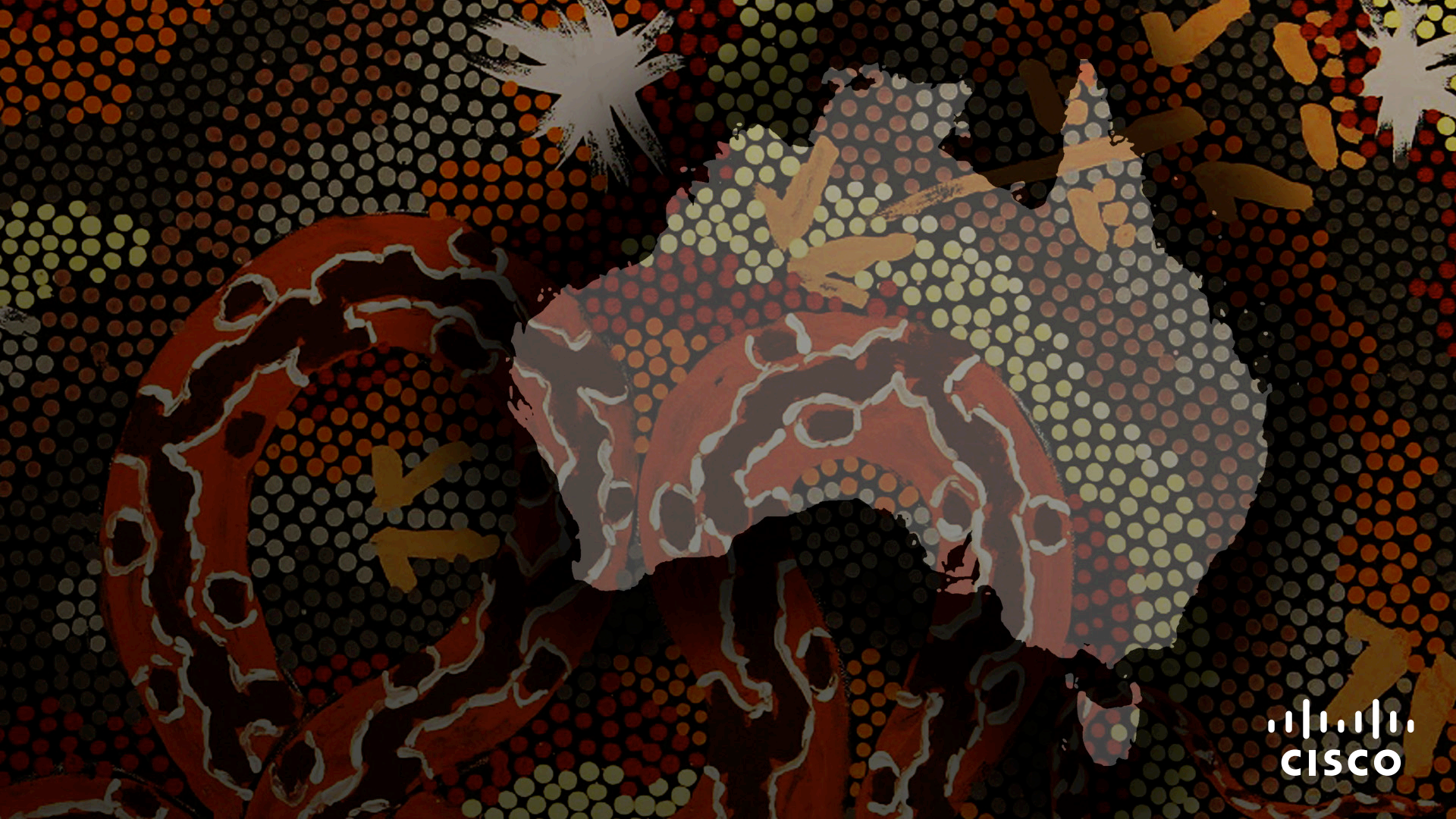
#CiscoLiveAPJC

The bridge to possible

Introduction to SRv6 uSID

Jakub Horn
Principal Technical Marketing Engineer

BRKSPG-3058


cisco

Cisco Webex App

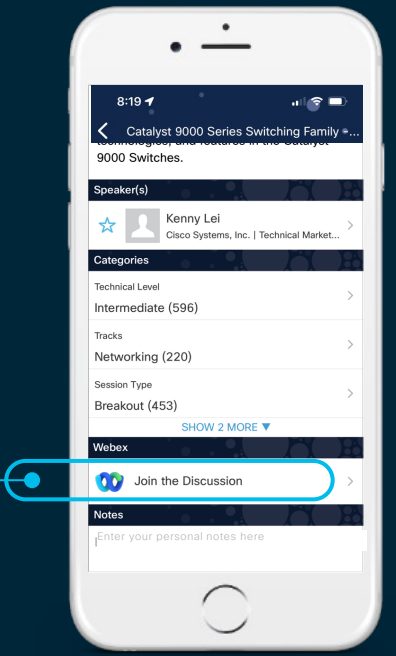
Questions?

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 Thursday 22 December, 2022.

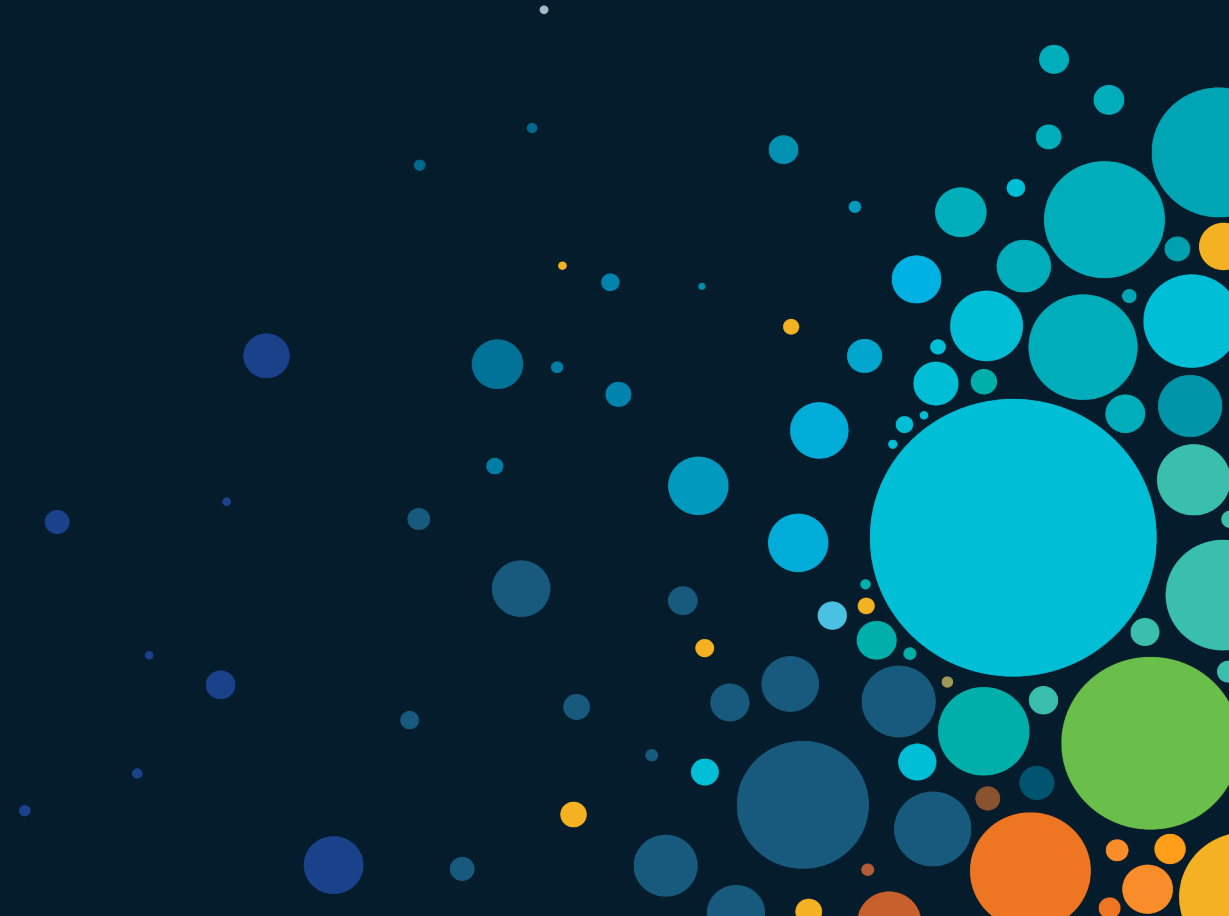


https://ciscolive.ciscoevents.com/ciscolivebot/#BRKSPG-3058

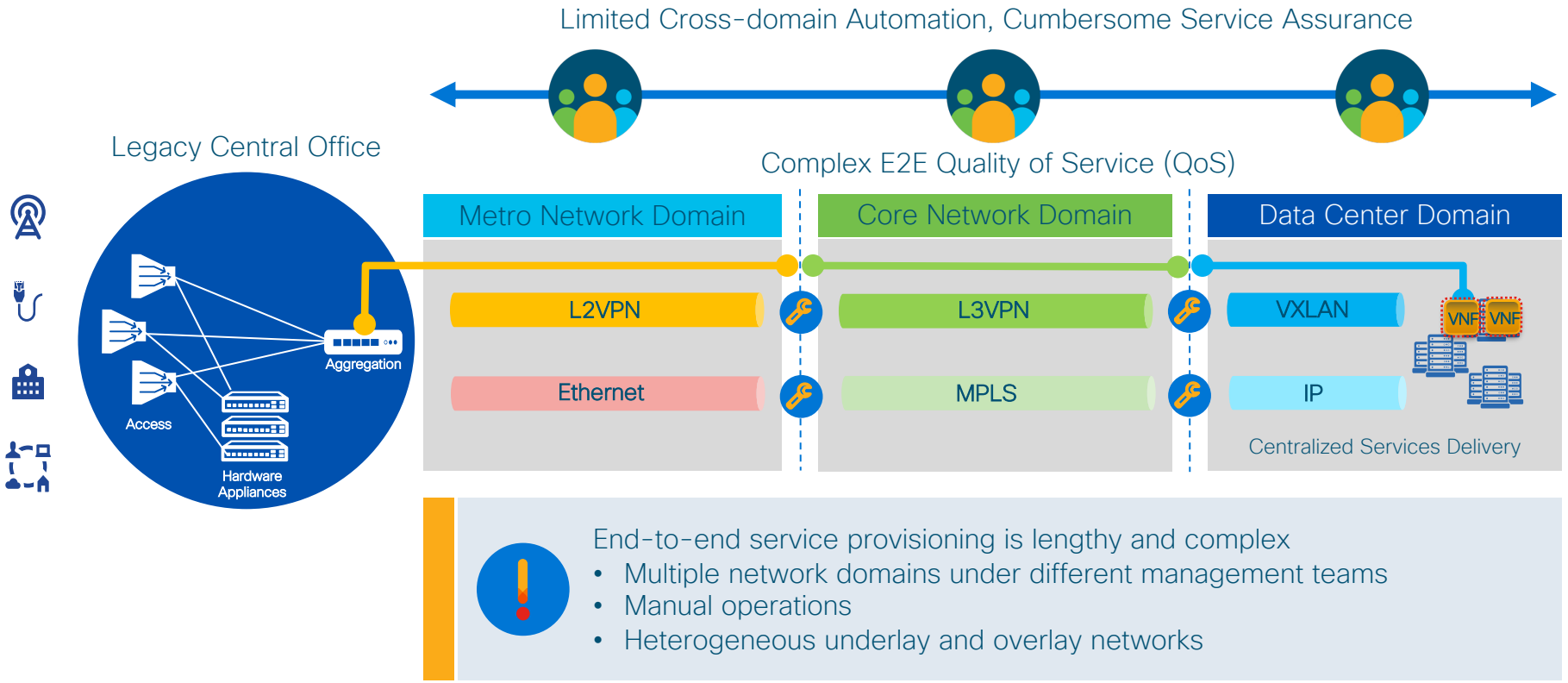
Agenda

- Introduction
- SRv6 uSID Dataplane
- SRv6 uSID Control Plane
- Flexible Algorithm
- SRv6 Addressing
- MPLS to SRv6 Migration
- Conclusion

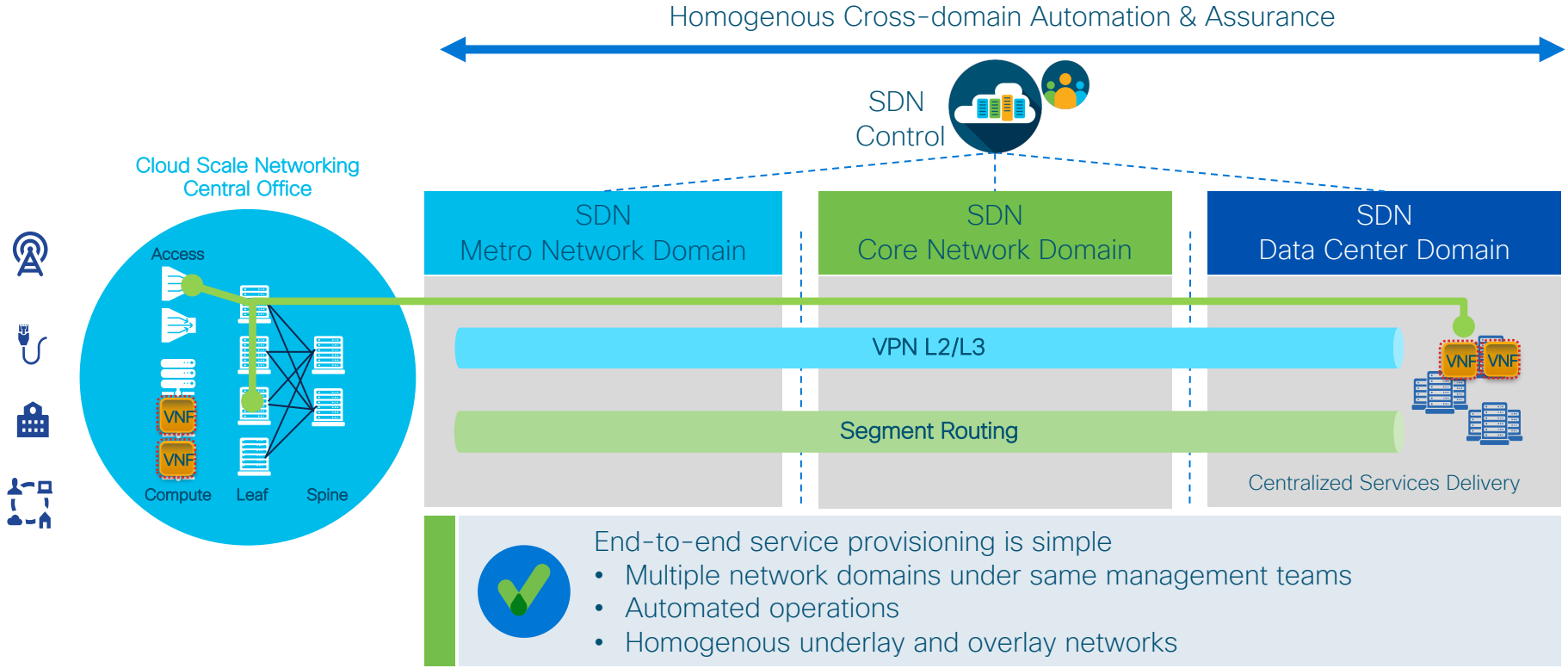
Introduction



Understanding Today's Service Creation



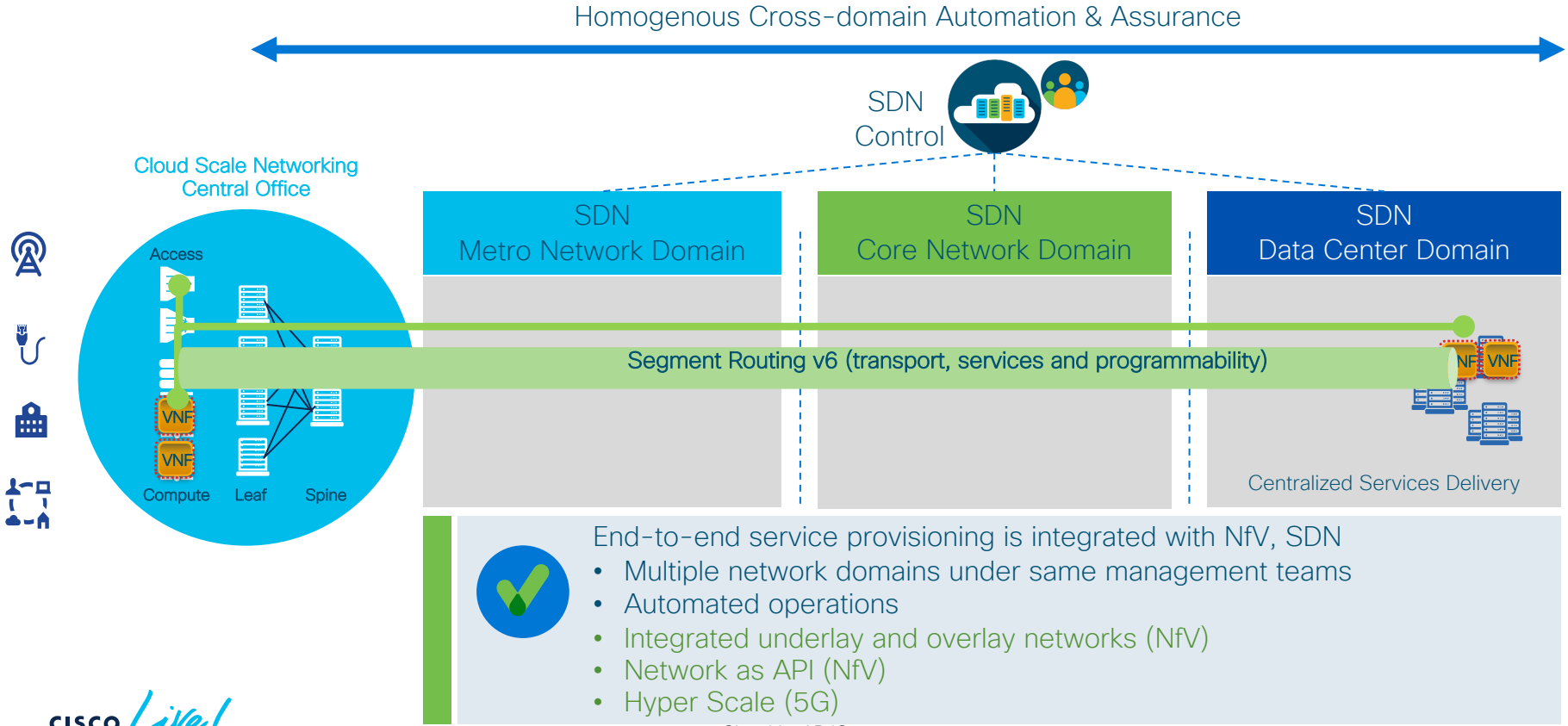
SR-MPLS: SDN ready
“Network as a Fabric” for Service Creation



End-to-end service provisioning is simple

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

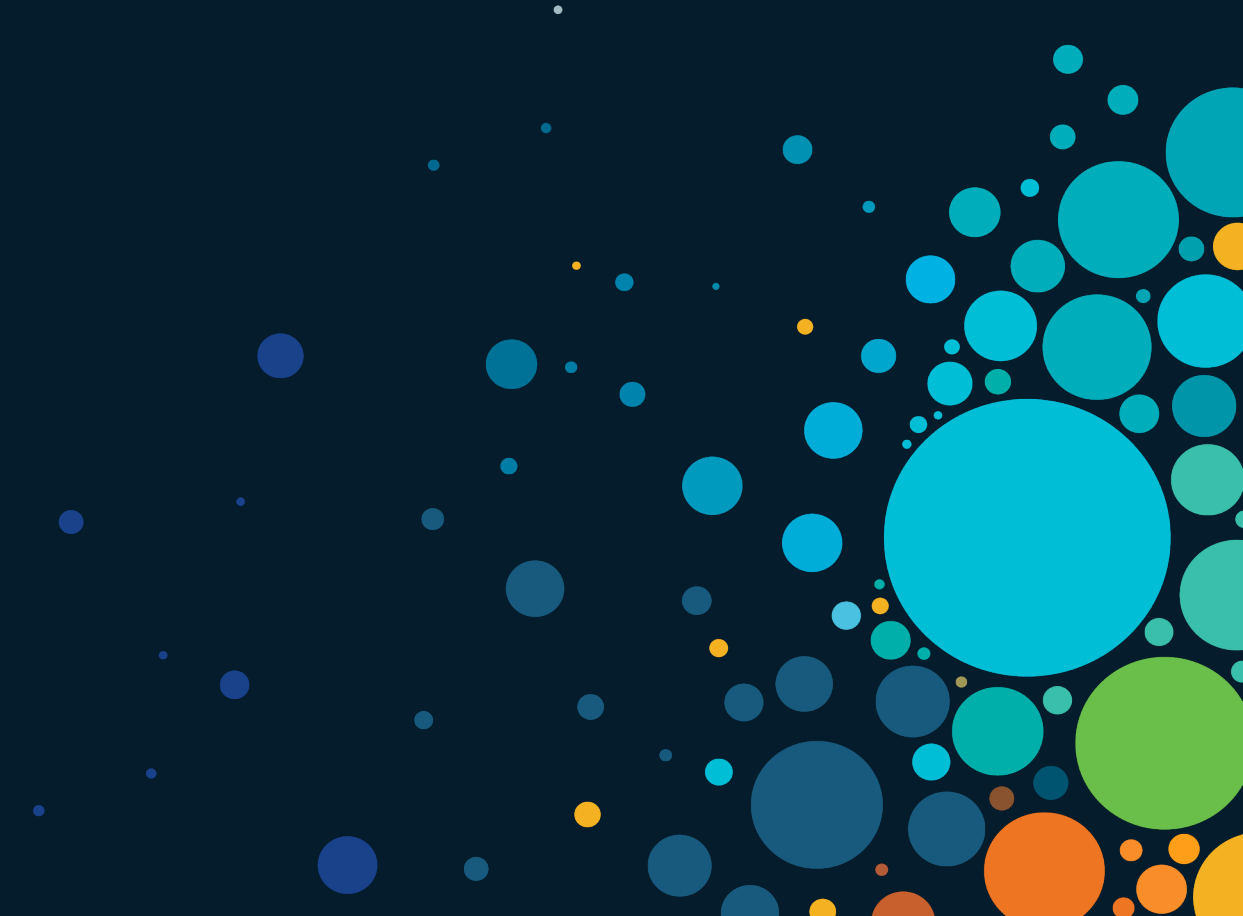
SRv6: SDN, NfV, 5G ready
“Network as an API” for Service Creation



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

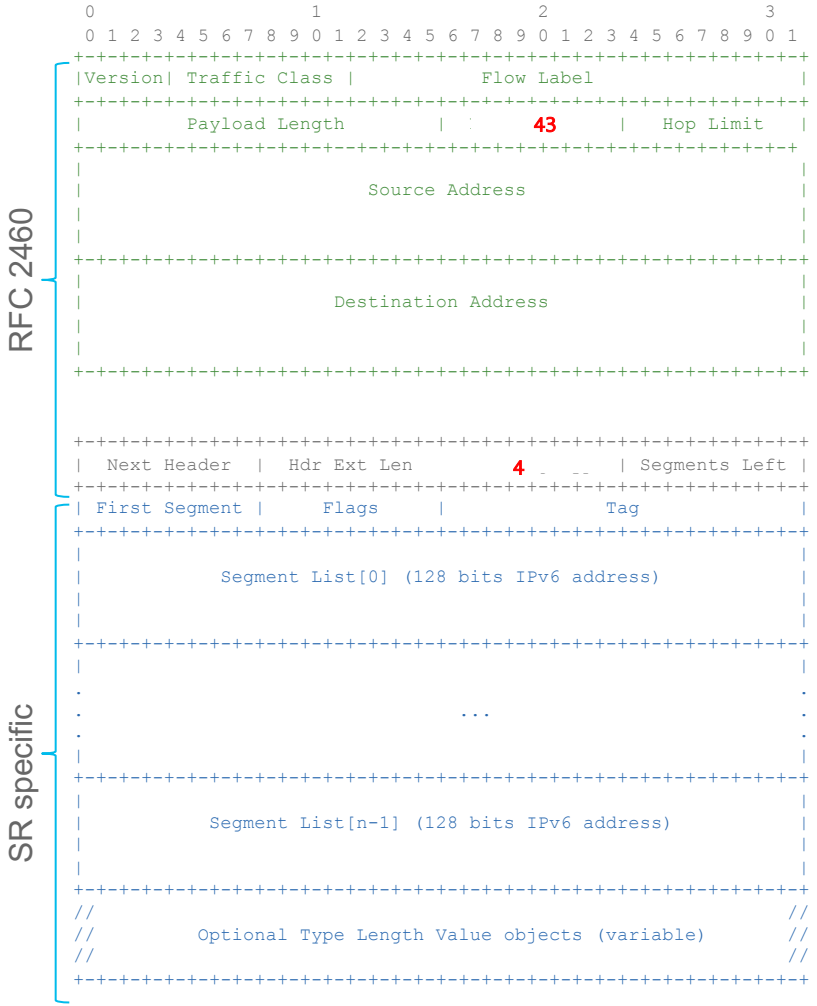
- Multiple network domains under same management teams
- Automated operations
- Integrated underlay and overlay networks (NfV)
- Network as API (NfV)
- Hyper Scale (5G)

Dataplane



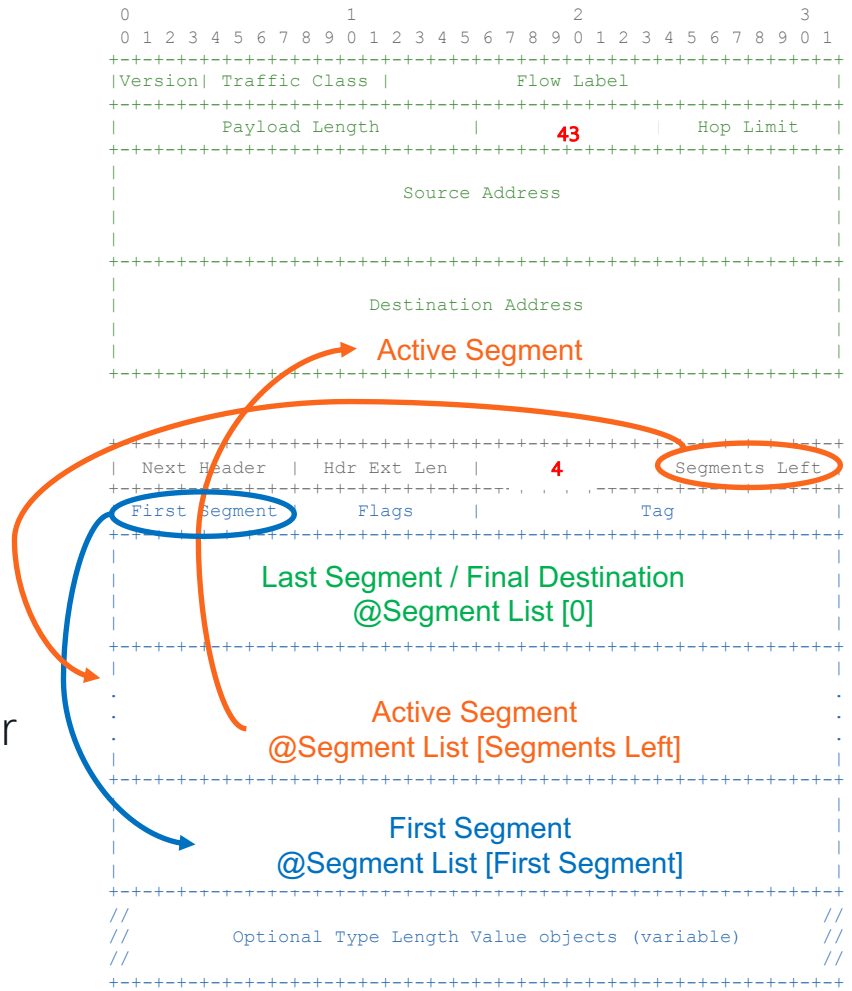
IPv6 SR Header

- IPv6 header
  - Next header field: **43 → Routing**
- IPv6 Routing extension header
  - Generic header format defined in RFC 2460
    - Next Header: IPv4, TCP, UDP, …
    - Hdr Ext Len: Any IPv6 device can skip this header
    - Segments Left: Ignore extension header if equal to 0
  - Specific data depends on Routing Type field:
    - ~~*0 Source Route*~~ (deprecated since 2007)
    - ~~*1 Nimrod*~~ (deprecated since 2009)
    - 2 Mobility (RFC 6275)
    - 3 RPL Source Route (RFC 6554)
    - **4 Segment Routing** (tentative)

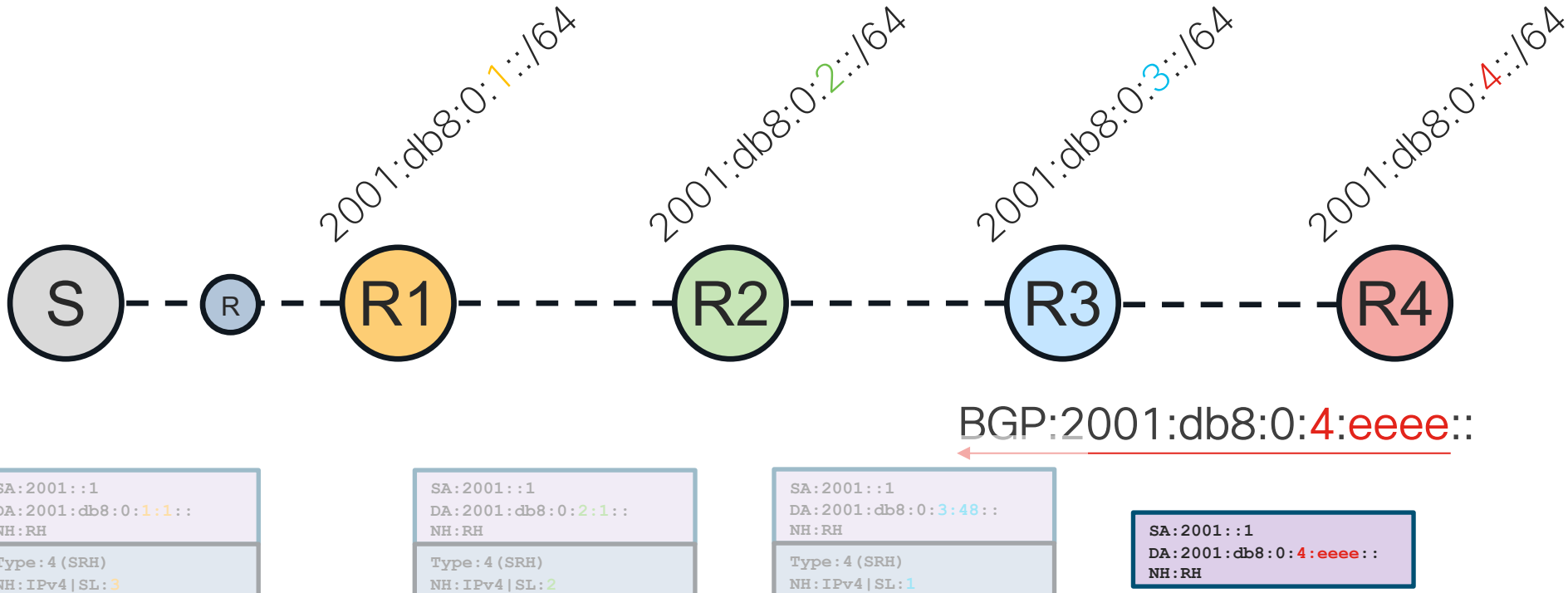


IPv6 SR Header

- Each segment is an IPv6 address
- Segments are encoded in reverse order
  - Last segment index is 0
  - First segment index is **First Segment**
  - Active segment index is **Segments Left**
- Active Segment is copied in the Destination Address field of the IP header
- Additional data can be stored in TLVs
  - Security (HMAC), NFV metadata, …

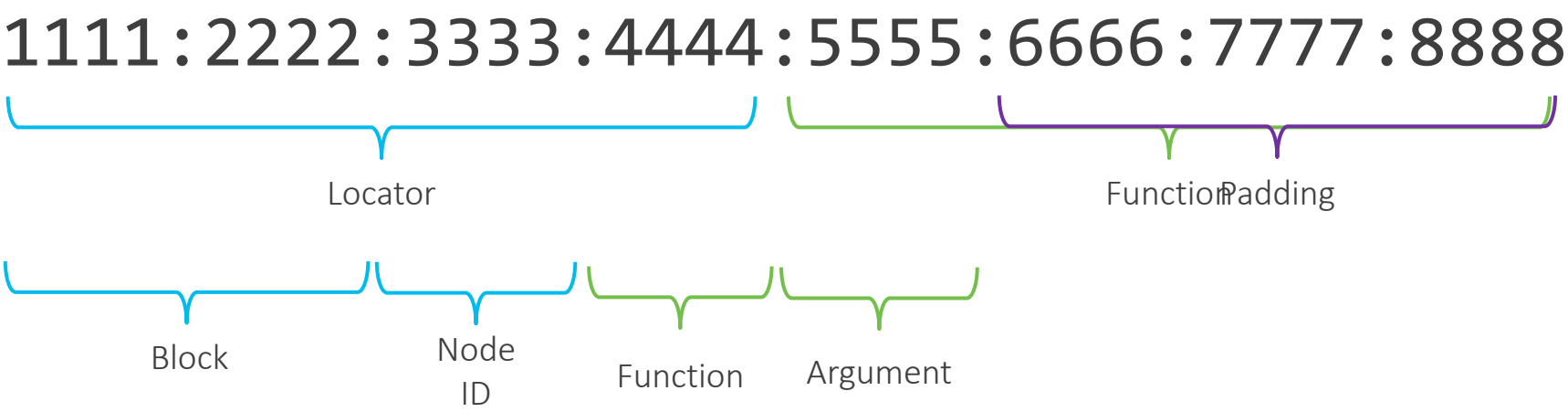


SRv6 Full SID



SID Structure

128 Bits, like IPv6 address but different semantics

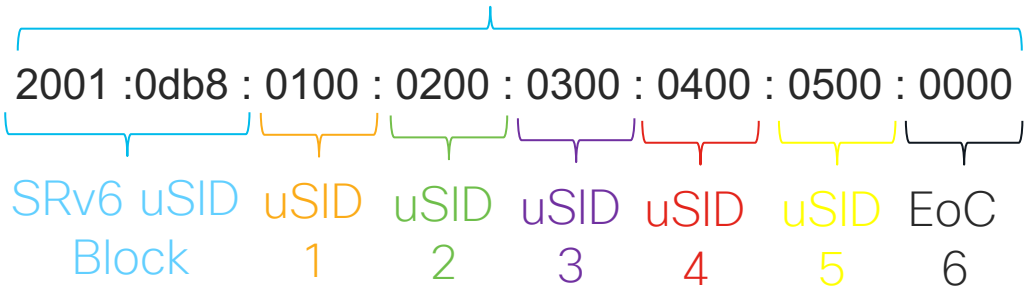


SRv6 uSID format

: 0100 : =SRV6 uSID

16 bits here, but can be anything

SRV6 uSID Carrier



32 bits here,
but can be anything

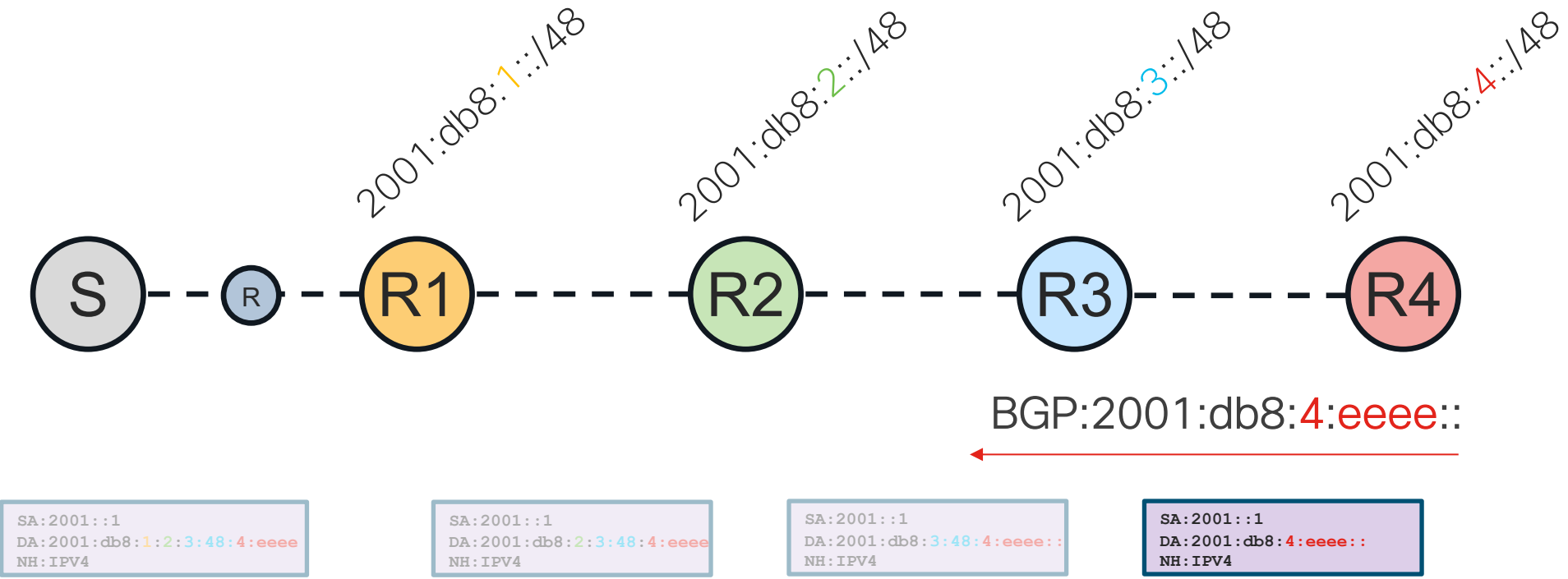
SRV6 Encapsulation

```
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
```

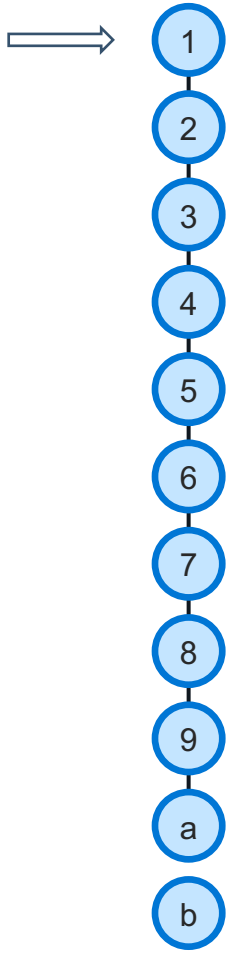
SRV6 uSID Encapsulation

```
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



Srv6 uSID More Than 6 SIDs?



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

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

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

```
SA:2001::1
DA:2001:db8:[illegible]
NH:RH
```

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

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

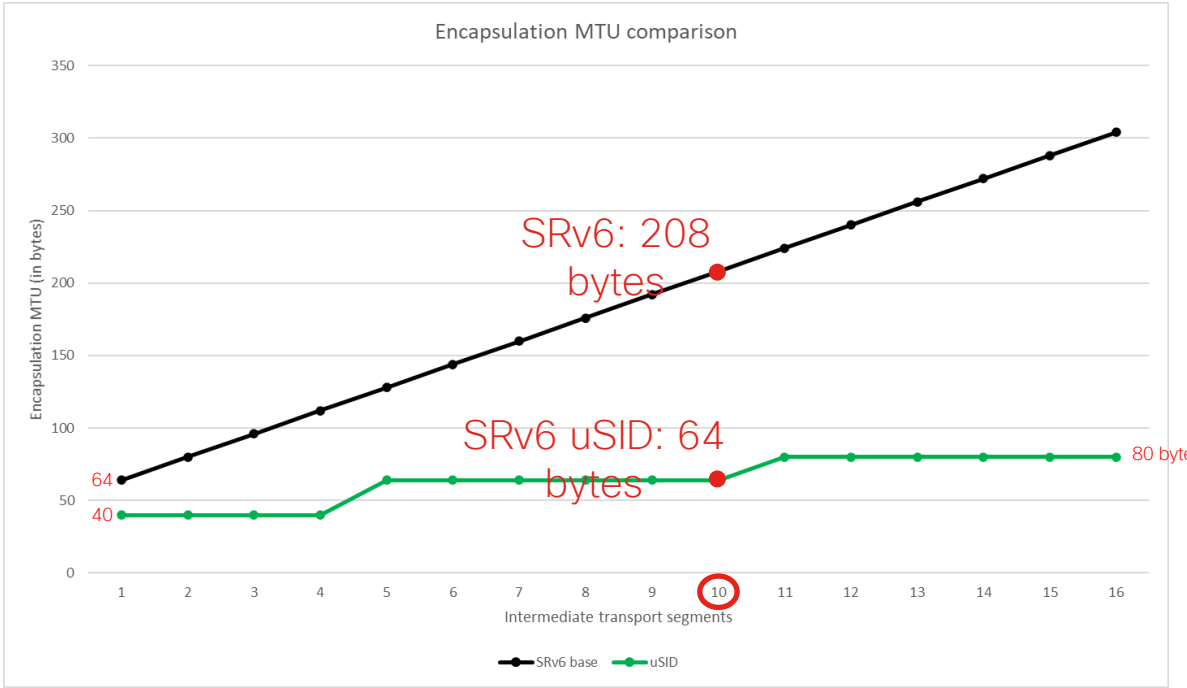
```
UDP Header/Data
```

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

The Power of SRv6 uSID

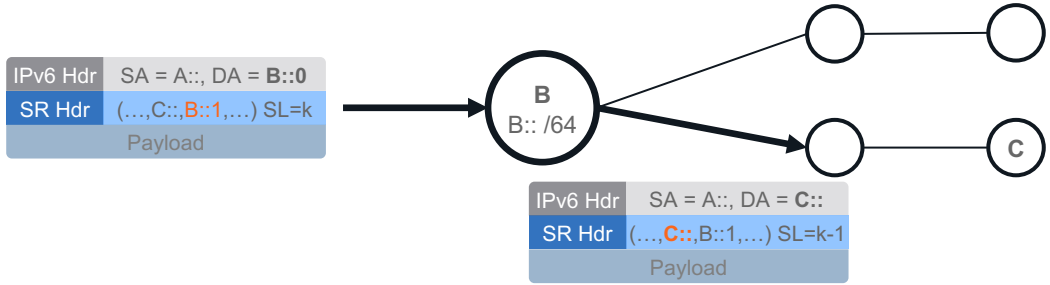
Best SRv6 Compression

- Realizing a network program with “n” intermediate segments …



END – Default endpoint

- *Default endpoint* behavior (node segment)
  - Decrement Segments Left, update DA
  - Forward according to new DA
- Node B advertises prefix B::/64 (B::/64 is the SID **locator**)
  - Packets are forwarded to B along the default routes (shortest path)
- On B, the *default endpoint* behavior is associated with ID 0 (0 is the **function**)
- The SID corresponding to the *default endpoint* behavior on node B is B::1

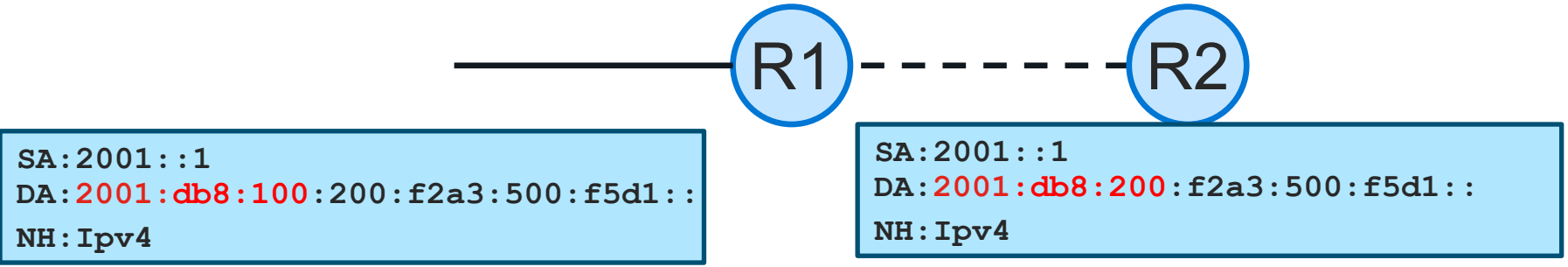


uN: shortest-path to a Node

- Global uSID that represents the shortest path to a node N
  - 2001:db8:0N00::
- Shift&Forward

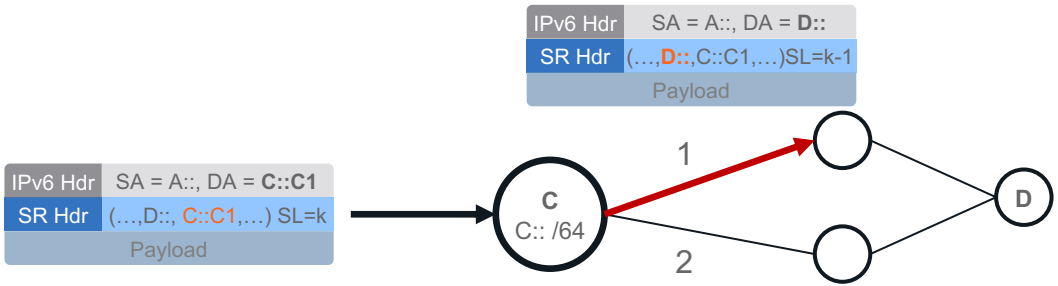
Just Like END

If EoC -> then it is exactly END example: 2001:db8:200:0::



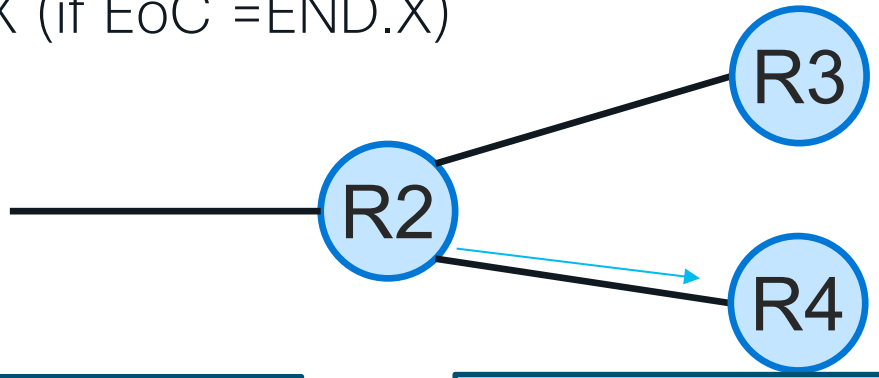
END.X – Endpoint then Xconnect

- *Endpoint xconnect* behavior (adjacency segment)
  - Decrement Segments Left, update DA
  - Forward on the interface associated with the Xconnect segment
- Node C advertises prefix C::/64
  - Packets are forwarded to C along the default routes (shortest path)
- On C, the *endpoint xconnect* behavior for link (C, E) is associated with ID CE
- The SID corresponding to *endpoint xconnect-(C,E)* behavior on node C is C::CE



uA: xconnect into L3 adjacency

- Local uSID that represents the L3 adjacency from node N to J
  - 2001:db8:FNAJ::
- Shift&Forward to xconnect
- Just Like END.X (if EoC =END.X)




```
SA:2001::1
DA:2001:db8:200:f2a4:500:f5d1::
NH:Ipv4
```

```
SA:2001::1
DA:2001:db8:500:f5d1::
NH:Ipv4
```

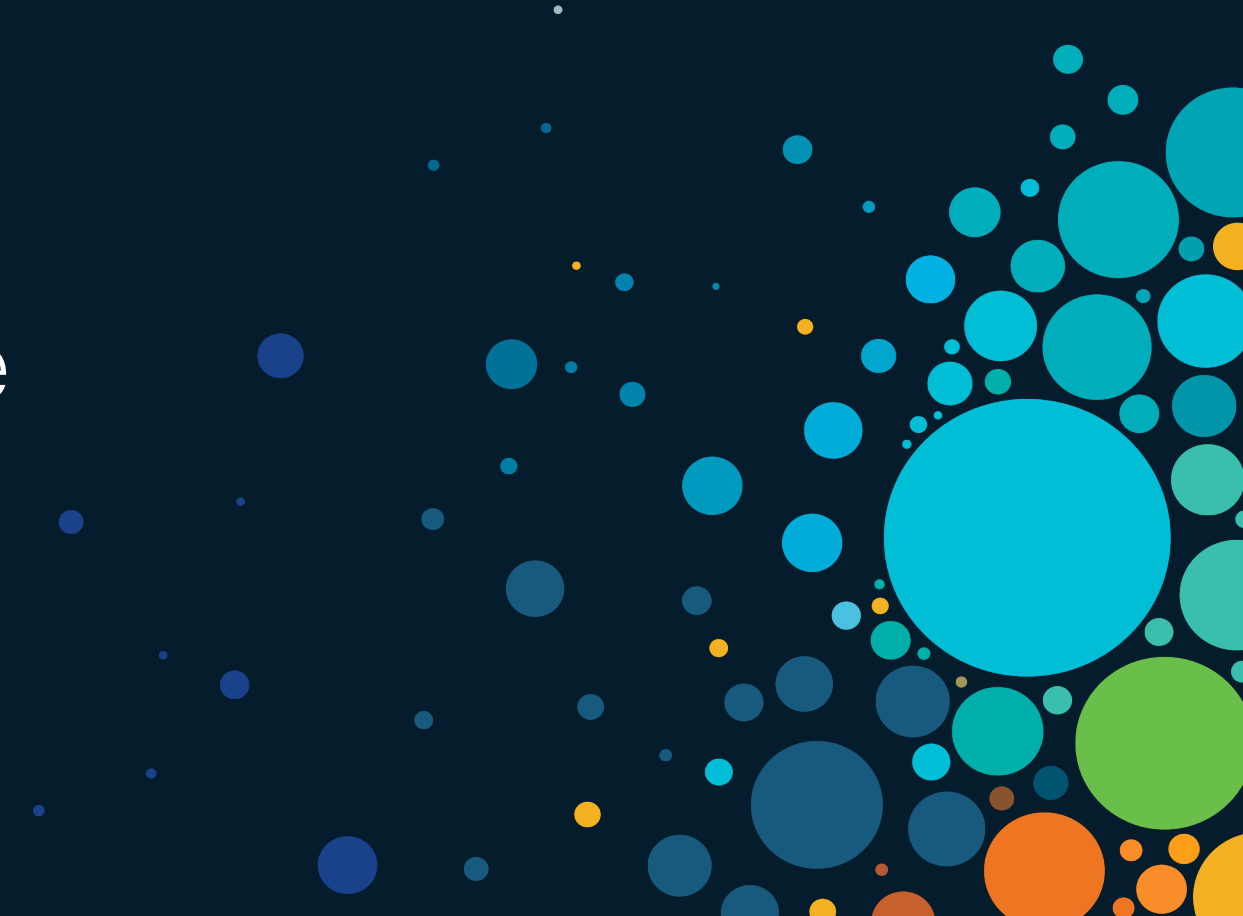
END.DX4 – Endpoint with Decapsulation and Xconnect (END.DX6, END.DX2)

- *Endpoint xconnect* behavior (adjacency segment)
  - Segments Left must be 0
  - NH must be IPv4 (or IPv6 or L2)
  - Decapsulate inner packet
  - Forward on the interface associated with the Xconnect
- Node D advertises prefix C::/64
  - Packets are forwarded to D along the default routes (shortest path)
- On C, the *endpoint xconnect* behavior for link (D, E) is associated with ID DE
- It is like L3 VPN with per CE label allocation

END.DT4 – Endpoint with Decapsulation and Table lookup(END.DT6)

- *Endpoint xconnect* behavior (adjacency segment)
  - Segments Left must be 0
  - NH must be IPv4 (or IPv6)
  - Decapsulate inner packet
  - Do the lookup for IPv4 destination of inner packet and forward accordingly
- Node D advertises prefix D::/64
  - Packets are forwarded to D along the default routes (shortest path)
- On C, the *endpoint* behavior for link (D, E) is associated with ID CE
- It is like L3 VPN with per VRF label allocation

Control Plane



Functions might be signaled differently

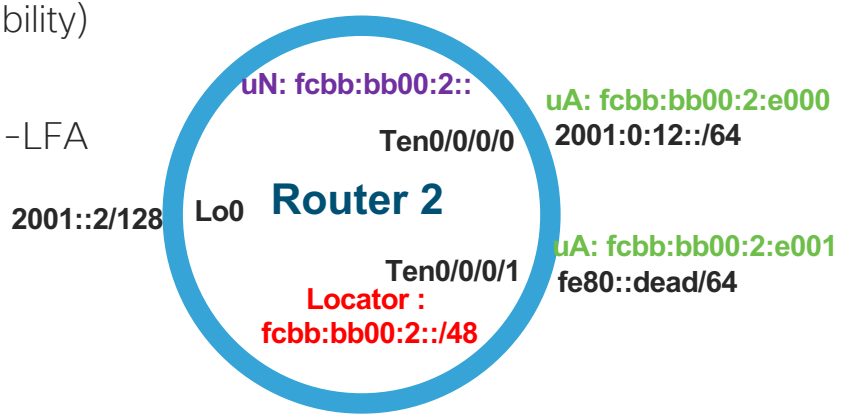
Signalling	IGP	BGP-LS	BGP-IP/VPN
End, uN	Yes	Yes	
End.X, uA	Yes	Yes	
End.T	Yes	Yes	
End.DX4,uDX4		Yes	Yes
End.DX6,uDX6	Yes	Yes	Yes
End.DX2,uDX2		Yes	Yes
END.DT4,uDT4		Yes	Yes
End.DT6,uDT6	Yes	Yes	Yes
End.B		Yes	

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

Locator – routing table

IGP for uSID

- Uses TLVs
- For Srv6:
  - Locator – for Reachability (twice for backward compatibility)
  - END function – TI-LFA and TE
  - END.X function for each interface in routing protocol TI-LFA and TE
  - Capabilities:
  - Max SID depth for different functions
- OSPF will follow



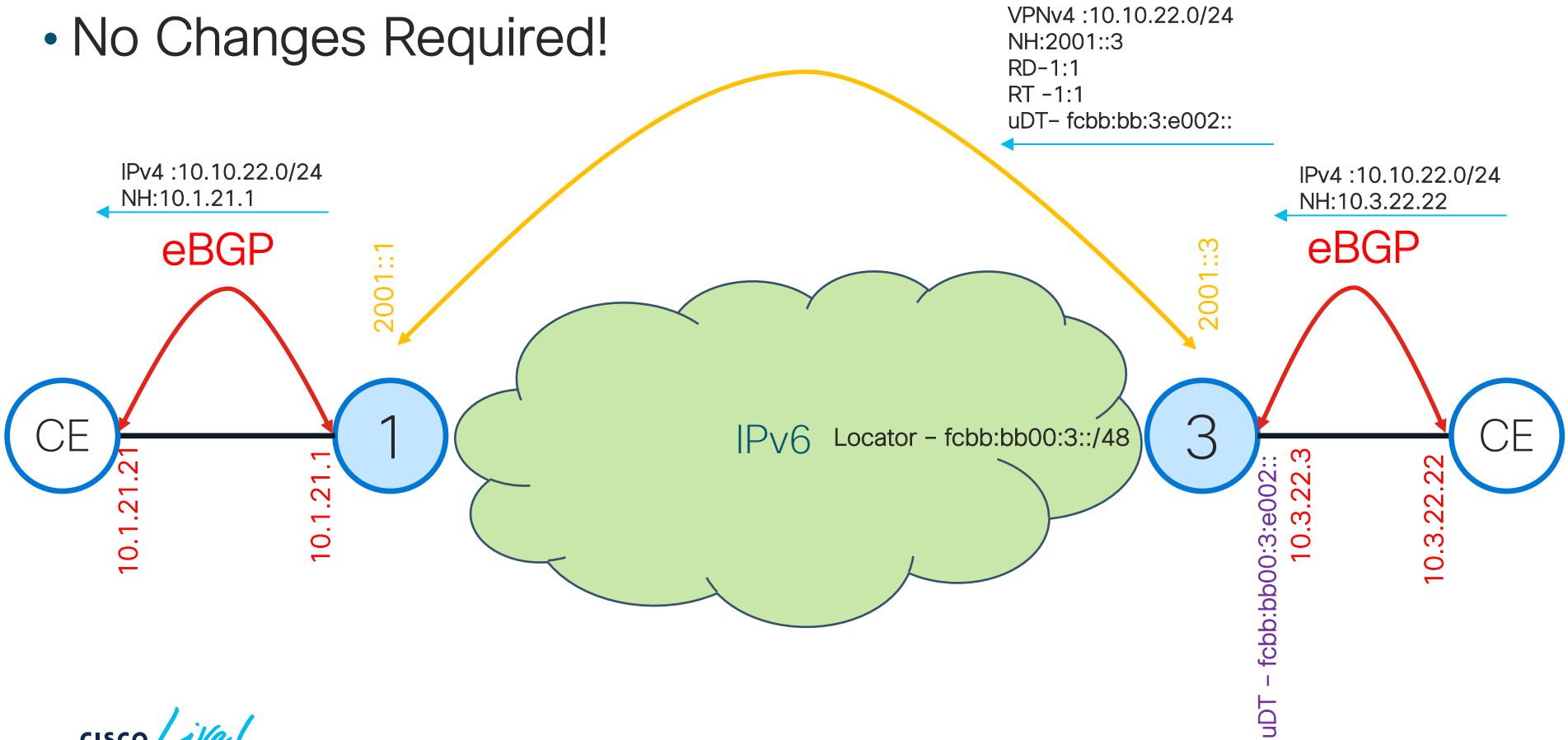
ISIS example

```
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 2001::2/128
    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 r2.00
    Local Interface ID: 6, Remote Interface ID: 6
    Interface IPv6 Address: 2001:0:0:12::1
    Neighbor IPv6 Address: 2001:0:0:12::2
    END.X SID: fcbb:bb00:2:e000:: 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
```

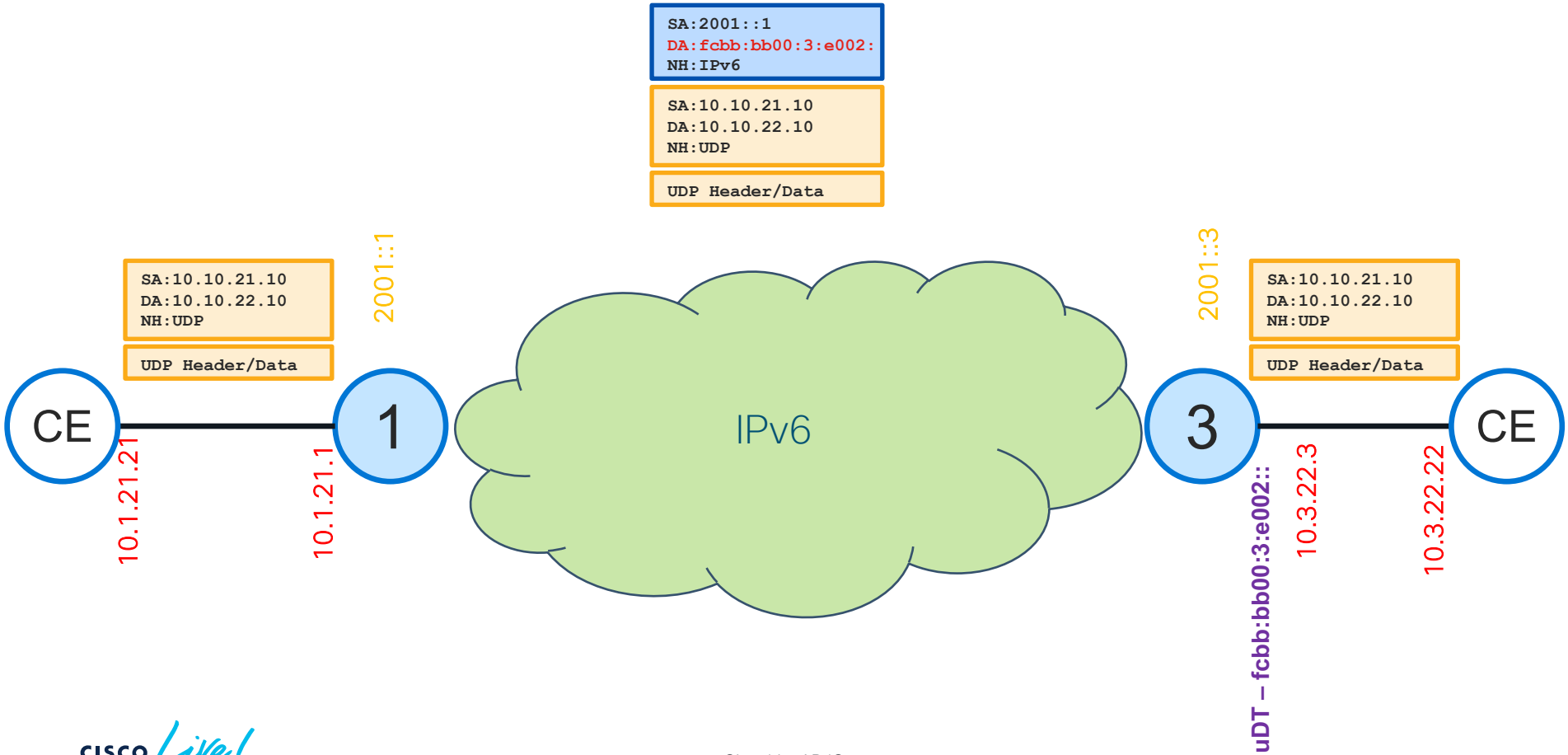
Locator
Capabilities
END
END.X
SID Structure

BGP

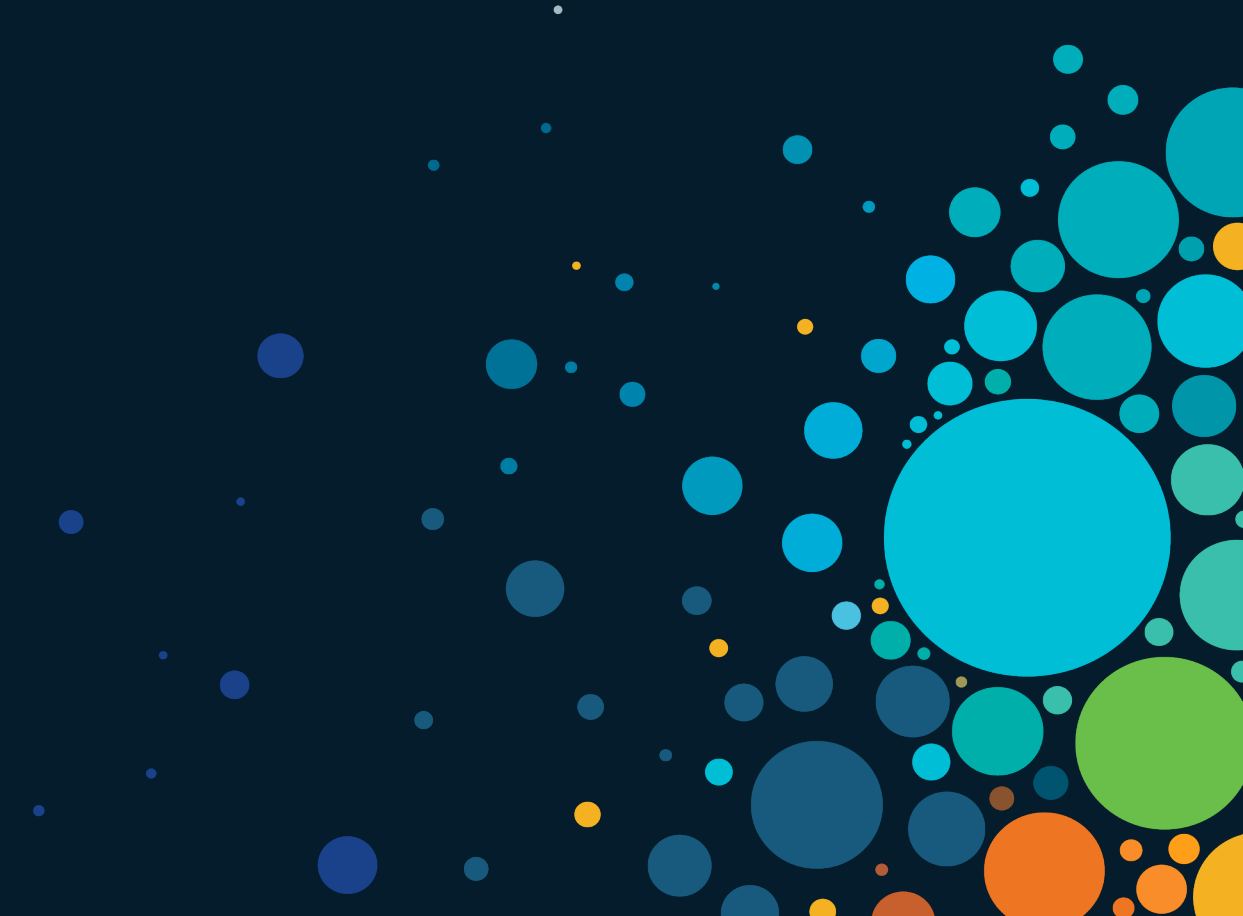
- No Changes Required!



BGP



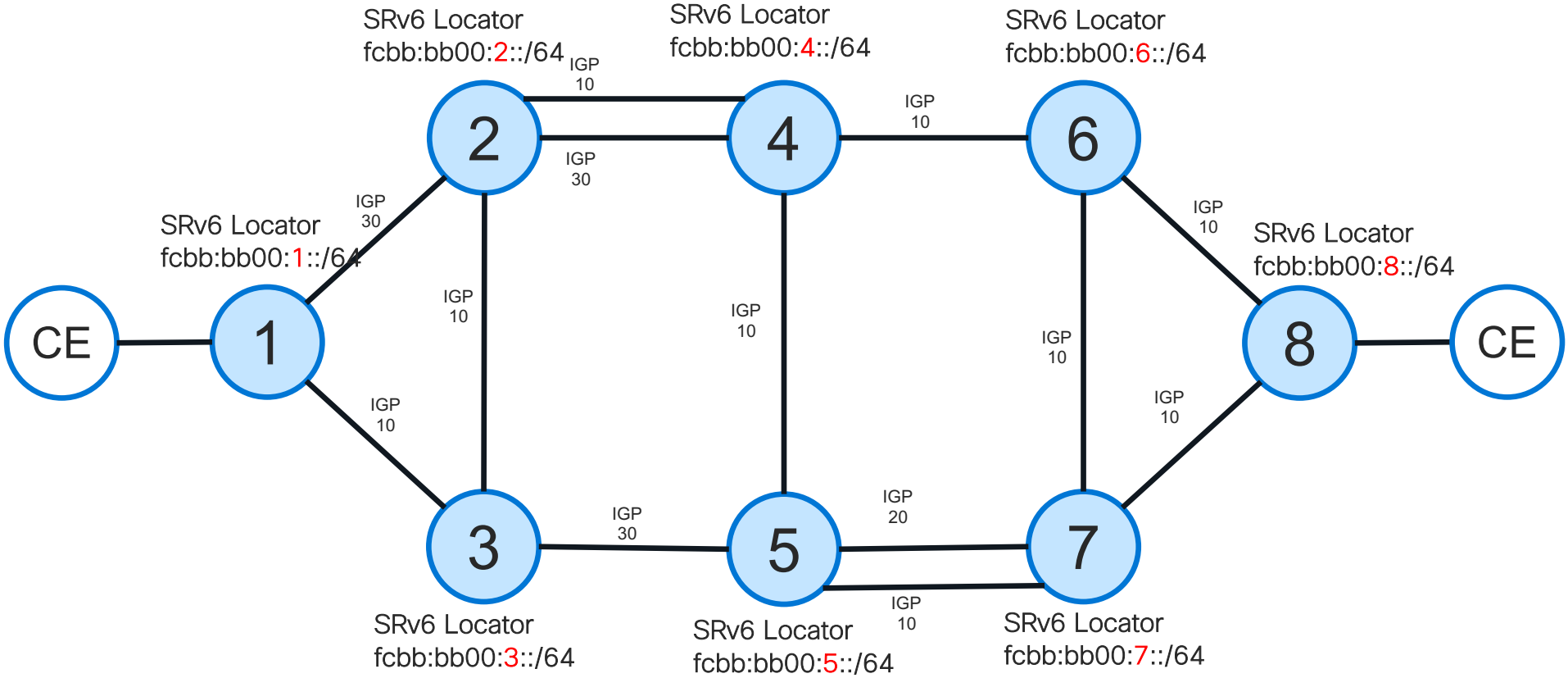
Flexible Algorithm



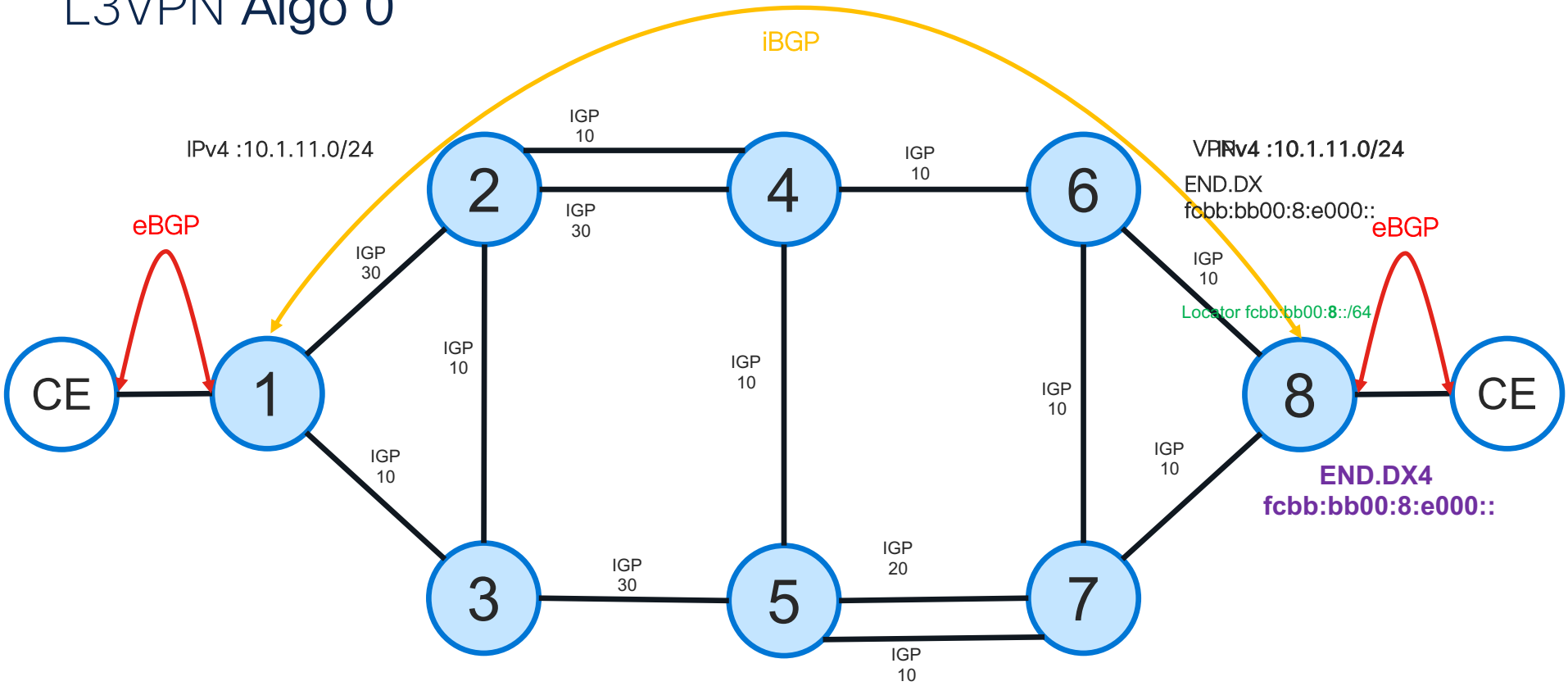
Flexible Algorithm

- 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”

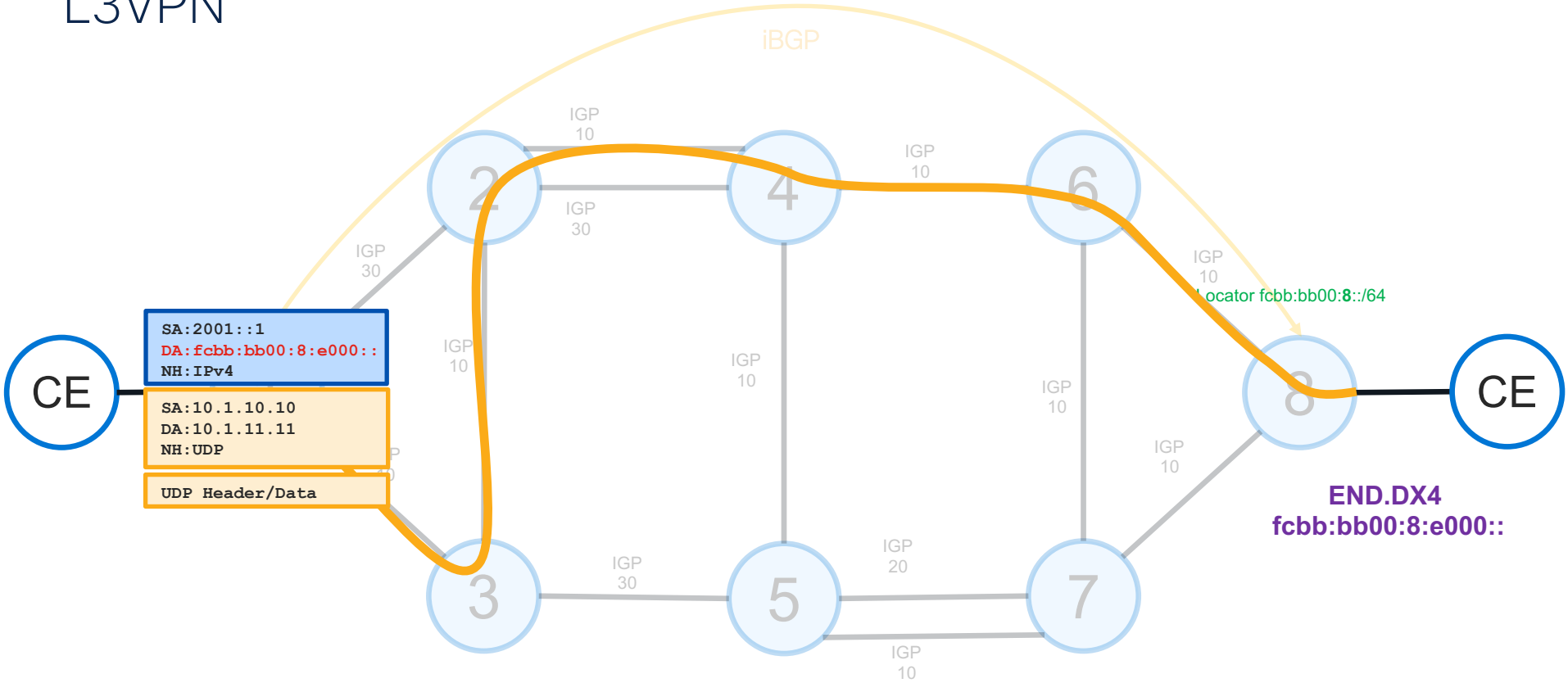
IPv6 Only –SRv6



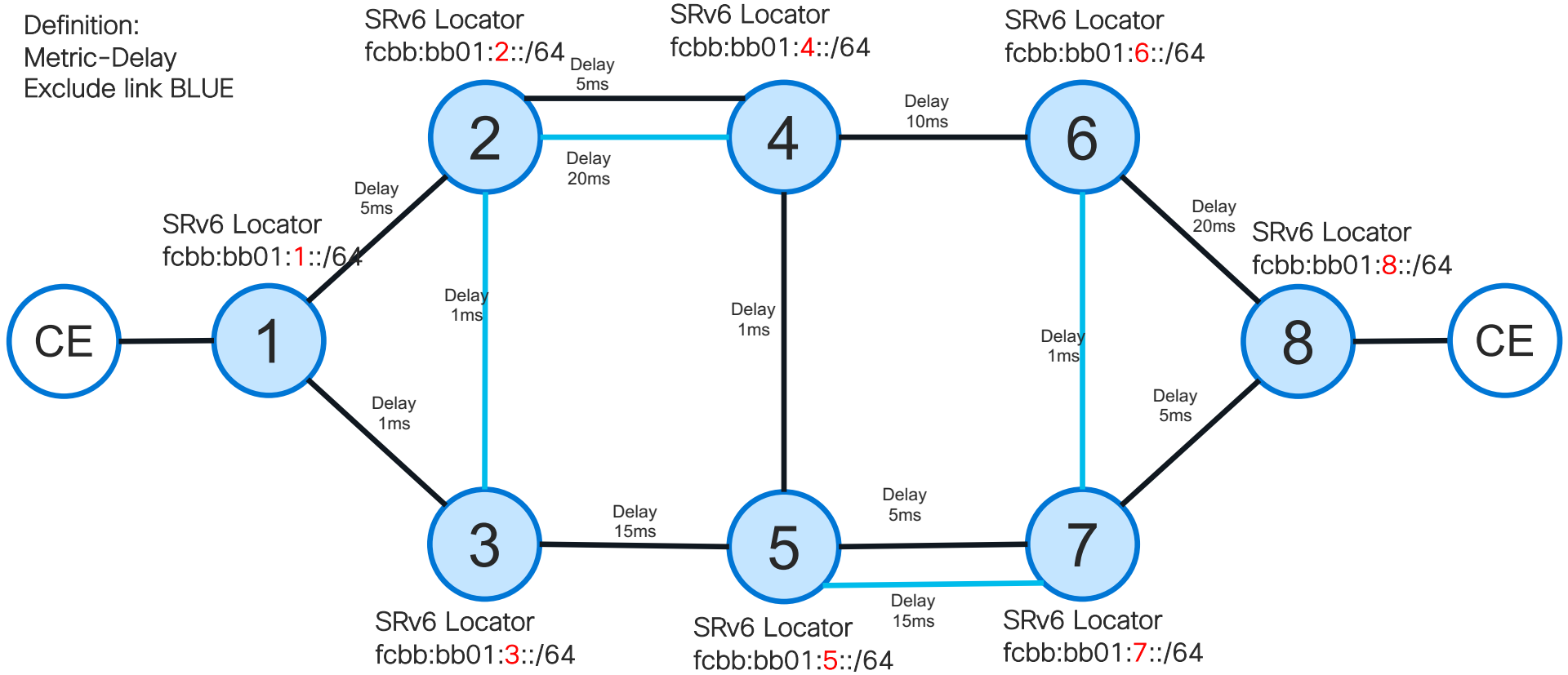
L3VPN Algo 0



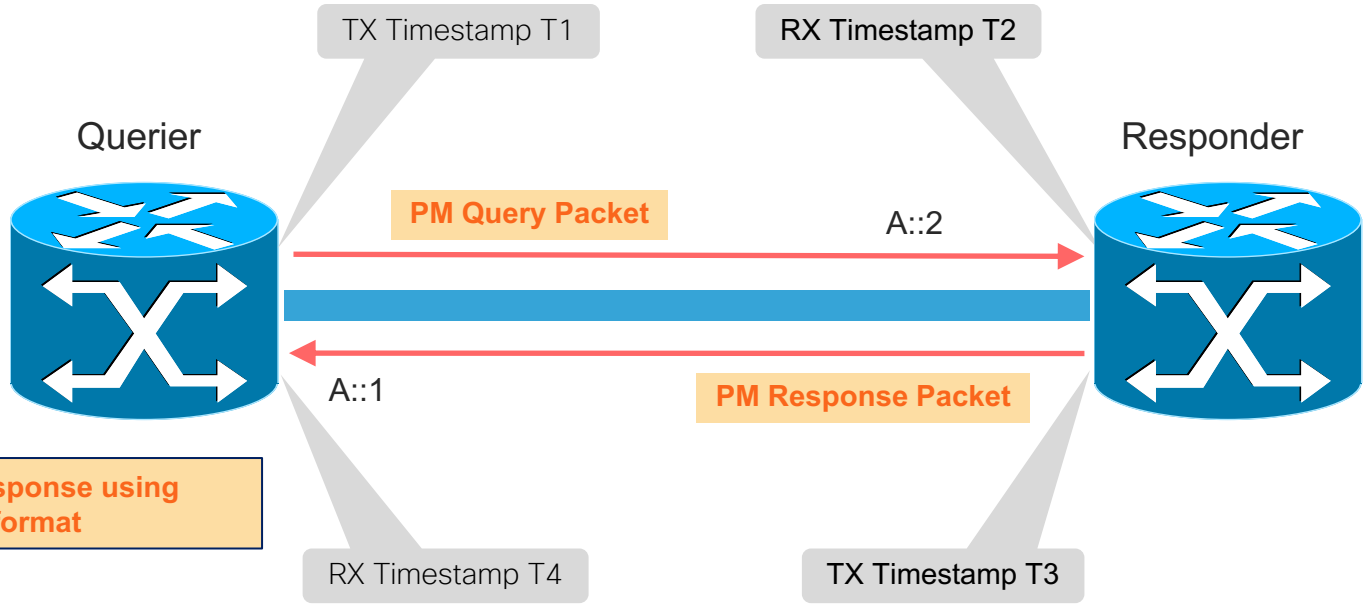
L3VPN



Flex ALGO 128



Link Delay Measurement Protocol

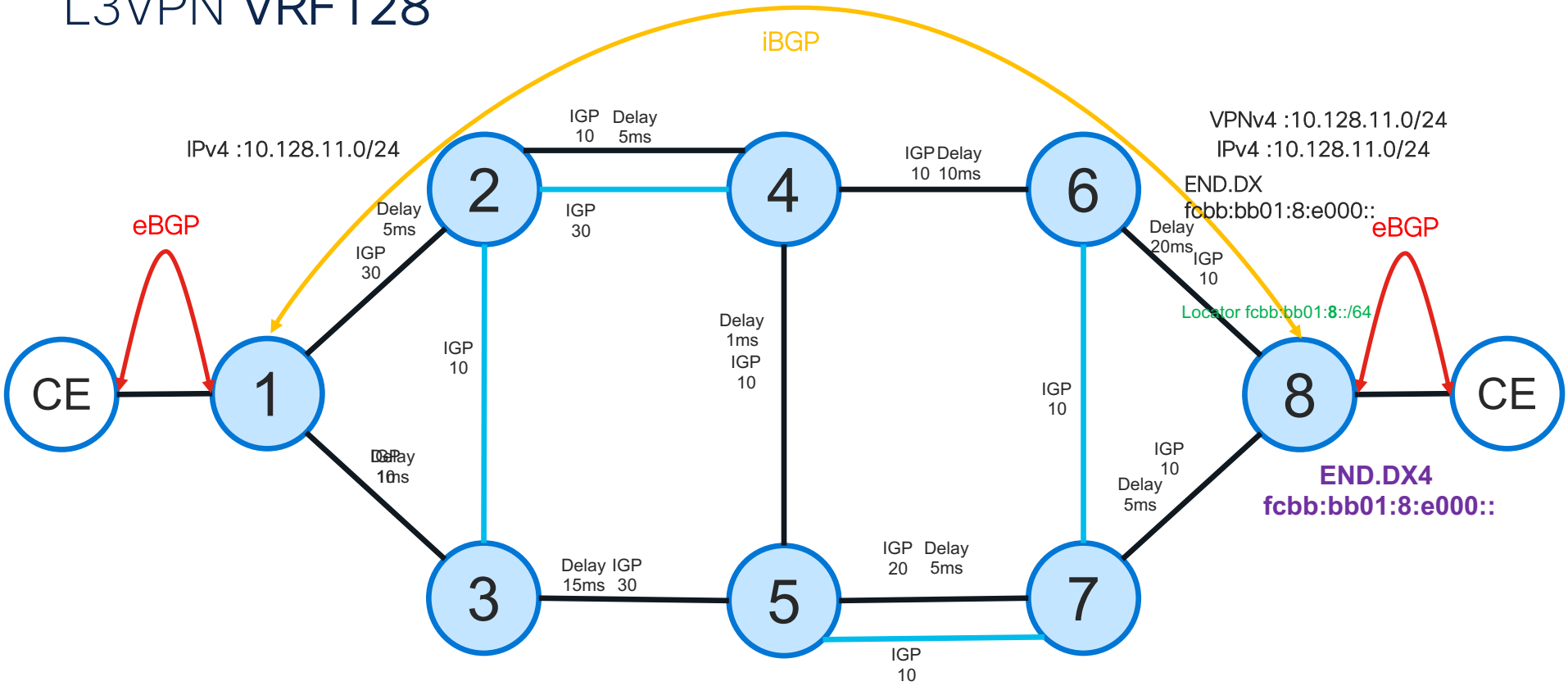


PM Query and Response using RFC 6374 packet format

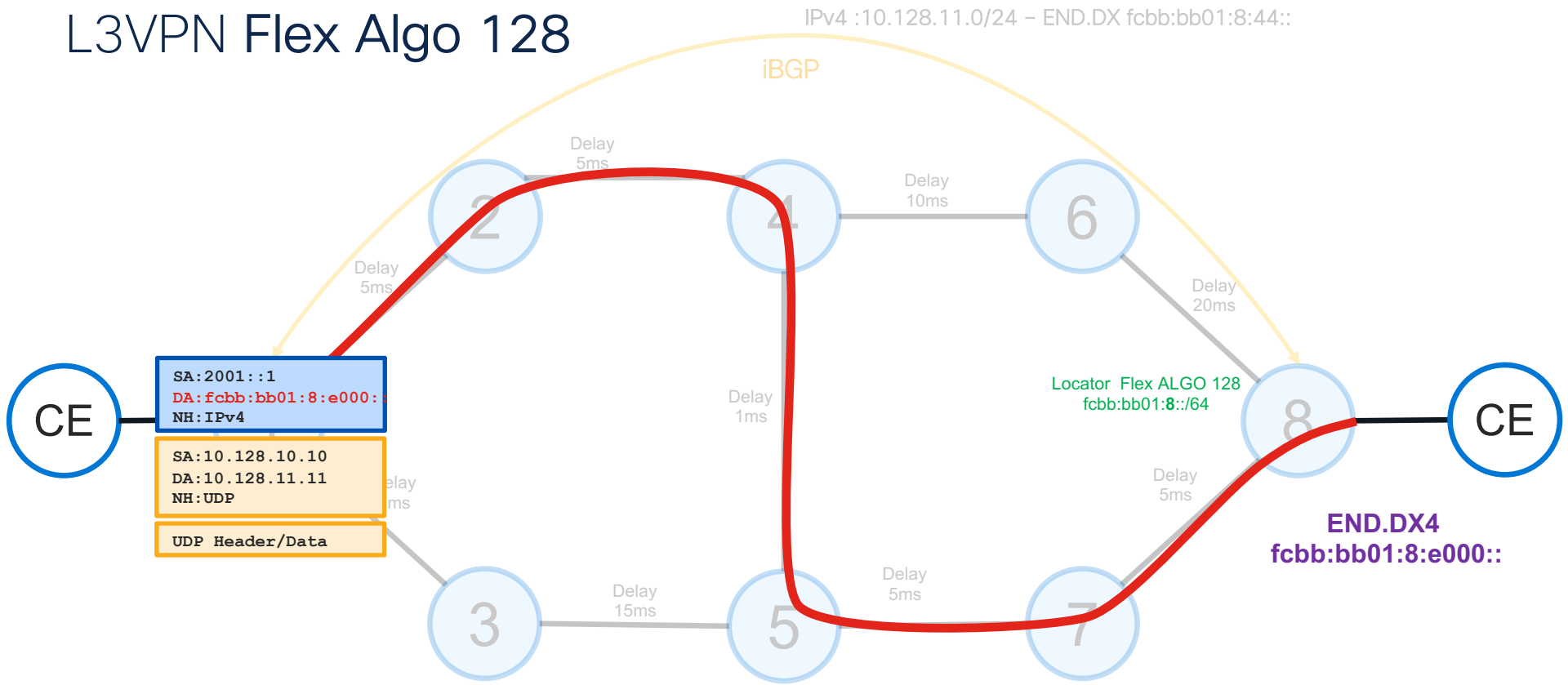
One-Way Delay = (T2 – T1)
Two-Way Delay = (T4 – T1) – (T3 – T2)
One-Way Delay = Two-Way Delay/2

- For one-way delay measurement, hardware clocks must be synchronized using PTP (IEEE 1588) between querier and responder nodes.

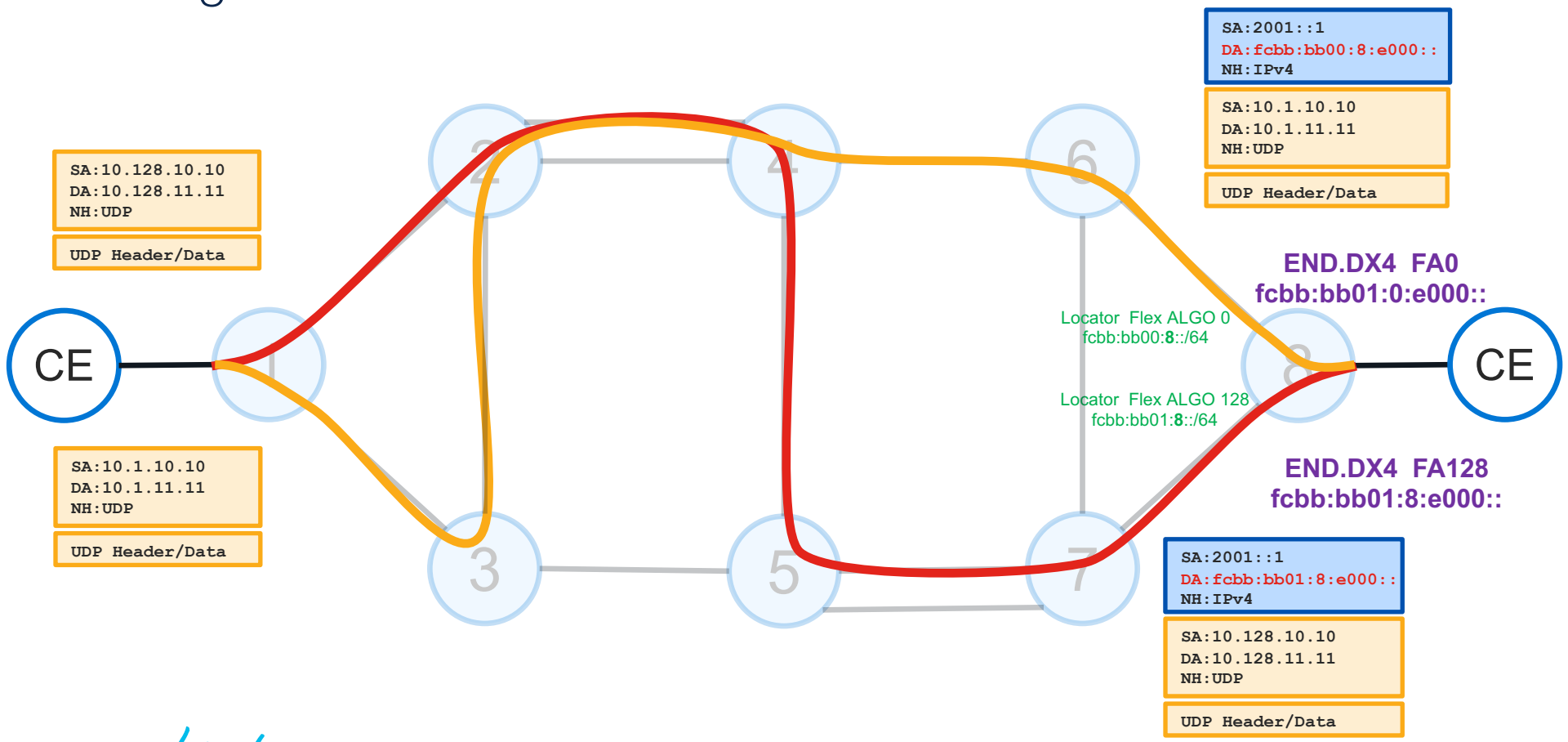
L3VPN VRF128



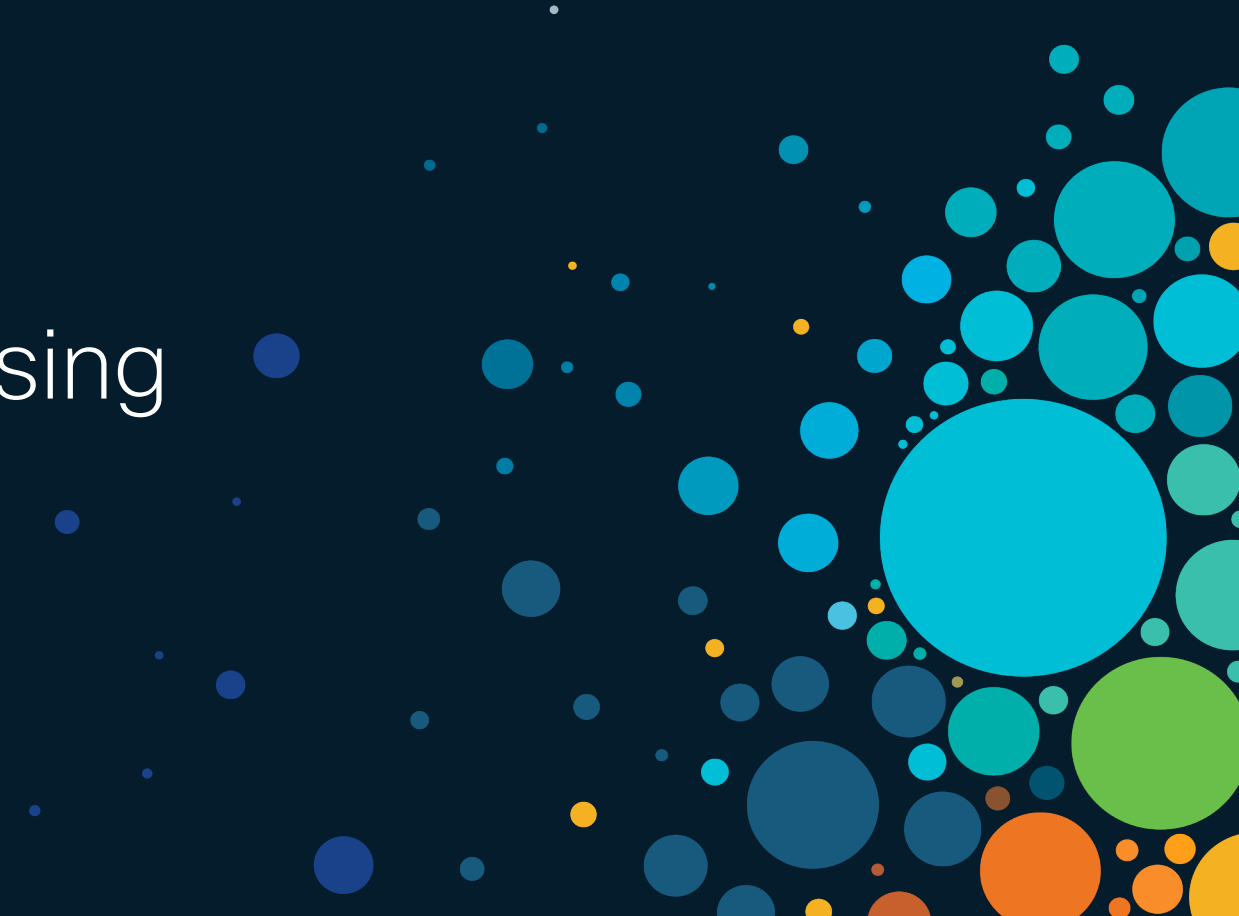
L3VPN Flex Algo 128



All Together



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.

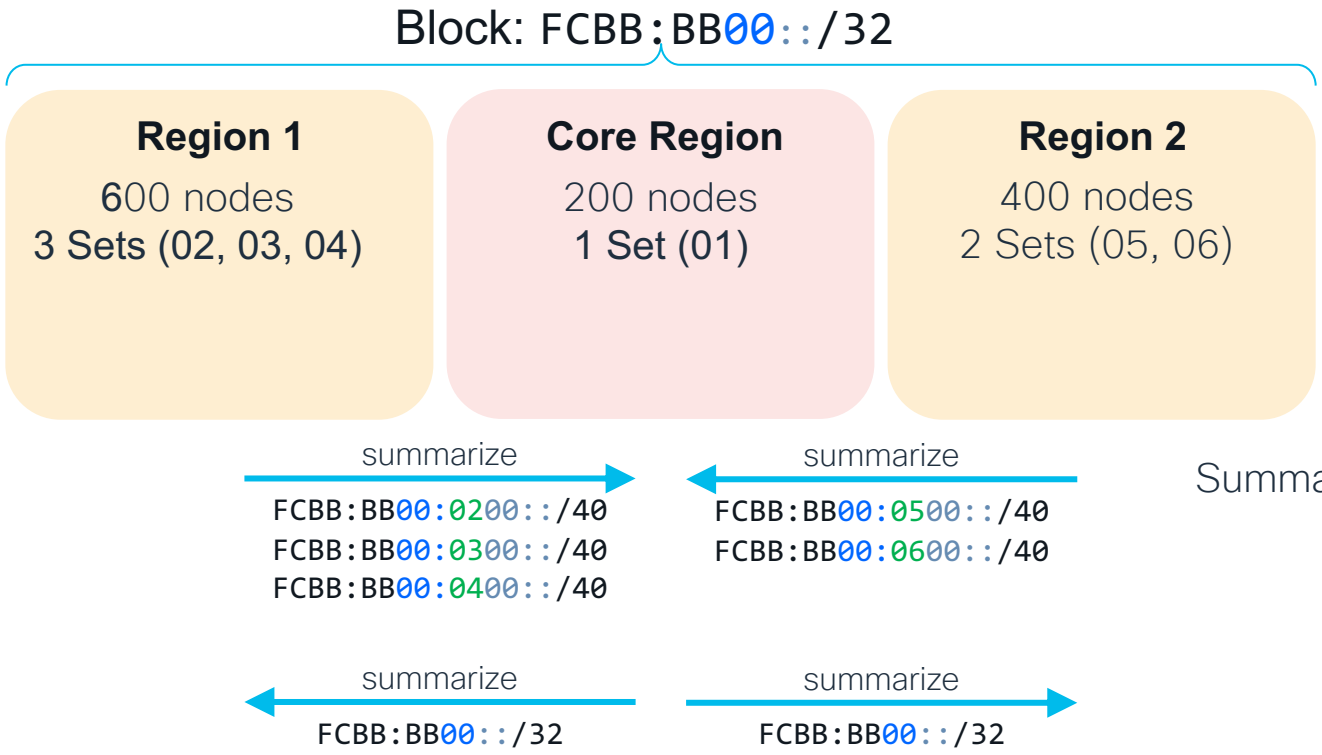
An existing IPv6 address plan is not a constraint
for a future SRv6 SID allocation plan.

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

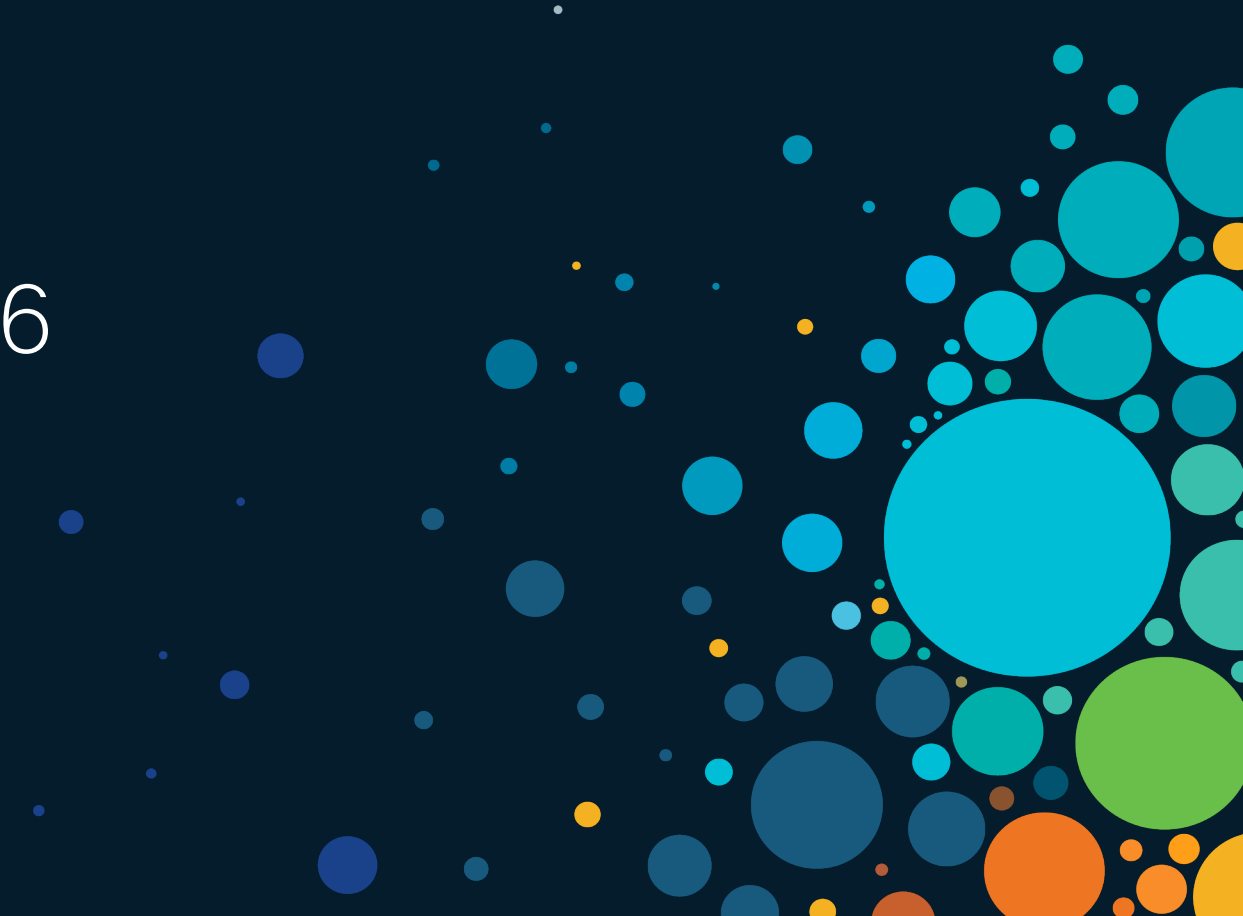
- Public range allocation
  - Supported, not advised
  - From allocated public GUA range

Summarization



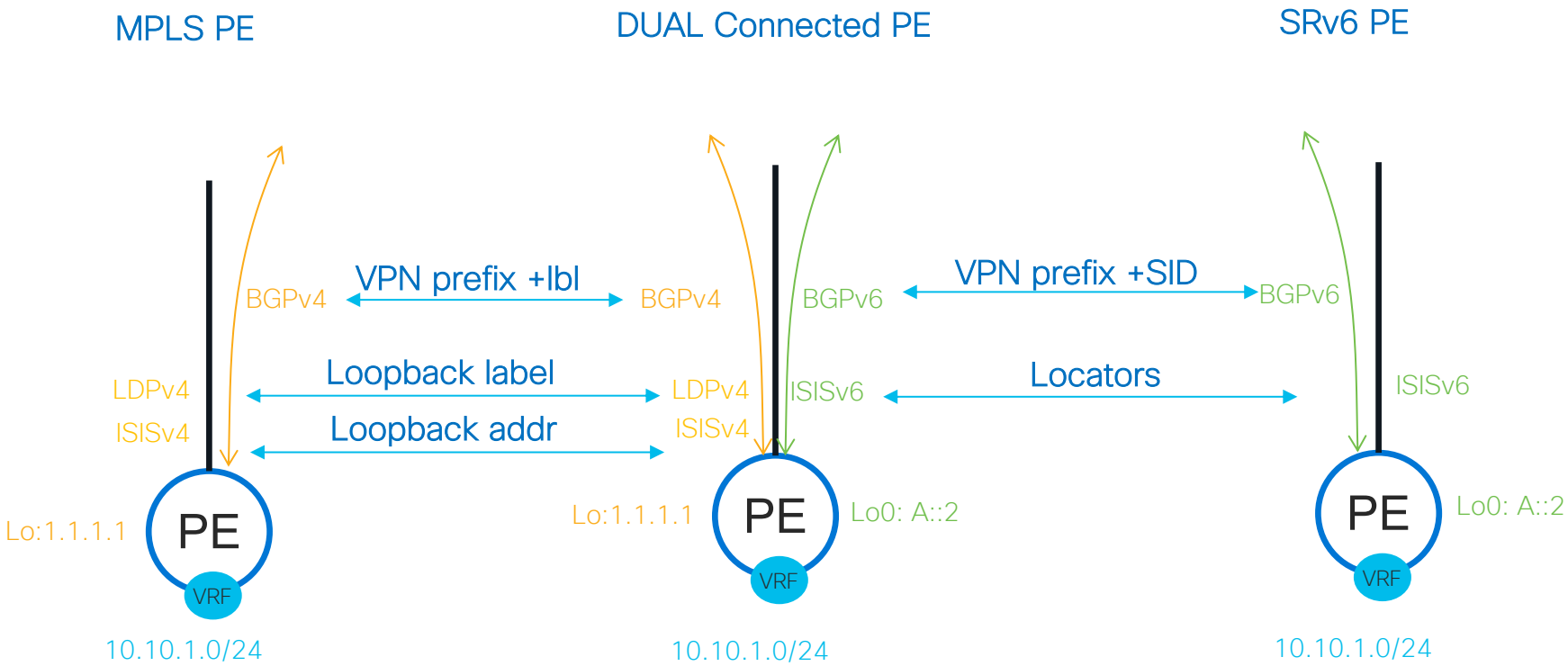
Summarization gain:
× 256

MPLS to SRv6 Migration

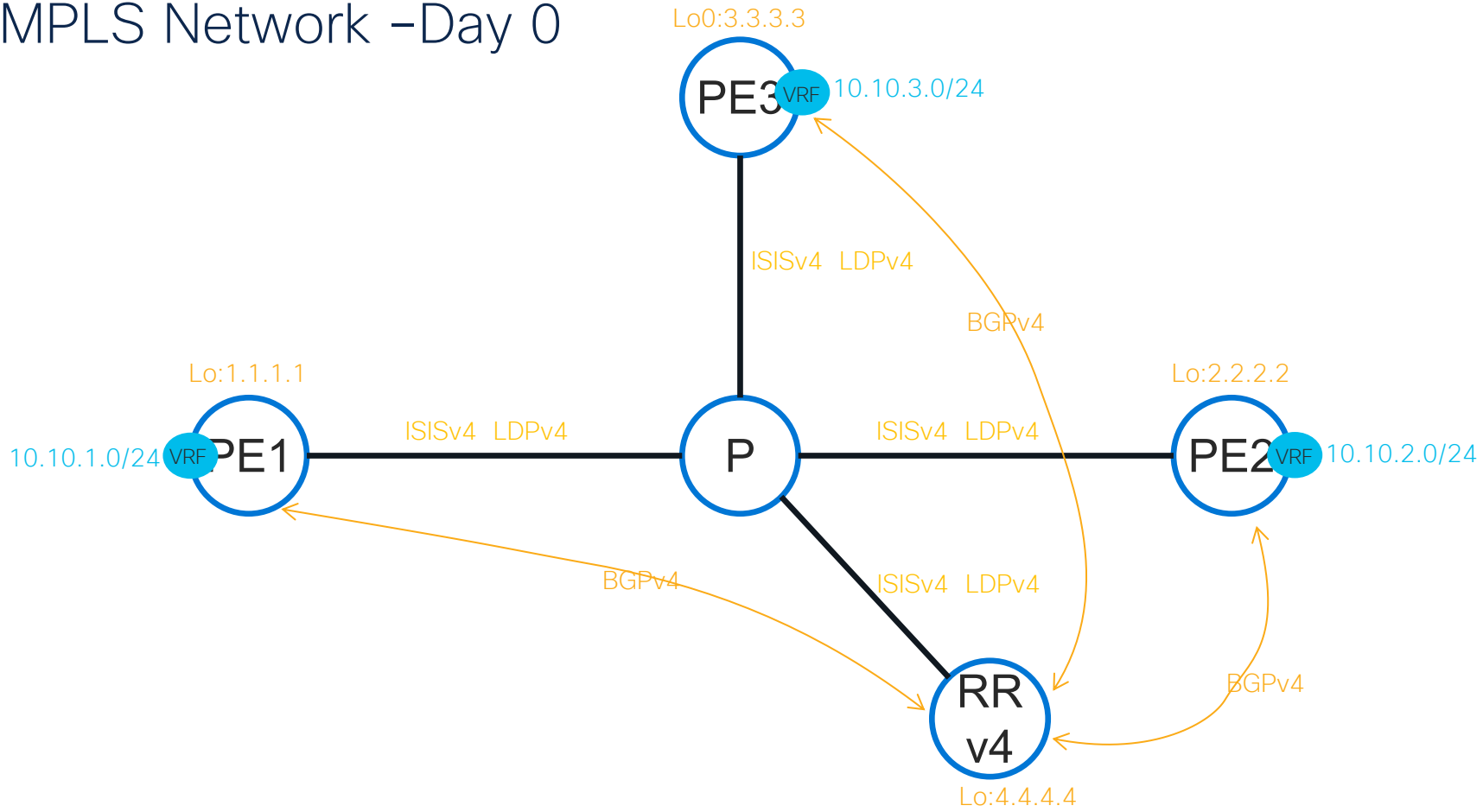


cisco Live!

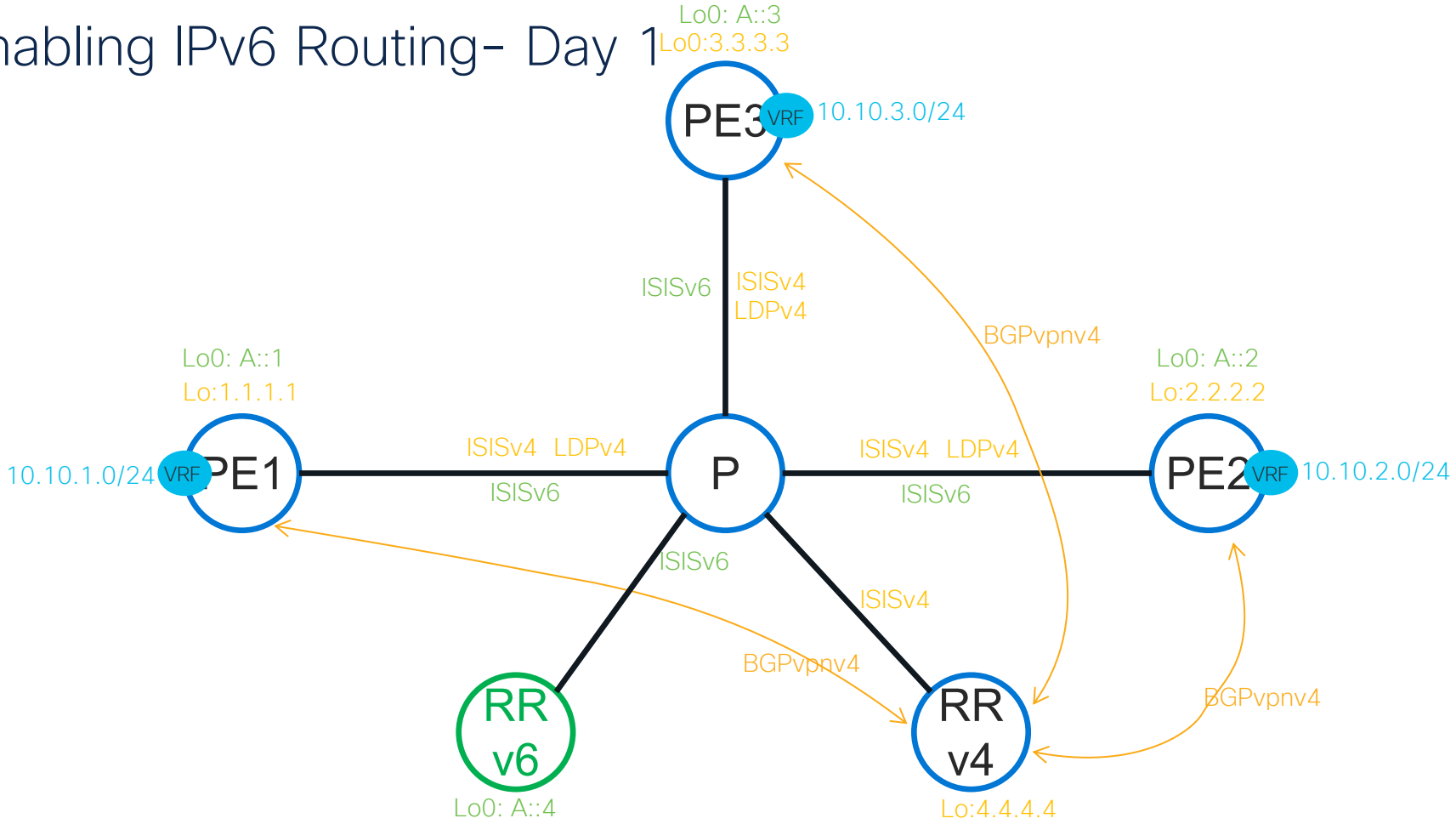
Dual Connected PE



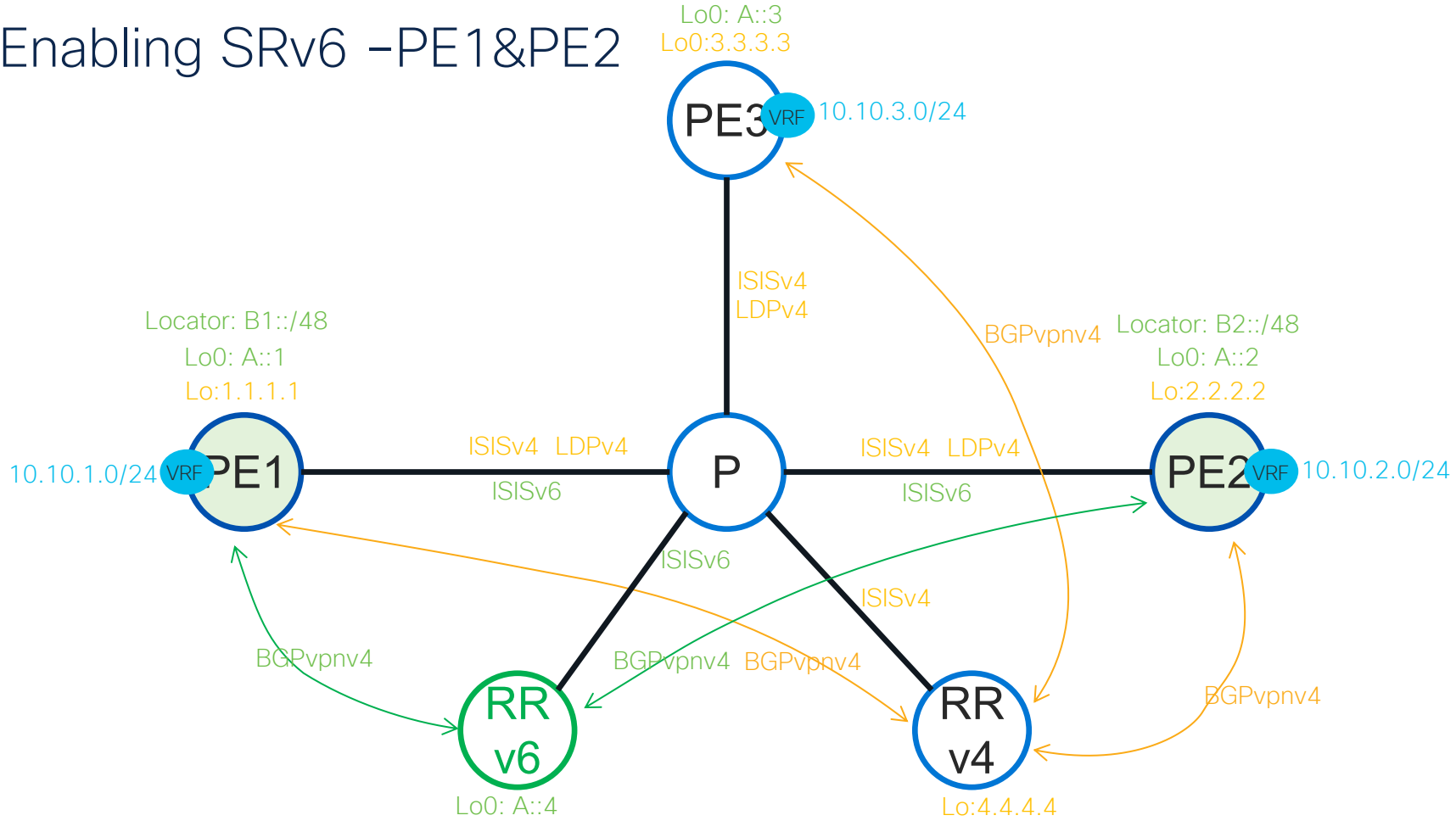
MPLS Network –Day 0



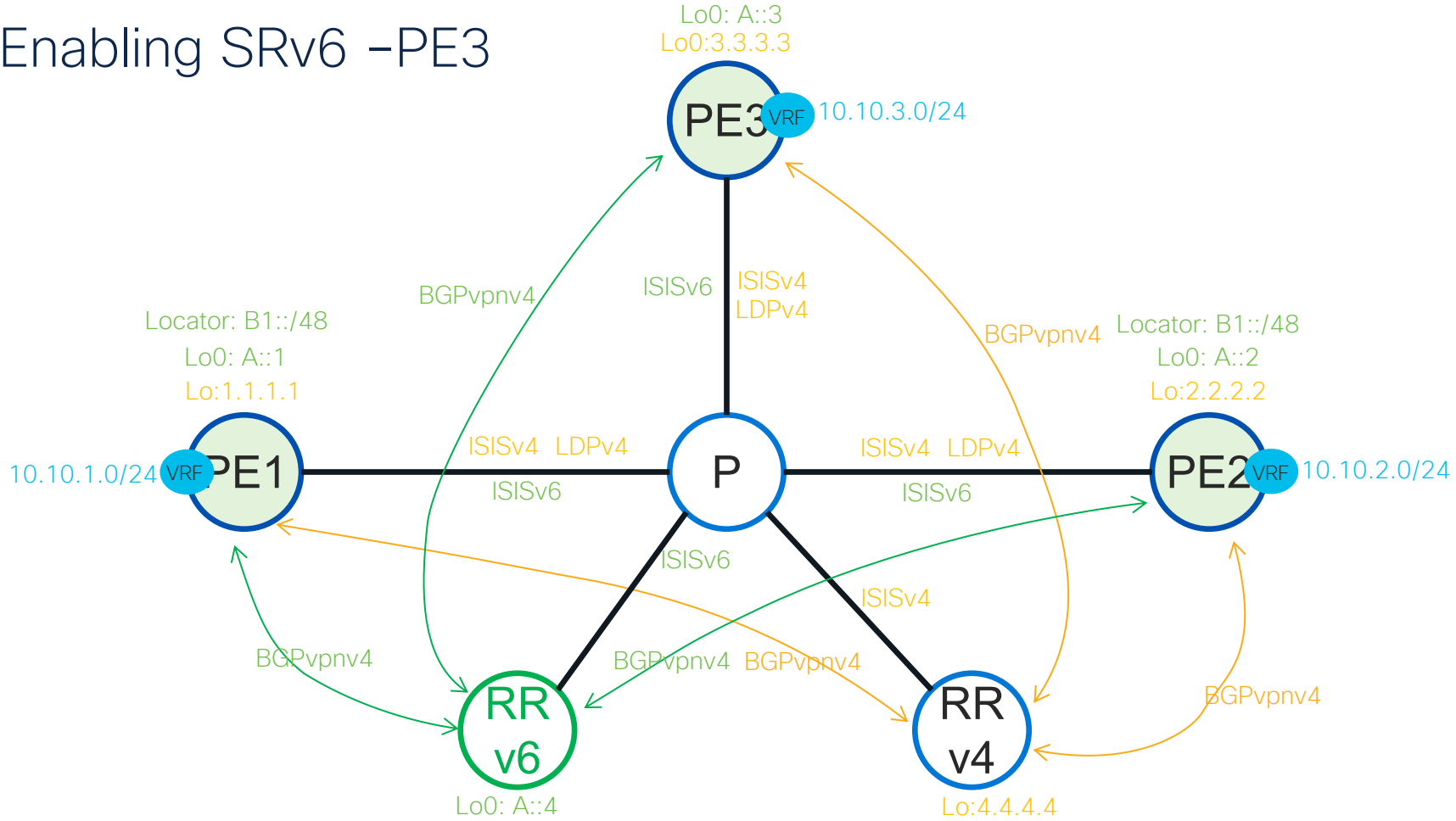
Enabling IPv6 Routing- Day 1



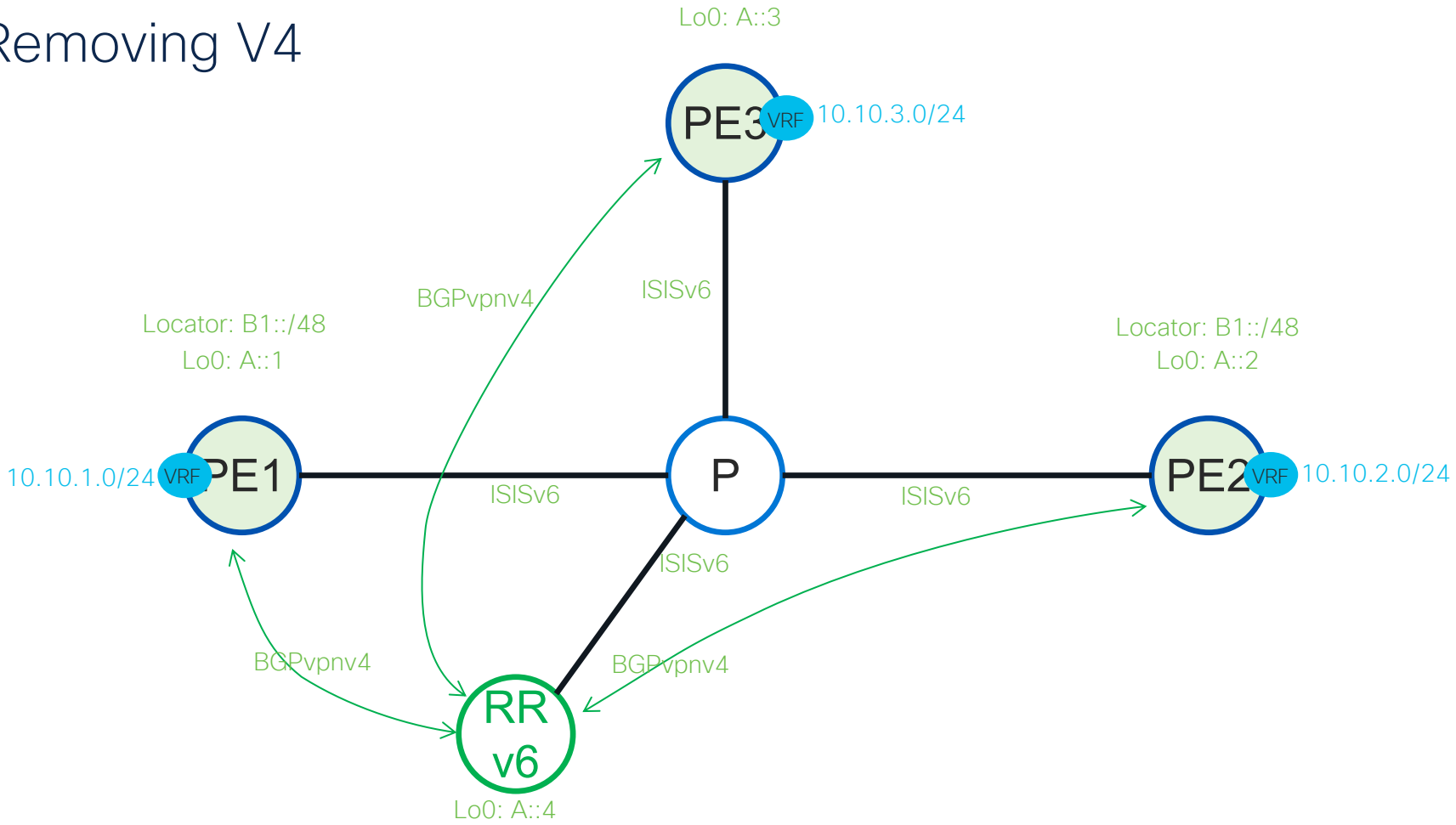
Enabling SRv6 –PE1&PE2



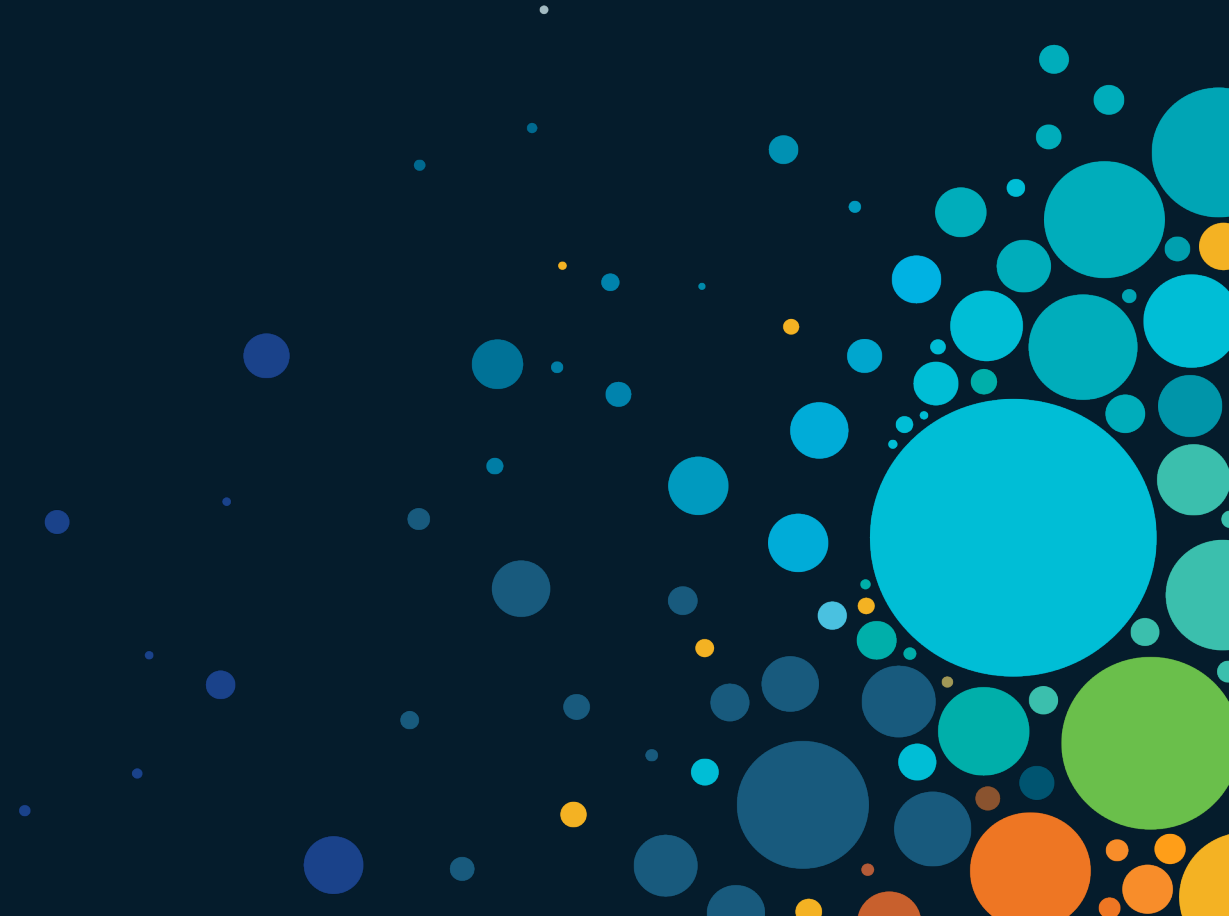
Enabling SRv6 –PE3



Removing V4



Conclusion



Rich Ecosystem

Network Equipment Manufacturers

Cisco, Huawei, Nokia, Ciena, Juniper Networks, Arrcus, ZTE, H3C

Merchant Silicon

Cisco, Broadcom, Barefoot Networks (Intel), Marvell

Open-Source Applications

Pyroute, IPTABLES, SERA, GoBGP, exa, BGP, eBPF, TCPDUMP, WIRESHARK, SNORT, BIRD, cilium

Open-Source Networking Stacks

SONiC, Linux, FD.io, P4, Open Compute Project, Kubernetes, OpenStack

Smart NIC

Intel, Mellanox Technologies

Partners

kaloom, Trend Micro, ENEA Qosmos Division, NoviFlow, ixia A Keysight Business, Spirent Promise. Assured.

Simplicity Always Prevails

~~LDP~~

~~RSVP-TE~~

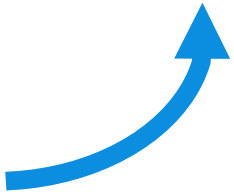
~~BGP 3108~~

~~MPLS~~

~~UDP/VxLAN~~

~~NSH~~

Furthermore, with more scale



and functionality

Session Surveys

We would love to know your feedback on this session!

- Complete the session surveys in the Cisco Events mobile app. You'll earn some points in the Cisco Live Game and potentially win a prize.
- Complete a minimum of four session and the overall event surveys to claim a Cisco Live cable bag.

Continue your education

Visit the Cisco Showcase for related demos

Book your one-on-one Meet the Expert 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

Cisco Learning and Certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. www.cisco.com/go/certs

Pay for Learning with Cisco Learning Credits

(CLCs) are prepaid training vouchers redeemed directly with Cisco.

Learn

Cisco U.
IT learning hub that guides teams and learners toward their goals

Cisco Digital Learning
Subscription-based product, technology, and certification training

Cisco Modeling Labs
Network simulation platform for design, testing, and troubleshooting

Cisco Learning Network
Resource community portal for certifications and learning

Train

Cisco Training Bootcamps
Intensive team & individual automation and technology training programs

Cisco Learning Partner Program
Authorized training partners supporting Cisco technology and career certifications

Cisco Instructor-led and Virtual Instructor-led training
Accelerated curriculum of product, technology, and certification courses

Certify

Cisco Certifications and Specialist Certifications
Award-winning certification program empowers students and IT Professionals to advance their technical careers

Cisco Guided Study Groups
180-day certification prep program with learning and support

Cisco Continuing Education Program
Recertification training options for Cisco certified individuals

The bridge to possible

Thank you

#CiscoLiveAPJC