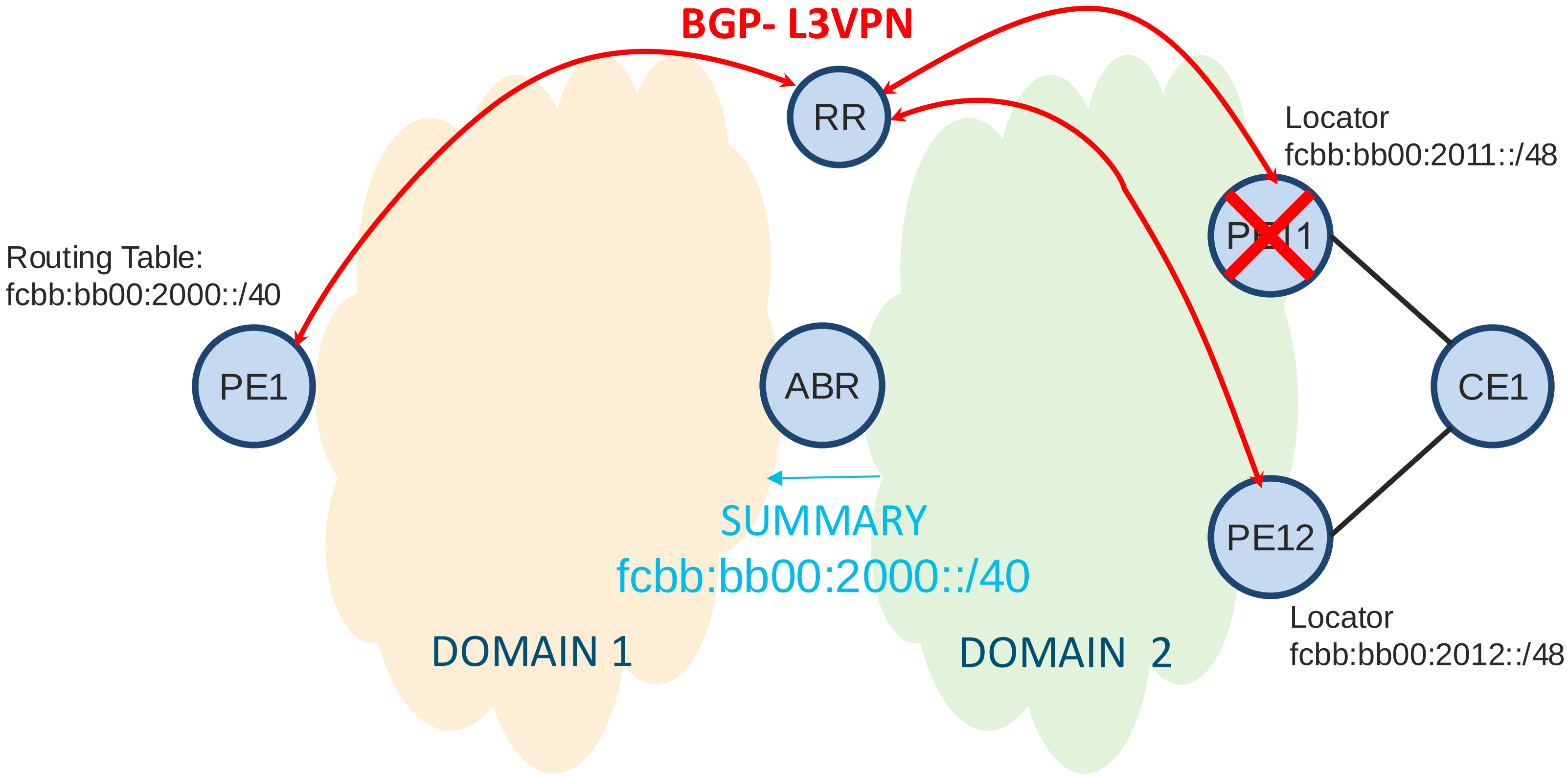
- Smaller LSP
- No loopback addressing planning
- Summarization!!

*Note: only last nibble can be altered!

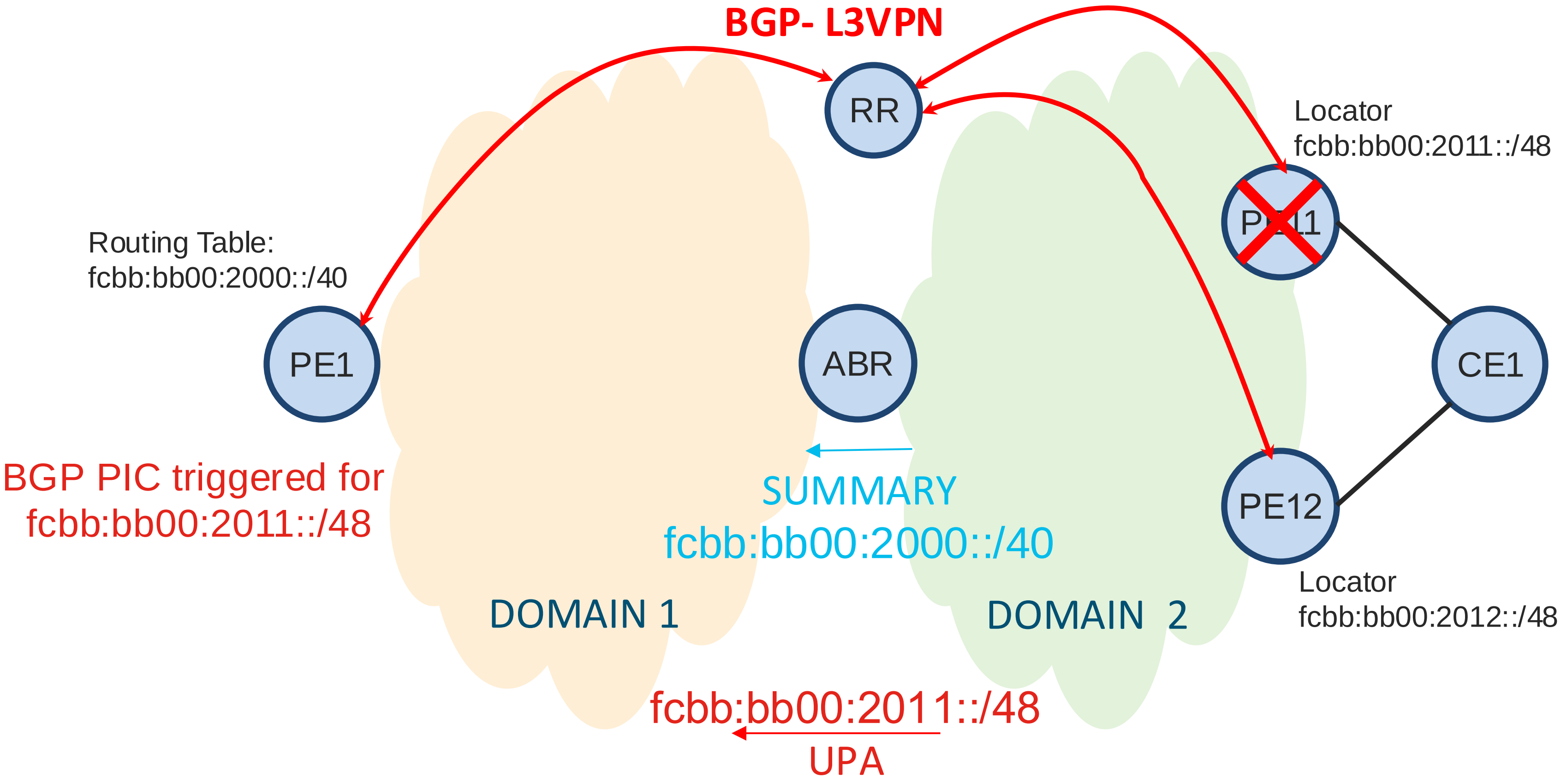
Unreachable Prefix Announcement



Unreachable Prefix Announcement



Unreachable Prefix Announcement



Routing Table:
fcbb:bb00:2000::/40

Locator
fcbb:bb00:2011::/48

Locator
fcbb:bb00:2012::/48

BGP PIC triggered for
fcbb:bb00:2011::/48

SUMMARY
fcbb:bb00:2000::/40

fcbb:bb00:2011::/48
UPA

SRv6 Summarization +UPA configuration

ABR:

Summarization for Algorithm 0, per Set /40

```
router isis 1

 address-family ipv6 unicast
  summary-prefix fcbb:bb00:2000::/40 adv-unreachable
  summary-prefix fcbb:bb01:2000::/40 algorithm 128 adv-unreachable
```

UPA

Summarization for Algorithm 128, per Set /40

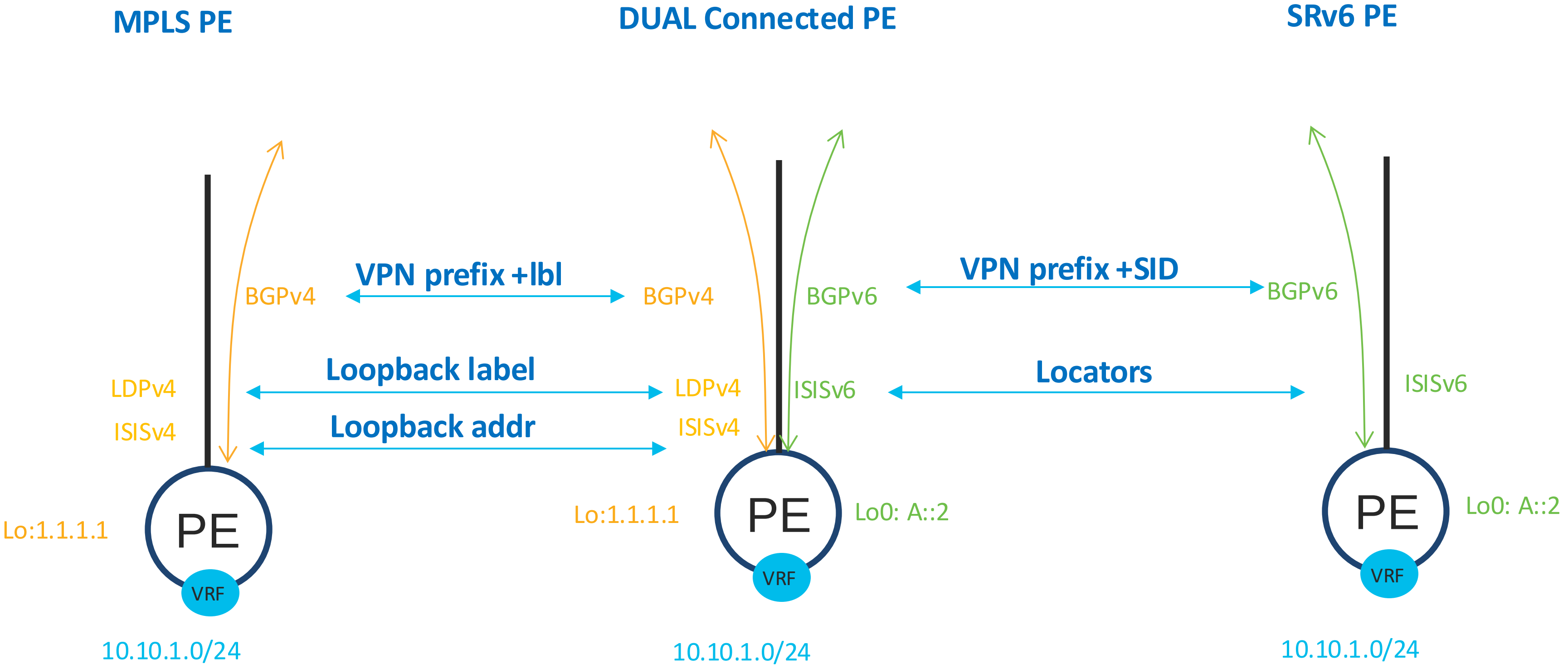
PE:

```
 router isis 1
  address-family ipv6 unicast
   prefix-unreachable
    rx-process-enable
```

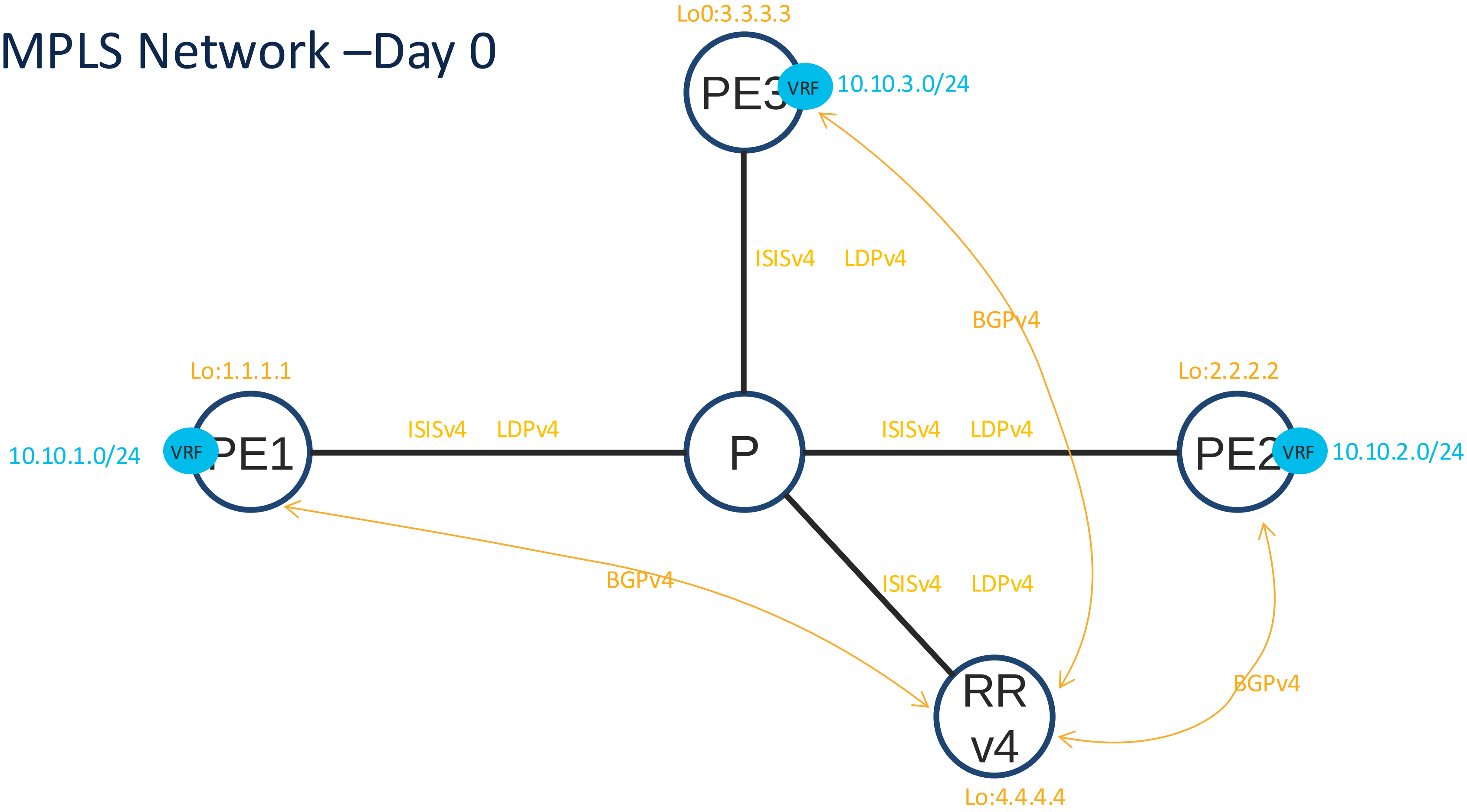
Triggers BGP PIC

SRv6 Migration

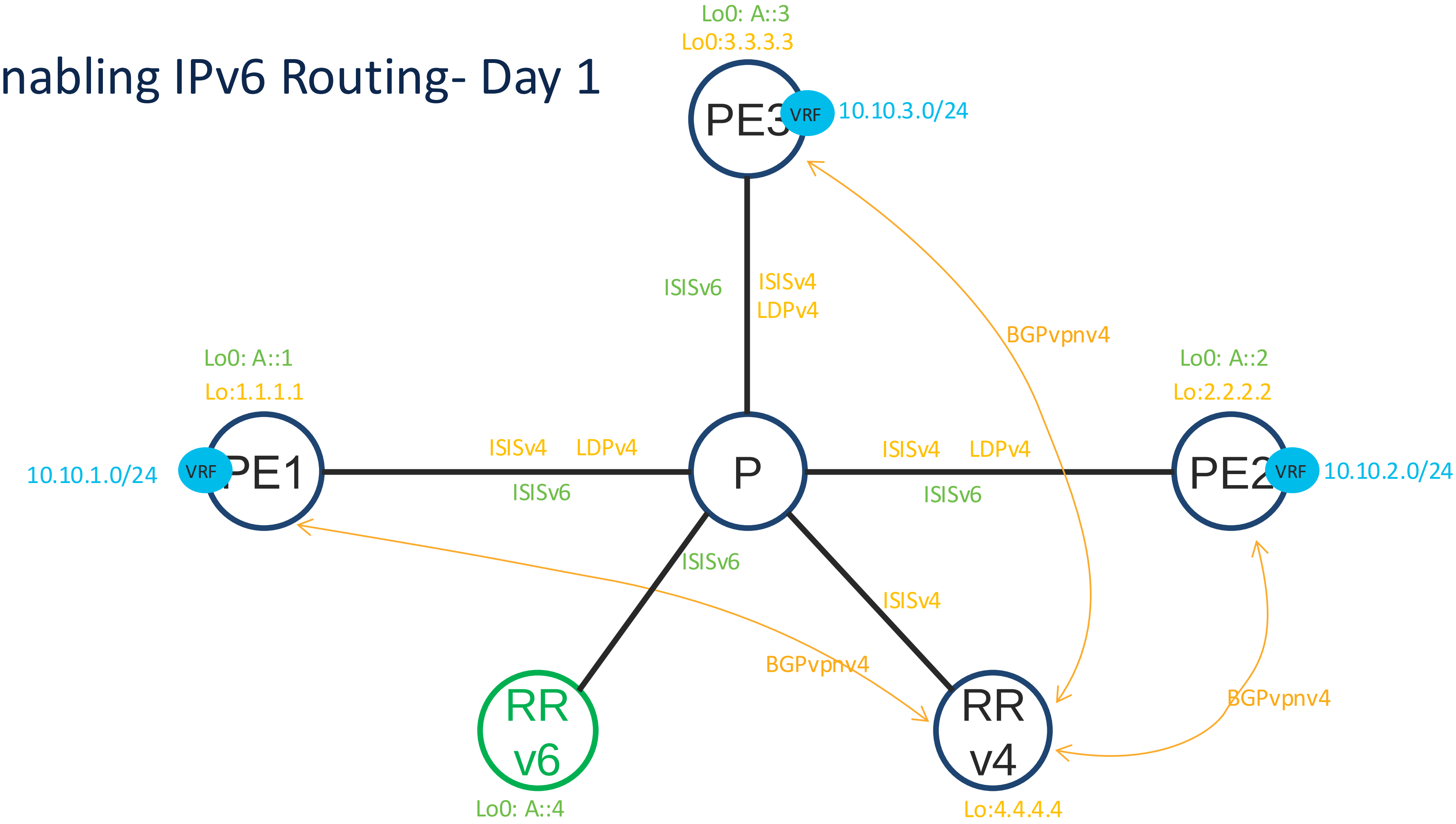
Dual Connected PE



MPLS Network –Day 0



Enabling IPv6 Routing- Day 1



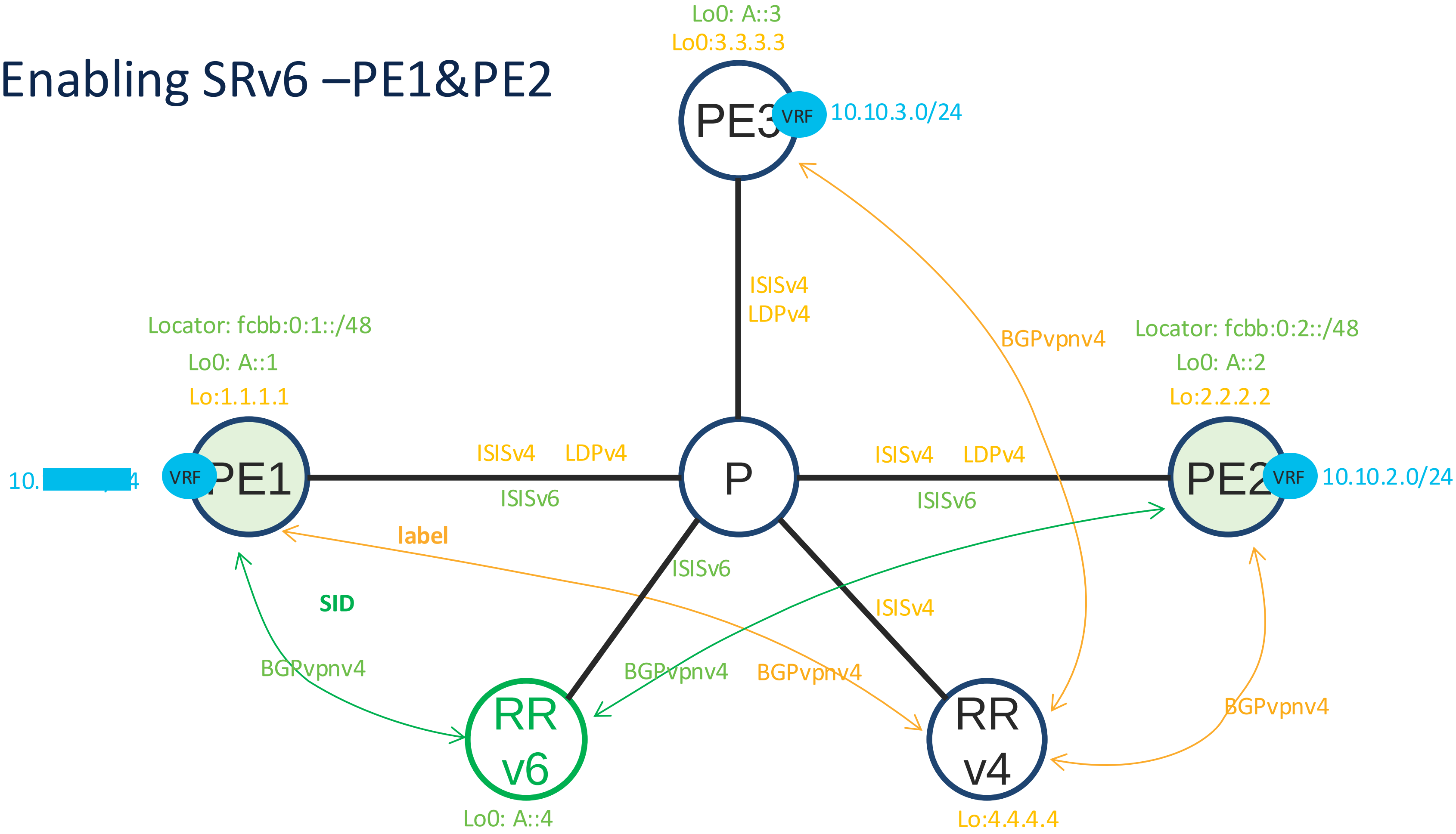
IPv6 Routing Configuration

```
interface Loopback0
 ipv6 address fcbb:bb00:2::1/128 interface
TenGigE 0/0/0/0
  ipv6 enable
router isis 1
 address-family ipv6 unicast
  metric-style wide
  advertise link attributes
  router-id Loopback 0
interface Loopback0
    passive
    address-family ipv6 unicast
  interface TenGigE 0/0/0/1
    circuit-type level-2-only
    point-to-point
    address-family ipv6 unicast
  !
```

← Loopback From Locator

← LL Only on Interface

Enabling SRv6 –PE1&PE2



Everything is BGP best path selection driven!! (ie Local Preference)

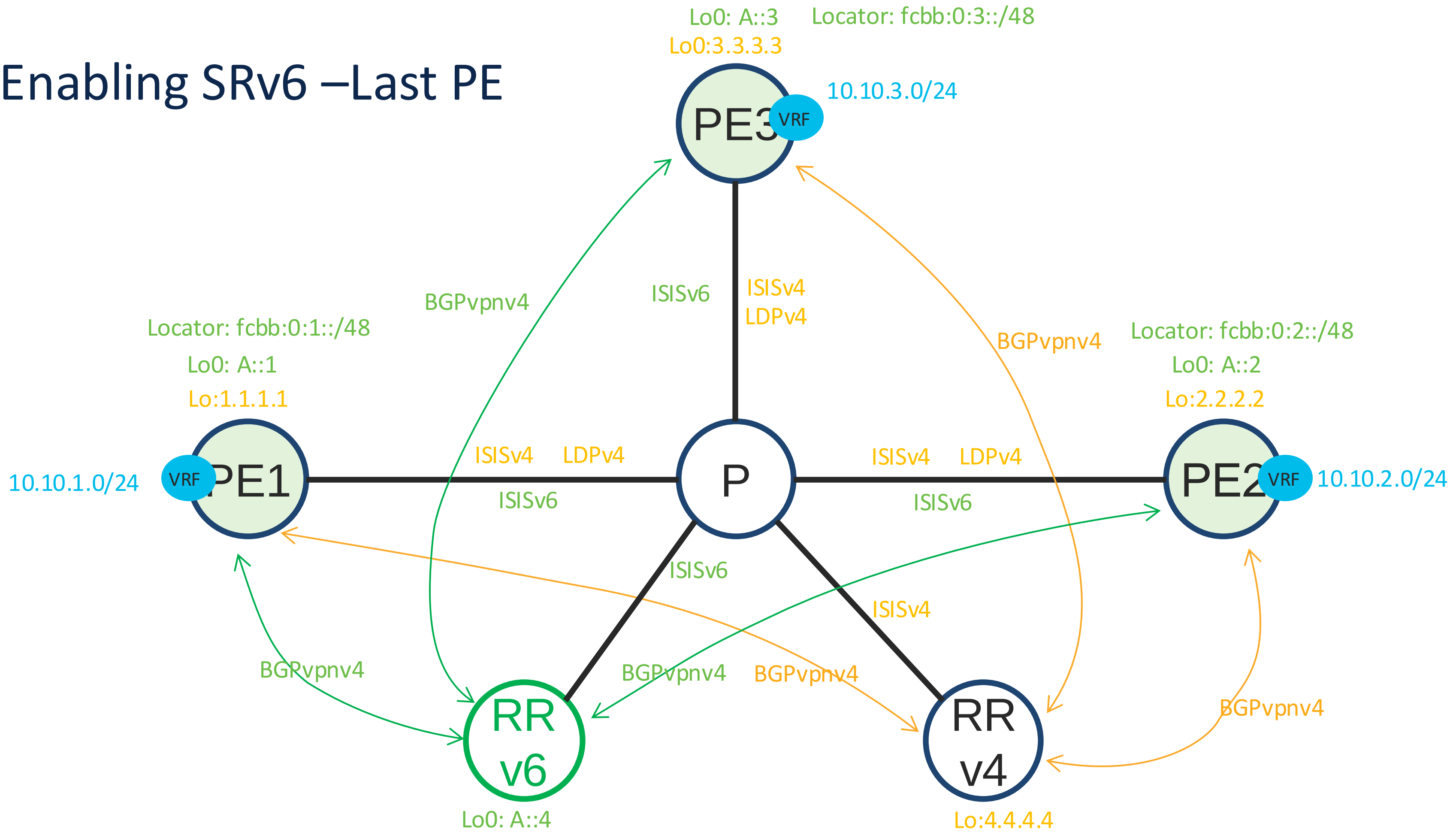
Enable SRv6

```
router isis 1
 address-family ipv6 unicast
  segment-routing srv6
   locator MAIN
```

Reference to Locator MAIN

```
segment-routing
 srv6
  locators
   locator MAIN
    micro-segment behavior unode psp-usd
    prefix fcbb:bb00:2::/48
```

Enabling SRv6 –Last PE



SRv6 Dual PE Configuration

```
router bgp 1
 neighbor A::4
  address-family vpnv4 unicast
   encapsulation-type srv6
   route-policy RRv6 out
 neighbor 4.4.4.4
  address-family vpnv4 unicast
   route-policy RRv4 out
vrf 1
  address-family ipv4 unicast
   mpls alloc enable
   segment-routing srv6
    locator MAIN
    alloc mode per-vrf
```

Via RPL we set specific BGP
attributes to to prefixes
ie Local Preference
towards RRv6 and RRv4

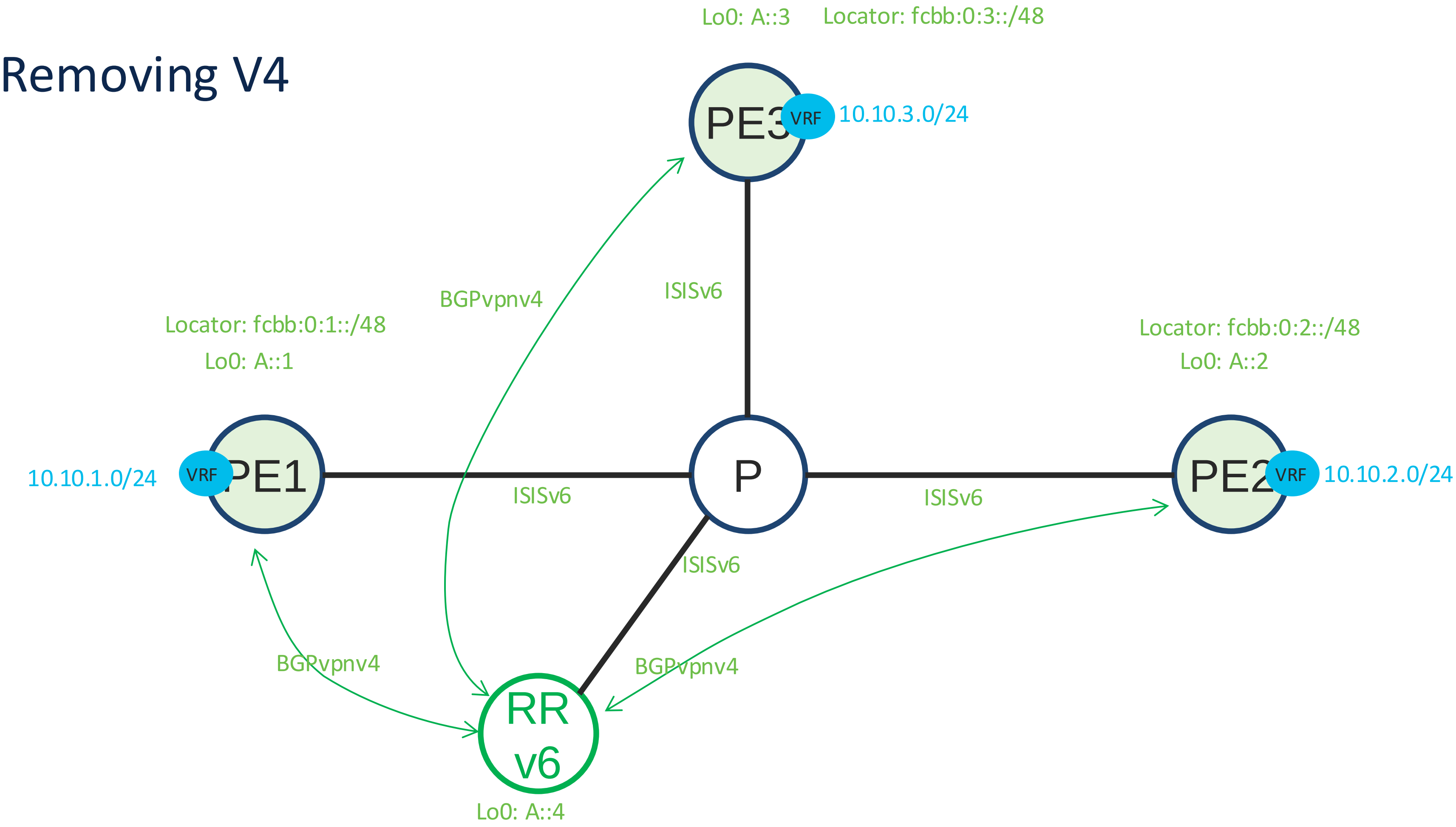
Policy towards v6 RR

Policy towards v4 RR

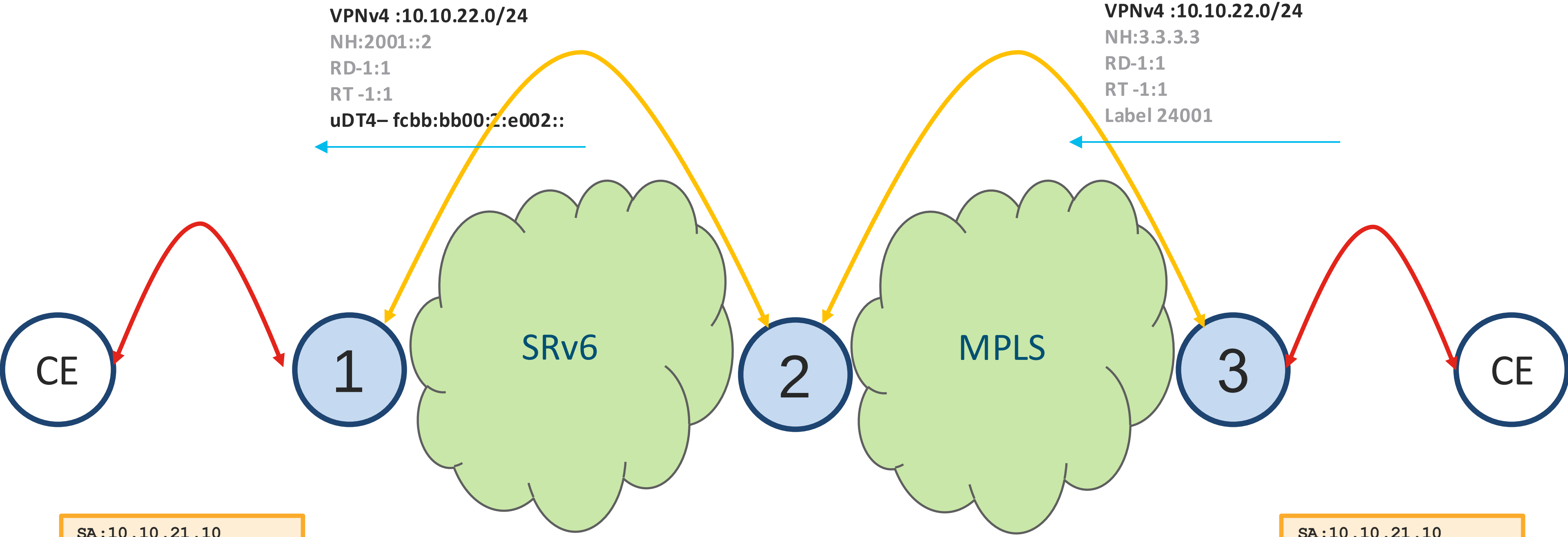
Allocates Labels for all prefixes in VRF

Allocates SIDs for all prefixes in VRF from Locator MAIN

Removing V4


Lo0: A::3
Locator: fcbb:0:3::/48
PE3
VRF
10.10.3.0/24
BGPvpnv4
ISISv6
Locator: fcbb:0:1::/48
Lo0: A::1
Locator: fcbb:0:2::/48
Lo0: A::2
10.10.1.0/24
VRF
PE1
P
PE2
VRF
10.10.2.0/24
ISISv6
ISISv6
ISISv6
BGPvpnv4
BGPvpnv4
RR
v6
Lo0: A::4

L3 VPN SRv6 MPLS Gateway



VPNv4 :10.10.22.0/24
NH:2001::2
RD-1:1
RT -1:1
uDT4– fcbb:bb00:2:e002::

VPNv4 :10.10.22.0/24
NH:3.3.3.3
RD-1:1
RT -1:1
Label 24001

```
SA:10.10.21.10
DA:10.10.22.10
NH:UDP
```

```
UDP Header/Data
```

```
SA:2001::1
DA:fcbb:bb00:2:e002::
NH:IPv4
```

```
SA:10.10.21.10
DA:10.10.22.10
NH:UDP
```

```
UDP Header/Data
```

```
16003
24001
```

```
SA:10.10.21.10
DA:10.10.22.10
NH:UDP
```

```
UDP Header/Data
```

```
SA:10.10.21.10
DA:10.10.22.10
NH:UDP
```

```
UDP Header/Data
```

L3VPN SRv6 to MPLS Gateway Configuration -VRF

```
vrf 1
  address-family ipv4 unicast
   import route-target
    1:1
    2:1 stitching
   export route-target
    1:1
    2:1 stitching
```

RT for MPLS

RT for SRv6

L3VPN SRv6 to MPLS Gateway Configuration -BGP

```
router bgp 1
 neighbor 4::4
  address-family vpnv4 unicast
   encapsulation-type srv6
   import reoriginate stitching-rt
   advertise vpnv4 unicast re-originated
 neighbor 3.3.3.3
  address-family vpnv4 unicast
   import stitching-rt reoriginate
   advertise vpnv4 unicast re-originated stitching-rt

 vrf 1
  address-family ipv4 unicast
   mpls alloc enable
   segment-routing srv6
    locator MAIN
```

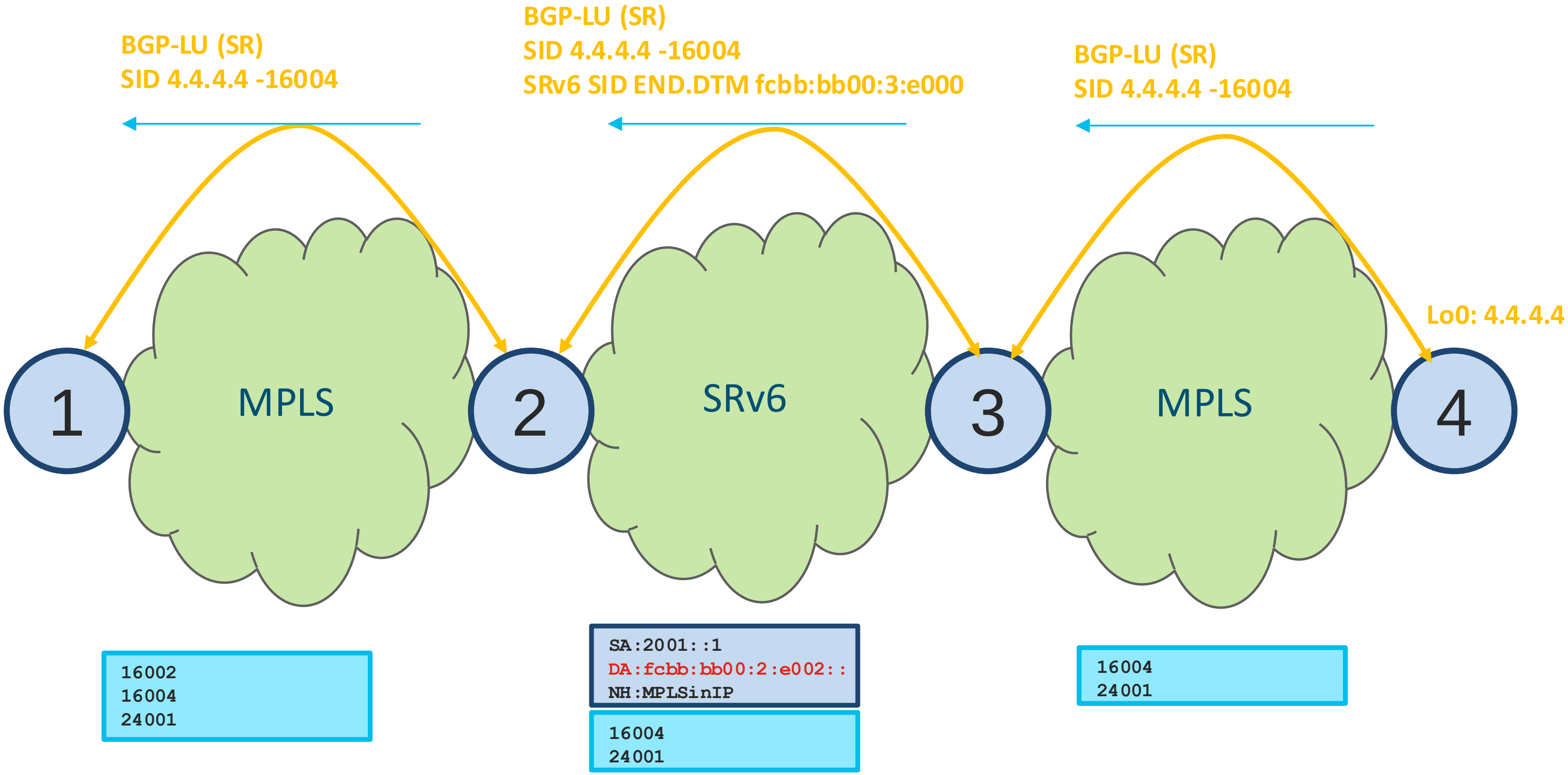
v6 RR

v4 RR

Allocates Labels for all prefixes in VRF

Allocates SIDs for all prefixes in VRF from Locator MAIN

Mo6 – (BGP-LU+SID)



BGP-LU (SR)
SID 4.4.4.4 -16004

BGP-LU (SR)
SID 4.4.4.4 -16004
SRv6 SID END.DTM fcbb:bb00:3:e000

BGP-LU (SR)
SID 4.4.4.4 -16004

Lo0: 4.4.4.4

```
16002
16004
24001
```

```
SA:2001::1
DA:fcbb:bb00:2:e002::
NH:MPLSinIP
```

```
16004
24001
```

```
16004
24001
```

Mo6 ABR Configuration

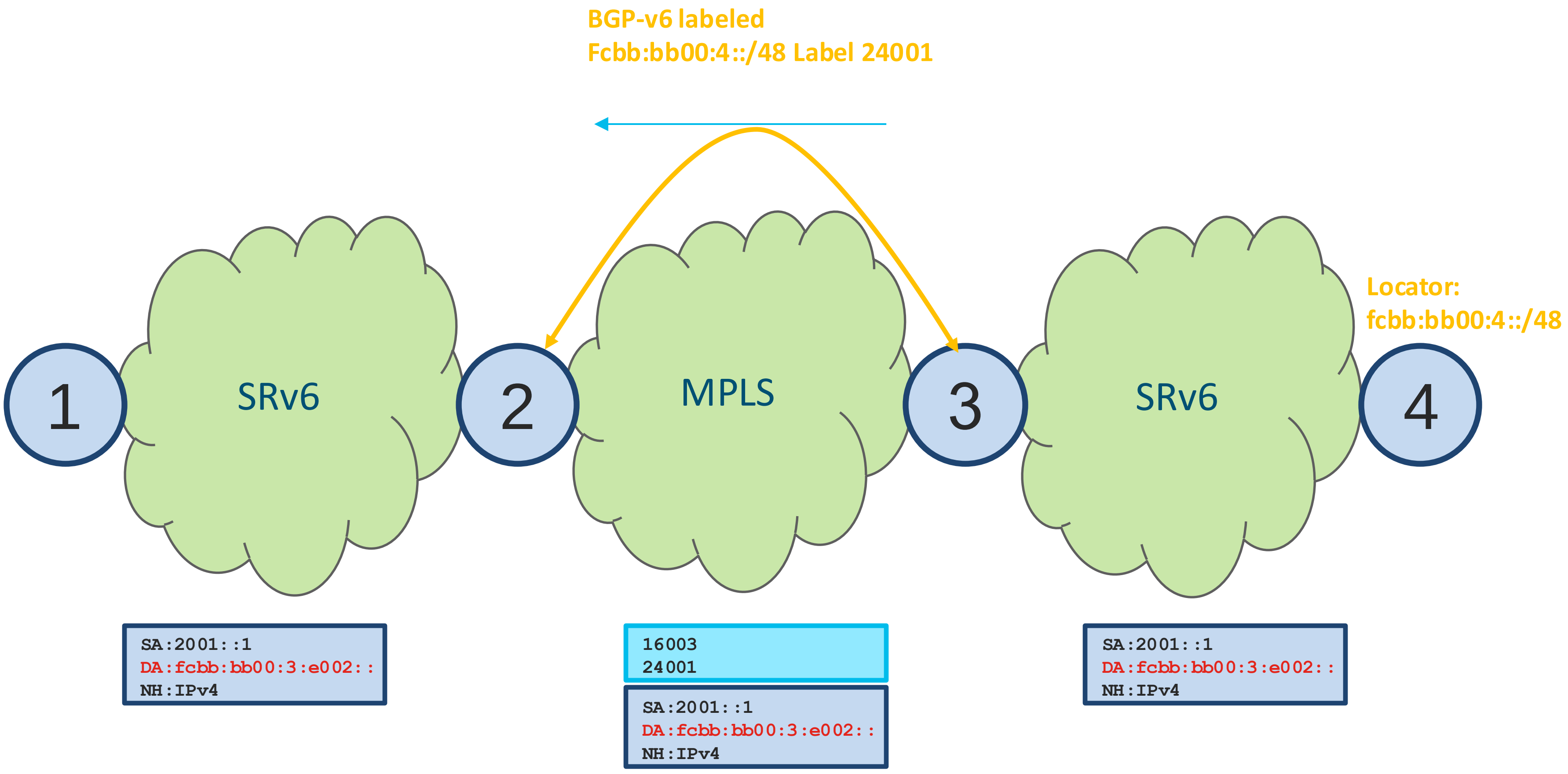
```
router bgp 1
 address-family ipv4 unicast
  segment-routing srv6
   locator MAIN
  labeled-unicast
!
 neighbor 1::1
  address-family ipv4 labeled-unicast
   encapsulation-type srv6
```

Allocates SRv6 SID

Allocate Label

Advertises SID in LU AF

6PE(6VPE)



BGP-v6 labeled
Fcbb:bb00:4::/48 Label 24001

Locator:
fcbb:bb00:4::/48

1

SRv6

2

MPLS

3

SRv6

4

```
SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4
```

```
16003
24001
```

```
SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4
```

```
SA:2001::1
DA:fcbb:bb00:3:e002::
NH:IPv4
```

SRv6 uSID Conclusion

SRv6 is Fully Standardized

Architecture

- SR Architecture – **RFC 8402**
- SRTE Policy Architecture – **RFC 9256**
- Compressed SRv6 Segment List – **WG Draft**

Data Plane

- SRv6 Network Programming – **RFC 8986**
- IPv6 SR header – **RFC 8754**

Control Plane

- SRv6 BGP Services – **RFC 9252**
- SRv6 ISIS – **RFC 9352**
- SR Flex-Algo – **RFC 9350**

Operation & Management

- SRv6 OAM – **RFC 9259**
- Performance Management – **RFC 5357**

Strong Cisco Commitment and Leadership

Editor of	96% IETF RFCs
Co-author of	100% IETF RFCs

Rich SRv6 uSID Ecosystem

Network Equipment Manufacturers

CISCO HUAWEI NOKIA ciena JUNIPER NETWORKS

ZTE ARRCUS NETWORK DIFFERENT ARISTA H3C

Merchant Silicon

BROADCOM MARVELL

BAREFOOT NETWORKS intel

Open-Source Applications

Pyroute2 IPTABLES SERA GoBGP exa BGP eBPF cilium

TCPDUMP WIRESHARK SNORT GoBMP

Open-Source Networking Stacks

.io P4

OPEN Compute Project SONiC FRRouting

SR v6 Segment Routing

Smart NIC

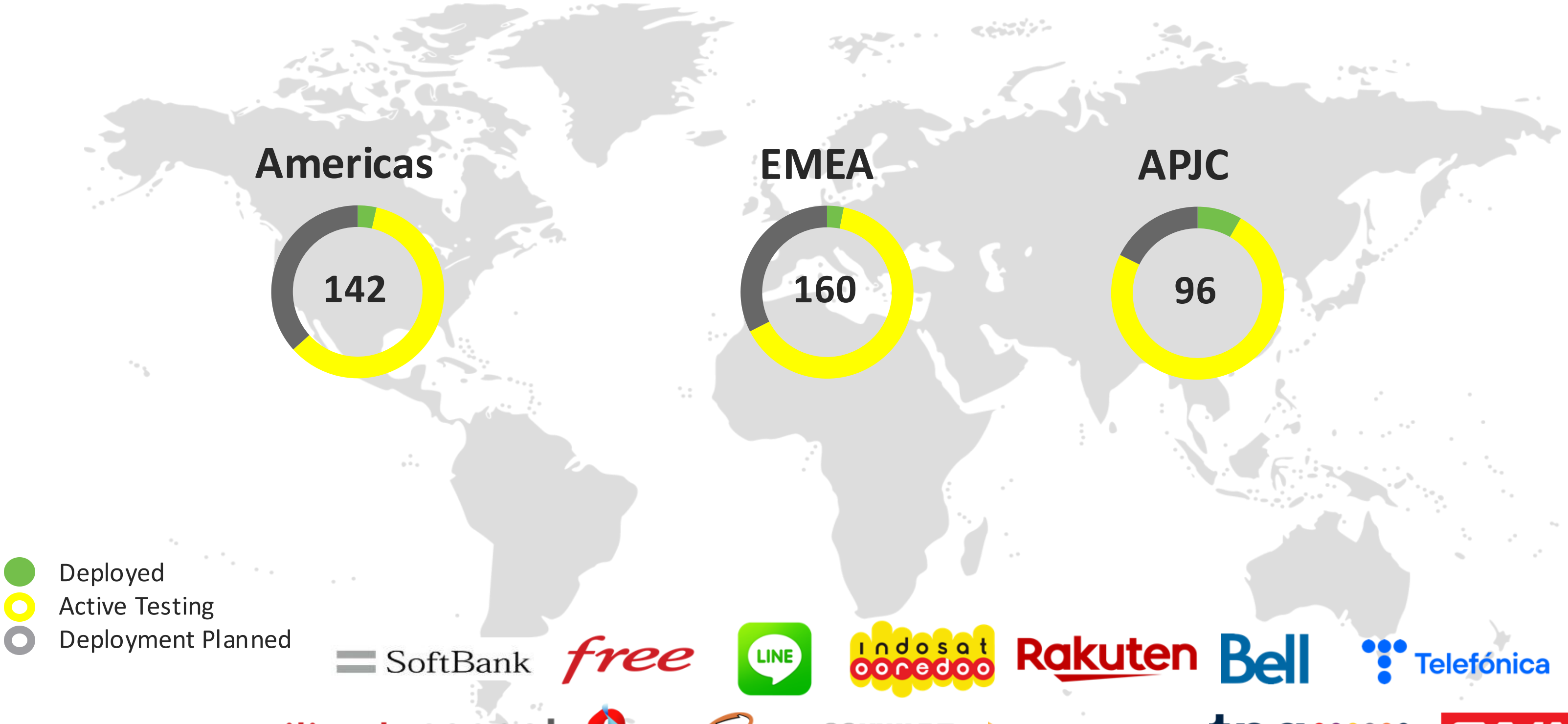
intel Mellanox TECHNOLOGIES

Partners

kaloom TREND MICRO ENEA Qosmos Division

NoviFlow ixia | A Keysight Business spirent Promise. Assured.

SRv6 … at Record-Speed



SoftBank free LINE indosat ooredoo Rakuten Bell Telefónica T Mobile

iliad cesnet swisscom Alibaba.com SCHWARZ ITKG XL axiata tpg TELECOM SYMA

Simplicity Always Prevails

~~LDP~~

~~RSVP-TE~~

~~BGP 3107~~

~~MPLS~~

~~UDP/VxLAN~~

~~NSH~~

Furthermore, with more scaleand functionality cisco

The bridge to possible