cisco Live! Let’s go
SRv6 for Next-Generation Transport Networks

Brent Foster
Principal Architect
Communications Service Provider Sales
BRKMPL-2205
Cisco Webex App

Questions?
Use Cisco Webex App to chat with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 9, 2023.

https://ciscolive.ciscoevents.com/ciscolivebot/#BRKMLE-2205
It’s time to put MPLS in the rear-view mirror!
Agenda

• Introduction
• Comparing MPLS and SRv6 (plus SRv6 101)
• SRv6 Architecture Advantages
• SRv6 Test-Cases
• Cisco Platform Support for SRv6
• Conclusion
Introduction
Segment Routing 101

An IP source-routing architecture that seeks the **right balance** between **distributed intelligence** and **centralized optimization**.

**Paths options**
- Dynamic (SPF computation)
- Explicit (expressed in the packet)

**Control Plane**
- Routing protocols with extensions (IS-IS, OSPF, BGP)

**Data Plane**
- MPLS (segment labels)
- IPv6 (+SR header)
- SR-MPLS
- SRv6

**Data**
- SID R5
- SID R7
- R1
- R2
- R3
- R4
- R5
- R6
- R7

**Path expressed in the packet**
- SID R7
- Data

**Dynamic path**

**Explicit path**

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Public
Rich SRv6 uSID Ecosystem

Network Equipment Manufacturers

- Cisco
- Huawei
- Nokia
- Ciena
- Juniper Networks
- ZTE
- Arrcus
- Arista
- H3C

Merchant Silicon

- Broadcom
- Marvell
- Barefoot Networks

Open-Source Applications

- Pyroute2
- SERA
- GoBGP
- eBPF
- cilium
- Wireshark
- NETMAP
- BGP

Open-Source Networking Stacks

- Open Compute Project
- SONIC
- FRRouting

Smart NIC

- Intel
- Mellanox Technologies

Partners

- kaloom
- Trend Micro
- ENEA
- Qosmos Division
- NoviFlow
- Ixia
- Spirent

#CiscoLive BRKMPL-2205
SRv6 Mature Standardization

- Proposed Standard RFCs
  - RFC8402 SR Architecture
  - RFC8754 SRv6 Data Plane
  - RFC8986 SRv6 Network Programming
  - RFC9352 SRv6 ISIS Extensions
  - RFC9350 IGP Flexible Algorithms
  - RFC9252 SRv6 BGP Extensions
  - RFC9256 SR Policy Architecture
  - RFC9259 SRv6 OAM

- WG Document: Proposed Standard
  - WG draft Compressed SRv6 Segment List encoding in SRH
  - WG draft SRv6 BGP Link State Extensions

Much faster standardization than usual
Sign of the SRv6 Industry Endorsement

Strong Cisco Commitment to IETF

- Editor of 96% IETF RFCs
- Co-author of 100% IETF RFCs
- Editor of 77% IETF WG Drafts
- Co-author of 84% IETF WG Drafts

Over 70 RFCs/ Drafts spanning 13 working groups
EANTC 2023

• Published on April 18th at MPLS-WC
• Industry endorsement of uSID solution
  • L3VPN over uSID
  • SRv6 OAM over uSID
  • BGP GRT with uSID
  • EVPN VPWS (Multihoming), ELAN, RT5
  • SRv6 Locator (and FA) summarization with uSID
  • SR TE Policy with uSID
• Arista, Arrcus, Huawei, Juniper, Nokia

Whitepaper: https://eantc.de/de/showcases/2023/mpls_sdn_interop.html
Comparing MPLS and SRv6 (plus SRv6 101)
Comparing (SR-)MPLS and SRv6

- MPLS provides a transport service by applying one or more labels to a service packet
  - Segment lists require one label per segment
  - MPLS requires 1:1 label to /32 prefix

- SRv6 provides a transport service by encapsulating the packet with IPv6
  - Segment lists are encoded as uSID in the IPv6 header
  - SRv6 enables summarization (huge benefit!)
Comparing MPLS and IPv6 Headers

MPLS Label Format (RFC 3032)

- **Function**: Path/Service encoding
  - **MPLS**: 20-bit Label
  - **IPv6**: 128-bit DA

- **Function**: Flow Identification
  - **MPLS**: FAT/Entropy Label(s)
  - **IPv6**: 20-bit Flow Label

- **Function**: QoS Identification
  - **MPLS**: 3-bit EXP
  - **IPv6**: 8-bit Traffic Class

- **Function**: Loop prevention
  - **MPLS**: 8-bit TTL
  - **IPv6**: 8-bit Hop Limit
IPv6 Addressing Review

• Representation of prefix is similar to IPv4, except length is 128 bits vs 32 bits
  • v4 address:
    • 198.10.0.0/16
  • v6 address:
    • 2001:db8:12::/48

• Leading zeros in contiguous block could be represented by (::)
  • 2001:0db8:0000:130F:0000:0000:087C:140B → 2001:0db8:0:130F:087C:140B

• Double colon only appears once in the address

• Only leading zeros are omitted. Trailing zeros are not omitted
  • 2001:0db8:0012::/48 = 2001:db8:12::/48
  • 2001:db8:1200::/48 ≠ 2001:db8:12::/48
SRv6 Addressing (F3216)

Overview

- Each node is assigned a 48-bit (/48) “locator”
  - /16 and /48 are also possible
  - Analogous to the v4 Loopback address used in SR-MPLS
- The first 32 bits (/32) will be common among all nodes in the same ISIS topology
- The next 16 bits are unique per node
- Recommended to use IPv6 Unique Local Address (FC00::/7)
SRv6 Segment List Encoding and Micro-SID (uSID) Compression

Original SRv6 SID List Encoding

- **SA:** 2001::1
- **DA:** fcaa:bb20:0:4:e002:0:0:0
- **NH:** RH

**Type:** 4 (SRH)
**NH:** IPv4|SL:1
**Segment List:**
0: fcaa:bb20:0:4:e002:0:0:0
1: fcaa:bb20:0:3:48:0:0:0
2: fcaa:bb20:0:2:0:0:0:0
3: fcaa:bb20:0:1:0:0:0:0

IPv4
- **SA:** 7.5.4.3
- **DA:** 11.6.19.71
- **Port:** UDP

UDP Header/Data

SRv6 uSID Encoding

- **SA:** 2001::1
- **NH:** IPv4

IPv4
- **SA:** 7.5.4.3
- **DA:** 11.6.19.71
- **Port:** UDP

UDP Header/Data

128-bit SIDs encoded in SR Extension Header

32-bit block

16-bit uSIDs

Last SID analogous to Service Label

16-bit uSIDs range from 0x0 to 0x7fff

Micro-SID Encoding

- 128-bit SIDs
- 32-bit block
- 16-bit uSIDs

Original SRv6 SID List Encoding

- **SA:** 2001::1
- **DA:** fcaa:bb20:0:4:e002:0:0:0
- **NH:** RH

**Type:** 4 (SRH)
**NH:** IPv4|SL:1
**Segment List:**
0: fcaa:bb20:0:4:e002:0:0:0
1: fcaa:bb20:0:3:48:0:0:0
2: fcaa:bb20:0:2:0:0:0:0
3: fcaa:bb20:0:1:0:0:0:0

IPv4
- **SA:** 7.5.4.3
- **DA:** 11.6.19.71
- **Port:** UDP

UDP Header/Data

16-bit uSIDs

Last SID analogous to Service Label
SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example

SRv6 uSID Example
Comparing SRv6 and MPLS SID Imposition

### Example MPLS Label Stack with 11 SID Labels + 1 Service Label

- **SID 11:** 16011
- **SID 10:** 16010
- **SID 9:** 16009
- **SID 8:** 16008
- **SID 7:** 16007
- **SID 6:** 16006
- **SID 5:** 16005
- **SID 4:** 16004
- **SID 3:** 16003
- **SID 2:** 16002
- **SID 1:** 16001
- **SVC:** 80022

### Example SRv6 with 24+2 uSID

- **SA:** 2001::1
- **NH:** RH
- **Type:** 4 (SRH)
- **NH:** IPv4 | SL: 4
- **Segment List:**
  - [0]: fcaa:bb20:119:e200:0:0:0:0:0:0

- **IPv4**
  - **SA:** 7.5.4.3
  - **DA:** 11.6.19.71
  - **Port:** UDP

### SRv6 Parameters

<table>
<thead>
<tr>
<th></th>
<th>BRCM Q/J/J+</th>
<th>BRCM Q2/J2</th>
<th>ASR 9k LSP</th>
<th>Cisco 8k Q200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum MPLS Label Imposition</td>
<td>3+3 (9+3)</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Maximum SRv6 uSID (Headend Tx)</td>
<td>3+3 (6+3)</td>
<td>24+3</td>
<td>12+3</td>
<td>6+3 (17+3)</td>
</tr>
</tbody>
</table>

**Line Rate (With Recirculation):**

- IPv4
  - **SA:** 7.5.4.3
  - **DA:** 11.6.19.71
  - **Port:** UDP

---

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Public
Segment Routing Flexible Algorithm ("Flex Algo") with SR Performance Monitoring

- SR Performance Monitoring – All nodes actively measure latency (using TWAMP-Lite probes) and report via ISIS link-state updates
- Flexible Algorithm - a numeric identifier in the range 128-255 that is associated via configuration with the Flexible-Algorithm Definition.
  - All nodes have Shortest Path First (SPF) Algo 0 by default
- Low-Latency Flex Algo is defined in order to steer prioritized traffic along lowest latency path

Node 0 will steer traffic toward either the best effort or low-latency slice depending on the network destination

Node 9 will advertise network reachability for best effort and low-latency prefixes

- Leverages IOS-XR initial SR-MPLS Flex Algo implementation
- Now fully supported for SRv6
TI-LFA for SRv6

• **50msec Protection** upon local link, node or SRLG failure

• **Simple** to operate and understand
  • automatically computed by the router’s IGP process
  • 100% coverage across any topology
  • predictable (backup = postconvergence)

• **Optimum backup path**
  • leverages the post-convergence path, planned to carry the traffic
  • avoid any intermediate flap via alternate path

• **Incremental deployment**

• **Distributed and Automated Intelligence**

Leveraging the existing TI-LFA SR-MPLS code
• FCS since 2014
• Numerous deployments

![Diagram of TI-LFA for SRv6](image-url)
SRv6 Architecture Advantages
L2/L3 Service over Classic MPLS + BGP-LU

• This common design requires multiple protocols with significant complexity
• BGP-LU is used as a “shim layer” primarily to reduce the size of the IGP domain
• In many cases LDP or SR-MPLS are used instead of RSVP-TE, but overall, the picture doesn’t change
Load-Balancing Challenges with MPLS

1. BGP, Service & FAT label imposed by A1 & A2 (note: some platforms cannot support this function)

2. LAG required for RSVP-TE load balancing

3. P2 may not find FAT label depending on silicon limitations

4. If no FAT label, P2 would have to speculative hash up to L4!

5. If FAT or speculative hashing fails, traffic will not be properly load-balanced
L2/L3 Service over SRv6

- SRv6 greatly simplifies the design
- Eliminates BGP-LU, RSVP-TE, LDP
# Load-Balancing with SRv6

A1 & A2 map traffic flows to IPv6 flow label in the header.

<table>
<thead>
<tr>
<th>Layer 4 (L4)</th>
<th>Layer 5 (IP)</th>
<th>Layer 6 (MPLS)</th>
<th>Layer 7 (ETH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETH</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. L4  
2. IP  
3. MPLS  
4. ETH

P2 will easily identify flows with IPv6 flow label in the header.

<table>
<thead>
<tr>
<th>Layer 4 (L4)</th>
<th>Layer 5 (IP)</th>
<th>Layer 6 (MPLS)</th>
<th>Layer 7 (ETH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETH</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pseudowire

ISISv6

3. No LAG required

Traffic is properly load-balanced across all ECMP links

A1 & A2 map traffic flows to IPv6 flow label in the header.
Inter-Domain using BGP-LU

- This common design requires BGP-LU to stitch between domains within the SP
- Also commonly used to stitch between sub-domains within a domain
- BGP-LU adds a substantial tax of complexity and limits scalability
Inter-Domain Using SRv6

- SRv6 eliminates the BGP-LU shim layer and significantly improves scalability through summarization

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Public
MPLS Does not Support Summarization

PE-10 Forwarding Table
(100x IPv4 /32 prefixes)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.0.0.1/32 via BR-12</td>
<td>24242</td>
</tr>
<tr>
<td>20.0.0.2/32 via BR-12</td>
<td>24356</td>
</tr>
<tr>
<td>20.0.0.3/32 via BR-12</td>
<td>24875</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>20.0.0.100/32 via BR-12</td>
<td>24148</td>
</tr>
</tbody>
</table>

Label Push Operation (ip2mpls)

IGP Area or Domain 1

BR-12 Advertises 100x /32 prefixes+labels via BGP-LU

100x /32 Loopback IPv4 Prefixes

<table>
<thead>
<tr>
<th>PE</th>
<th>Loopback</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE-1</td>
<td>20.0.0.1/32</td>
</tr>
<tr>
<td>PE-2</td>
<td>20.0.0.2/32</td>
</tr>
<tr>
<td>PE-3</td>
<td>20.0.0.3/32</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>PE-100</td>
<td>20.0.0.100/32</td>
</tr>
</tbody>
</table>

100 PEs

Label Swap Operation (mpls2mpls)

IGP Area or Domain 2

Packet for PE-1
Packet for PE-2
Packet for PE-3
...   
Packet for PE-100
Summarization with SRv6

PE-10 Forwarding Table
(1x IPv6 /32 prefix)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>fcaa:bb20::/32 via BR-12</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Longest Prefix Match forwards all Domain 2 packets via BR-12

BR-12 Advertises 1x IPv6 Summary Route (fcaa:bb10::/32)

100x /48 Locator IPv6 Prefixes

<table>
<thead>
<tr>
<th>PE</th>
<th>Locator</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE-1</td>
<td>fcaa:bb20:15/48</td>
</tr>
<tr>
<td>PE-2</td>
<td>fcaa:bb20:16/48</td>
</tr>
<tr>
<td>PE-3</td>
<td>fcaa:bb20:17/48</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>PE-100</td>
<td>fcaa:bb20:78/48</td>
</tr>
</tbody>
</table>

100 PEs

IGP Area or Domain 1

Simple IP Forwarding Operation

IGP Area or Domain 2

Longest Prefix Match forwards all Domain 2 packets via BR-12

Packet for PE-1

Packet for PE-2

Packet for PE-3

... Packet for PE-100

Packet for PE-1

Packet for PE-2

Packet for PE-3

... Packet for PE-100
SRv6 End-to-End Routing Example

s1 BGP Table
- c2 via s2
  (SID fcaa:bb20:202:e002)

s1 Routing Table
- fcaa:bb20::/32 via g2

r1 Routing Table
- fcaa:bb20::/32 via r2

r2 Routing Table
- fcaa:bb20::/32 via g3

r3 Routing Table
- fcaa:bb20:202::/48 via s2

g3 Locator:
- fcaa:bb20:202::/48

IPv6 Only Network

IPv6 Summary Route

IPv6 Only Network

SRv6 Network 1

SRv6 Network 2

SA:2001::1
DA:fcaa:bb20:202:e002::
NH:IPv4

{IPv4 Packet}

SA:2001::1
DA:fcaa:bb20:202:e002::
NH:IPv4

{IPv4 Packet}

SA:2001::1
DA:fcaa:bb20:202:e002::
NH:IPv4

{IPv4 Packet}

s2 announces via BGP:
c2 via s2,
with SID: fcaa:bb20:202:e002

g3 considers first 48 bits of IPv6 Destination Address (Longest Prefix Match)

s1, g2, r1, r2 only consider first 32 bits of IPv6 Destination Address (Longest Prefix Match)
Inter-Domain Routing with BGP-LU

<table>
<thead>
<tr>
<th>PE-1</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP-LU FIB Entries</td>
<td>5000</td>
</tr>
<tr>
<td>BGP-LU Backup FIB Entries</td>
<td>4000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ASBR-1/ASBR-2</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP-LU FIB Entries</td>
<td>5000</td>
</tr>
<tr>
<td>BGP-LU Backup FIB Entries</td>
<td>5000</td>
</tr>
</tbody>
</table>

9k FIB entries  10k FIB entries
Inter-Domain Routing with SRv6 and Summarization

<table>
<thead>
<tr>
<th>PE-1</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6 inter-domain (via ASBR-1)</td>
<td>4</td>
</tr>
<tr>
<td>IPv6 inter-domain (via ASBR-2)</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ASBR-1/ASBR-2</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6 inter-domain (via each metro ASBR pair)</td>
<td>8</td>
</tr>
</tbody>
</table>

9k FIB entries reduced to 8!

10k FIB entries reduced to 8!
### MPLS vs SRv6 Hardware Programming

**MPLS**
- MPLS can have IP routing table indicating next-hop reachability, but there is no label switched path (broken LSP)

**SRv6**
- SRv6 will always have consistency between RIB and FIB since it is forwarding with SRv6 as native IPv6 routing
- Path tracing can identify any issues with forwarding down-stream

---

**RIB**
- 20.0.0.0/24 via PE-2
- PE-2 ISIS
- 30.0.0.0/24 via PE-3
- PE-3 ISIS

**FIB**
- 20.0.0.0/24 via PE-2
- PE-2 LSP12
- 30.0.0.0/24 via PE-3
- PE-3 Unresolved

---

**SRv6 (IPv6 routing!)**
- 20.0.0.0/24 via PE-2
- PE-2 IPv6
- 30.0.0.0/24 via PE-3
- PE-3 IPv6
How did the packet arrive from A to F?

- 3 possible “valid” ECMP paths
  - Any drop?
  - End-to-End Latency homogeneity?
- An invalid path is possible
  - Routing or FIB corruptions
- 40-year-old unsolved IP problem
SRv6 Path Tracing

- Each transit router records in PT header:
  - Outgoing interface ID
  - Timestamp (with 60µs accuracy)
  - Egress Queue Load
- Highly compressed for low MTU overhead
  - Only 3 bytes per hop!
- Implemented at line rate: Reports true packet experience
- Native interworking with legacy nodes
  - Seamless deployment
- Hardware/XR feature with analytics app
Data Plane & Service Chaining with MPLS

- SP Access/Aggregation utilizes MPLS transport which must be decapsulated at the DC
- Other mechanisms (e.g. VXLAN, NSH) must provide service-chaining and routing/switching through the DC
- Transporting services through the public cloud requires additional tunneling mechanism (e.g. GRE)
Unified IPv6 Dataplane with SRv6

- Enables native routing (IPv6) to cloud/virtual data center providers
  - Unified IPv6 dataplane from socket to Internet peering through DC, Access, Metro, Core
- Can route traffic through devices without SRv6 functionality
- Greatly optimizes and simplifies service-chaining

Path Tracing enables:
- Deterministic confirmation of NFV processing
- Deterministic latency measurement of the NFV processing
SRv6 Test-Cases
What can I do with SRv6?

• L3VPN
• L2VPN
• BGP Free Core
• Segment Routing Traffic Engineering (SR-TE)
• Network “slicing” with FlexAlgo
1) Configure SRv6 Locator

```
segment-routing
srv6
locators
locator MAIN
    micro-segment behavior unode psp-usd
prefix fcbb:bb00:100::/48
```

2) Enable Interfaces for IPv6

```
interface Loopback0
    ipv6 address 2010:0:100::1/128
! interface HundredGigE0/0/0/2
    ipv6 enable
    ! interface HundredGigE0/0/0/3
    ipv6 enable
```

Note:
- No link addressing required (can use IPv6 link local addressing)

3) Enable SRv6 for the IGP

```
router isis 1
    is-type level-2-only
    net 39.0100.0000.0000.0100.00
    address-family ipv6 unicast
    metric-style wide
    router-id Loopback0
    segment-routing srv6
    locator MAIN
    ! interface Loopback0
    passive
    address-family ipv6 unicast
    ! interface GigabitEthernet0/0/0/0
    point-to-point
    address-family ipv6 unicast
    ! interface GigabitEthernet0/0/0/1
    point-to-point
    address-family ipv6 unicast
```
L3VPN Service over SRv6 Example
R100 Example Configurations

1) Define VRF

```configuration
vrf BLUE_VRF
  address-family ipv4 unicast
  import route-target 1:123
  export route-target 1:123
  interface Loopback222
    vrf BLUE_VRF
    ipv4 address 172.16.10.100/32
  interface GigabitEthernet0/0/0/3.231
    vrf BLUE_VRF
    ipv4 address 10.0.231.2/30
    encapsulation dot1q 231
```

2) Configure VRF for SRv6 under BGP

```configuration
router bgp 10
  bgp router-id 100.0.0.100
  address-family vpnv4 unicast
  {configure neighbor for vpnv4 unicast}
  vrf BLUE_VRF
    rd 1:123
    address-family ipv4 unicast
    segment-routing srv6
    locator MAIN
    alloc mode per-vrf
    redistribute connected
  !
```

Similar configurations applied to R200

---

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Public
L3VPN Service over SRv6 Example
BGP Update for BLUE_VRF 172.16.200.0/24

RP/0/RP0/CPU0:xr9kv-100#show bgp vpnv4 uni vrf BLUE_VRF 172.16.200.0/24
<snip>
BGP routing table entry for 172.16.200.0/24, Route Distinguisher: 1:123
<snip>
Paths: (1 available, best #1)
<snip>
2001:0:200::1 (metric 30) from 2001:0:30::1 (30.0.0.30)
Received Label 0xe0050
<snip>
Extended community: Color:12905 RT:1:123
PSID-Type:L3, SubTLV Count:1
SubTLV:
T:1(Sid information), Sid:fcbb:bb00:200::, Behavior:63, SS-TLV Count:1
SubSubTLV:
T:1(Sid structure):
Source AFI: VPNv4 Unicast, Source VRF: BLUE_VRF, Source Route Distinguisher: 1:123
L3VPN Service over SRv6 Example
R100 Packet Capture Towards R200

<table>
<thead>
<tr>
<th>CE-1</th>
<th>R100</th>
<th>SRv6</th>
<th>R200</th>
<th>CE-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRv6 MAIN: fcbb:bb00:100::/48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRv6 MAIN: fcbb:bb00:200::/48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lo0: 2001:0:100::1
Lo0: 2001:0:200::1

CE-1
Lo0: 2001:0:200::1
SRv6: fcbb:bb00:200::/48

CE-2
172.16.200.0/24

### Flow Stock

| Frame 1: 1232 bytes on wire (9856 bits), 1232 bytes captured (9856 bits) |
| Ethernet II, Src: Realtek_03:0f:1d (52:54:00:03:0f:1d), Dst: Realtek_02:79:ea (52:54:00:02:79:ea) |
| Internet Protocol Version 6, Src: 2001:0:100::1, Dst: fcbb:bb00:200::e005:: |
| 0110 .... = Version: 6 |
| .... 0000 0000 .... .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT) |
| .... 1000 0000 1110 1010 1101 Flow Label: 0x8e0ad |
| Payload Length: 1138 |
| Next Header: IPIP (4) |
| Hop Limit: 255 |
| Source Address: 2001:0:100::1 |
| Destination Address: fcbb:bb00:200::e005: |
| Source Teredo Server IPv4: 1.0.0.0 |
| Source Teredo Port: 85555 |
| Source Teredo Client IPv4: 255.255.255.254 |
| Internet Protocol Version 4, Src: 10.0.231.1, Dst: 172.16.200.29 |
| Data (1138 bytes) |

SRv6 SID list encoded as uSID in IPv6 DA

Flow label computed
No SRH, next-header IPv4
EVPN-VPWS over SRv6 Service Example
R100 Example Configuration

Similar configurations applied to R200

Note: EVPN BGP configurations are not shown but do not require any SRv6 unique configurations
R100 EVPN-VPWS Control Plane State

```
RP/0/RP0/CPU0:xr9kv-100#show l2vpn xconnect detail
<snip>
Group EVPN-VPWS, XC EVPN-VPWS-123, state is up; Interworking none
  AC: GigabitEthernet0/0/0/3.123, state is up
<snip>
Statistics:
  packets: received 3325526, sent 0
  bytes: received 3977329096, sent 0
  drops: illegal VLAN 0, illegal length 0
EVPN: neighbor ::ffff:10.0.0.1, PW ID: evi 123, ac-id 12123, state is up ( established )
  XC ID 0xa0000005
  Encapsulation SRv6
  Encap type Ethernet
  Ignore MTU mismatch: Enabled
  Transmit MTU zero: Enabled
  Reachability: Up
  Load Balance Hashing: src-dst-ip
```

<table>
<thead>
<tr>
<th>SRv6</th>
<th>Local</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>uDX2</td>
<td>fcbb:bb00:100:e008::</td>
<td>fcbb:bb00:200:e008::</td>
</tr>
<tr>
<td>AC ID</td>
<td>12123</td>
<td>12123</td>
</tr>
<tr>
<td>MTU</td>
<td>1514</td>
<td>0</td>
</tr>
<tr>
<td>Locator</td>
<td>MAIN</td>
<td>N/A</td>
</tr>
<tr>
<td>Locator Resolved</td>
<td>Yes</td>
<td>N/A</td>
</tr>
<tr>
<td>SRv6 Headend</td>
<td>H.Encaps.L2.Red</td>
<td>N/A</td>
</tr>
</tbody>
</table>
EVPN-VPWS over SRv6 Service Example

R100 Packet Capture Towards R200

EVPN-VPWS PW

SRv6 MAIN: fcbb:bb00:100::/48
SRv6 MAIN: fcbb:bb00:200::/48

Lo0: 2001:0:100::1
Lo0: 2001:0:200::1

CE-1
VLAN 123

R100

CE-2
VLAN 123

R200

172.16.200.0/24

Flow label computed
No SRH, next-header Ethernet

SRv6 SID list encoded as uSID in IPv6 DA

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Public
“BGP Free” Core Example Configuration

- Configure global IPv4 AF for SRv6 under BGP
- Configure neighbor for SRv6 encapsulation
R100 BGP Entry for 199.0.0.0/8

RP/0/RP0/CPU0:xr9kv-100#show bgp ipv4 uni 199.0.0.0/8
Sat Jun  3 19:54:30.283 UTC
BGP routing table entry for 199.0.0.0/8
Versions:
  Process bRIB/RIB SendTblVer
  Speaker  31  31
Last Modified: Jun  3 18:56:26.833 for 00:58:03
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  2
    2001:0:200::1 (metric 30) from 2001:0:30::1 (30.0.0.30)
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best
      Received Path ID 1, Local Path ID 1, version 31
      PSID-Type:L3, SubTLV Count:1
      SubTLV:
      T:1(Sid information), Sid:fcbb:bb00:200:e006::, Behavior:63, SS-TLV Count:1
      SubSubTLV:
      T:1(Sid structure):
“BGP Free” Core Packet Capture R100 to R200

Frame 1: 1232 bytes on wire (9856 bits), 1232 bytes captured (9856 bits)
Ethernet II, Src: RealtekU_1c:55:d8 (52:54:00:1c:55:d8), Dst: RealtekU_1d:75:02 (52:54:00:1d:75:02)
Internet Protocol Version 6, Src: 2001:0:100::1, Dst: fccb:bb00:200::e005::
0110 .... = Version: 6
> .... 0000 0000 ........ .......... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
> .... 1101 1010 1001 0000 1001 = Flow Label: 0xda909
Payload Length: 1178
Next Header: IPiP (4)
Hop Limit: 255
Source Address: 2001:0:100::1
Destination Address: fccb:bb00:200::e006::
  [Source Teredo Server IPv4: 1.0.0.0]
  [Source Teredo Port: 55535]
  [Source Teredo Client IPv4: 255.255.255.254]
Internet Protocol Version 4, Src: 10.0.231.1, Dst: 199.128.64.5
Data (1138 bytes)

SRv6 SID list encoded as uSID in IPv6 DA
Flow label computed
No SRH, next-header IPv4
SRv6 with SR-TE Example

- default IGP & TE metric is 10

```
segment-routing
traffic-eng
  on-demand color 200
srv6
    locator MAIN binding-sid dynamic behavior ub6-insert-reduced
  !
dynamic
  pcep
  !
  metric
type te
```

```
  extcommunity-set opaque COLOR-200
    200
  end-set
!
route-policy SET-COLOR-IPV4-GRT
  if destination in (99.0.0.0/8) then
    set extcommunity color COLOR-200
    pass
  else
    pass
  endif
end-policy
```
# R100 SR-TE Policy for Color 200

```bash
RP/0/RP0/CPU0:xr9kv-100#show segment-routing traffic-eng policy color 200

Color: 200, End-point: 2001:0:200::1
Name: srte_c_200_ep_2001:0:200::1
Status:
  Admin: up  Operational: up for 01:03:11 (since Jun  3 17:00:18.901)
Candidate-paths:
Preference: 100 (BGP ODN) (active)

Dynamic (pce 2001:0:30::1) (valid)
Metric Type: TE, Path Accumulated Metric: 50
SID[0]: fcbb:bb00:103::/48 Behavior: uN (PSP/USD) (48)
  Format: f3216
  LBL:32 LNL:16 FL:0 AL:0
  Address: 2001:0:103::1
SID[1]: fcbb:bb00:201::/48 Behavior: uN (PSP/USD) (48)
  Format: f3216
  LBL:32 LNL:16 FL:0 AL:80
  Address: 2001:0:201::1
SID[2]: fcbb:bb00:203::/48 Behavior: uN (PSP/USD) (48)
  Format: f3216
  LBL:32 LNL:16 FL:0 AL:0
  Address: 2001:0:203::1
SID[3]: fcbb:bb00:200::/48 Behavior: uN (PSP/USD) (48)
  Format: f3216
  LBL:32 LNL:16 FL:0 AL:80
  Address: 2001:0:200::1
```
Packet Capture on R100 → R103 Link

- SRv6 SID list encoded as uSID in IPv6 DA
- Flow label computed
- No SRH, next-header IPv4
Interdomain Network “Slicing” with Flex Algo

- SR-PM, “DELAY” Locator & Flex Algo configured on all routers

---

```
segment-routing
sv6
  locators
  locator MAIN
    micro-segment behavior unode psp-usd
    prefix fcbb:bb00:100::/48
  locator DELAY
    micro-segment behavior unode psp-usd
    prefix fcbb:bb03:100::/48
  algorithm 130
```

---

```
router isis 1
  is-type level-2-only
  net 39.0100.0000.0000.0100.00
  flex-algo 130
  metric-type delay
  advertise-definition
    address-family ipv6 unicast
    metric-style wide
    router-id Loopback0
  segment-routing srv6
    locator MAIN
  segment-routing srv6
    locator DELAY
```

---

```
summary-prefix fcbb:bb03:200::/40 level 2 algorithm 130 explicit
```

---

SRv6 MAIN: fcbb:bb00:100::/48
SRv6 DELAY: fcbb:bb03:100::/48

Lo0: 2001:0:100::1

SRv6 MAIN: fcbb:bb00:100::/48
SRv6 DELAY: fcbb:bb03:100::/48
Interdomain Network “Slicing” with Flex Algo
R200 Example Configurations

• All routers forward packets along lowest delay path according to “DELAY” Flex Algo topology

• No controller or PCE required

```
route-policy SET-ALGO
    if destination in (172.16.210.0/24) then
        set srv6-alloc-mode per-vrf locator DELAY
        pass
    else
        set srv6-alloc-mode per-vrf locator MAIN
        pass
    endif
end-policy
```

• default delay 100
R100 BGP Entry for 172.16.210.0/24

show bgp vpnv4 uni vrf BLUE_VRF 172.16.210.0/24

BGP routing table entry for 172.16.210.0/24, Route Distinguisher: 1:123

Paths: (1 available, best #1)

2001:0:200::1 (metric 30) from 2001:0:30::1 (30.0.0.30)
Received Label 0xe0040

Extended community: RT:1:123
PSID-Type:L3, SubTLV Count:1
SubTLV:
  T:1(Sid information), Sid:fcbb:bb03:200::, Behavior:63, SS-TLV Count:1
SubSubTLV:
  T:1(Sid structure):
Source AFI: VPNv4 Unicast, Source VRF: BLUE_VRF, Source Route Distinguisher: 1:123

show isis database xr9kv-201.00-00 detail

IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
xr9kv-201.00-00       0x00000066   0xdb76        785  /1200         0/0/0
Area Address:   39.0100

SRv6 Locator:  MT (IPv6 Unicast) fcbb:bb00:200::/40 D:0 Metric: 11 Algorithm: 0
SRv6 Locator:  MT (IPv6 Unicast) fcbb:bb03:200::/40 D:0 Metric: 201 Algorithm: 130
Interdomain Network “Slicing” with Flex Algo
Packet Capture R100→R200 (DELAY FA)

- No Controller or PCE!

172.16.210.0/24
locator DELAY
Lo0: 2001:0:200::1
SRv6 MAIN: fcbf:bb00:200::/48
SRv6 DELAY: fcbf:bb03:200::/48

SRv6 SID list encoded as uSID in IPv6 DA

Flow label computed
No SRH, next-header IPv4

- default delay 100
Cisco Platform Support for SRv6
### SRv6 Feature Support for Cisco IOS-XR Platforms

<table>
<thead>
<tr>
<th>Feature name</th>
<th>NCS 5500 NCS 540</th>
<th>NCS 560</th>
<th>NCS 5700 NCS540-Q2A</th>
<th>ASR9K (LSP)</th>
<th>8000 (Q200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSR: ISIS (incl. Ti-LFA / uLoop / Flex-Algo)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>OAM (Ping, Traceroute, SID Verification)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>SRv6 PM (Delay, Loss, Liveness)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>Seamless Migration (F1 -&gt; uSID + Dual-mode)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>L3 Services: VPNv4 / VPNv6</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>L3 Services: IPv4 / IPv6 Internet (GRT)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>L2 Services: EVPN-VPWS (ELINE P2P)</td>
<td>Supported</td>
<td></td>
<td></td>
<td></td>
<td>Not supported</td>
</tr>
<tr>
<td>L2 Services: EVPN (ELAN BD)</td>
<td>Supported</td>
<td></td>
<td>Roadmap</td>
<td>Supported</td>
<td>Not supported</td>
</tr>
<tr>
<td>SRv6TE: SRv6 PCE (ODN)</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>SRv6TE: Headend w/ Explicit Path</td>
<td></td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
<tr>
<td>Path Tracing</td>
<td>Not supported</td>
<td></td>
<td></td>
<td>Supported</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion
Key Takeaways

• SRv6 is gaining significant traction with network operators globally

• SRv6 is fully standardized and ready for deployment

• Services delivered today with MPLS can be delivered with SRv6 with greater simplicity and scalability

• Cisco is making significant investments in SRv6 across our portfolio
Stay up to date with...

Segment Routing
www.segment-routing.net/

LinkedIn
www.linkedin.com/groups/8266623

Twitter
www.twitter.com/SegmentRouting

Facebook
www.facebook.com/SegmentRouting/

Segment Routing, Part I / II Textbooks
Available on Kindle and in paperback
## More SRv6 @ Cisco Live Las Vegas 2023

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Time &amp; Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRKSPG-2039</td>
<td>Architecting Modern Broadband Networks</td>
<td>Monday, Jun 5 3:00 PM - 4:30 PM PDT Level 2, Mandalay Bay L</td>
</tr>
<tr>
<td>BRKSPG-2043</td>
<td>Simplify your journey to SR and SRv6 with Crosswork Automation</td>
<td>Monday, Jun 5 4:00 PM - 5:00 PM PDT Level 2, Oceanside F</td>
</tr>
<tr>
<td>BRKMPL-2203</td>
<td>SRv6 Fundamentals</td>
<td>Tuesday, Jun 6 3:00 PM - 4:30 PM PDT Level 3, South Seas B</td>
</tr>
<tr>
<td>BRKMPL-2117</td>
<td>SRv6 based IP Transport – Design, Deployment Best Practices &amp; Challenges</td>
<td>On-Demand Video</td>
</tr>
<tr>
<td>LABMPL-1201</td>
<td>SRv6 Basics</td>
<td>Walk-in Lab</td>
</tr>
<tr>
<td>LABSP-3393</td>
<td>Implementing Segment Routing v6 (SRv6) Transport on NCS 55xx/5xx platforms</td>
<td>Walk-in Lab</td>
</tr>
</tbody>
</table>

Cisco Group Discount available for SRv6 Tattoos 😊
Fill out your session surveys!

Attendees who fill out a minimum of four session surveys and the overall event survey will get **Cisco Live-branded socks** (while supplies last)!

Attendees will also earn 100 points in the **Cisco Live Challenge** for every survey completed.

**These points** help you get on the leaderboard and increase your chances of winning daily and grand prizes.
Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand
Thank you
Gamify your Cisco Live experience!
Get points for attending this session!

How:
1. Open the Cisco Events App.
2. Click on ‘Cisco Live Challenge’ in the side menu.
3. Click on View Your Badges at the top.
4. Click the + at the bottom of the screen and scan the QR code: