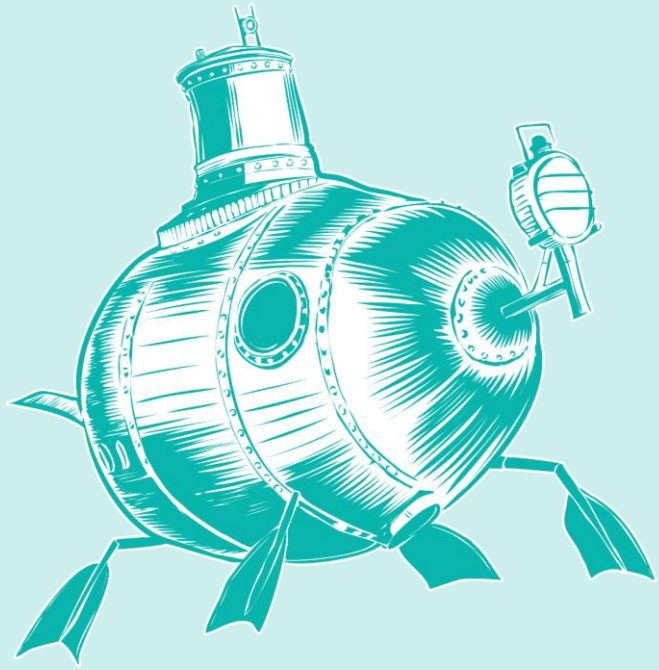
Trust the Strong

Evoluce k SRv6 – teorie a využití

Vladimír Bureš
vladimir.bures@alef.com

CSNOG 2025



ALEF

Segment Routing MPLS a SRv6

Classic MPLS

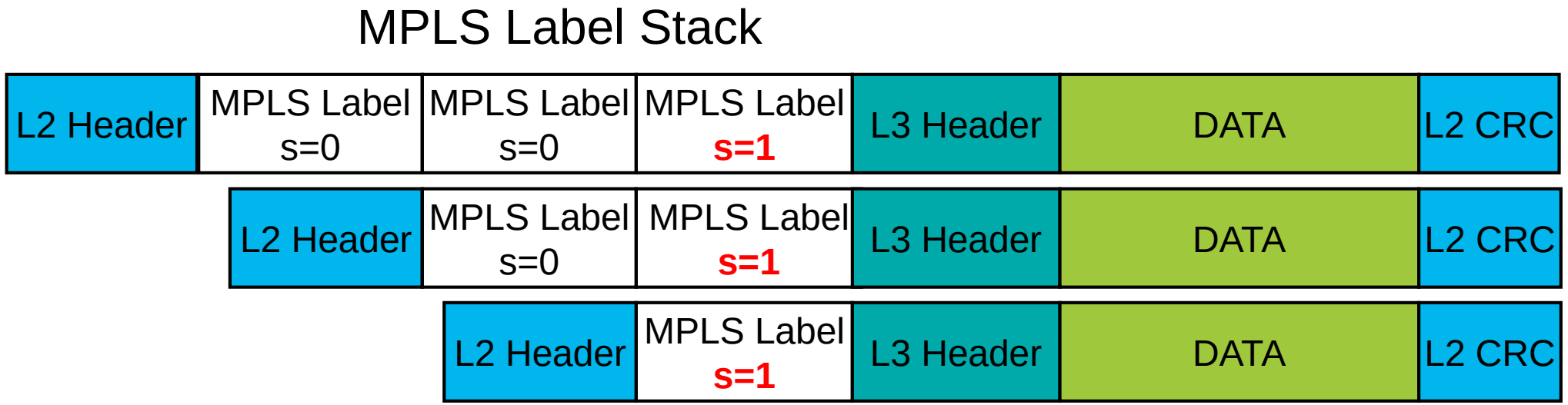
Classic MPLS (LDP) krátké shrnutí

MPLS

- MPLS - Multi Protocol Label Switching
- MPLS je přepínací mechanismus kde síťové prvky přepínají pakety podle značek (MPLS Labels) nesených v hlavičce MPLS paketu.
- MPLS Labels jsou obvykle svázány s cílovou IP adresou, může to být však L3/L2, QoS skupina, MPLS TE Tunel, ...
- Dynamická alokace MPLS labelů, lokální platnost
- Původně Cisco proprietální řešení, následně RFC standard, TDP->LDP, přidány další funkcionality (L2VPN, mLDP, ..)
- IETF's MPLS Working Group založena 1997
- 1998 První nasazení MPLS VPNs
- 1999 MPLS TE

Label Encapsulation

- „Label stack“ je umístěn mezi hlavičkou druhé (L2) a třetí vrstvy (L3)
- Začátek „label stack“ (top label) je přímo za L2 hlavičkou
- L3 hlavička je umístěna bezprostředně za MPLS labelem s nastaveným „**s**“ bitem
- Veškerý forwarding se vždy odehrává podle „**top labelu**“



Classic MPLS a jeho slabiny

- Komplikovaný control plane (LDP, RSVP-TE)
  - LDP
    - další protokol k IGP
    - nutno řešit LDP-IGP synchronizaci
  - RSVP-TE (RSVP)
    - Komplikované nasazení i následná údržba (troubleshooting), omezená škálovatelnost
    - Problematické rozkládání (ECMP)
    - MPLS FRR pro rychlou konvergenci časem “nahrazeno” IP LFA/R-LFA
- Neřeší „Microloop avoidance“

Segment Routing MPLS a SRv6

Segment Routing

Obecný úvod

Proč Segment Routing ?

- Zjednodušení
  - automatizace: TILFA sub-50msec FRR
  - eliminace používaných protokolů: LDP, RSVP-TE, VxLAN, ...
- Application Engineered Routing
  - Škalovatelnost – síť nedrží žádnou stavovou informaci „per-flow“ pro Traffic Engineering či NFV
  - Bezstavovost - Source Routing – cesta (topologická a/nebo služeb (NFV)) je kódována přímo v hlavičce paketu
  - Programovatelnost
  - Jednoduchá konfigurace a troubleshooting
- Lze nasadit „End-to-End“
  - DC, Metro, WAN

Co je to „Segment Routing“ ?

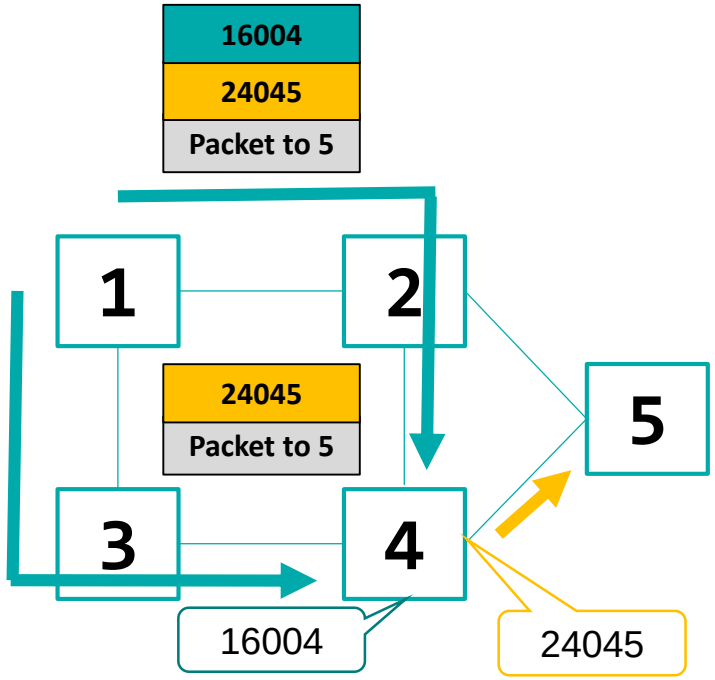
- Source Routing
  - zdroj určí cestu, kterou má paket použít a zakóduje ji do jeho hlavičky jako uspořádaný seznam.
  - zbytek sítě pouze provádí instrukce zakódované v hlavičce paketu
- Segment
  - jednotlivá instrukce (forwarding či služba)
- Segment List
  - uspořádaný seznam instrukcí

Segment Routing Data Plane

- MPLS -> **SR-MPLS**
  - uspořádaný seznam segmentů je reprezentován jako **MPLS label stack**
    - 1 segment = 1 MPLS label, segment list = MPLS label stack
  - Segment Routing **využívá MPLS data plane (bez jakýchkoliv změn)**
  - Segment je reprezentován jako MPLS label
  - Lze aplikovat jak IPv4 tak i IPv6
- IPv6 -> **SRv6**
  - uspořádaný seznam segmentů je vložen do „**routing extension header**“ v IPv6 hlavičce paketu
    - 1 segment = 1 adresa, segment list = uspořádaný seznam adres v SRH
  - RFC8200 – „Routing Header“ část

Kombinace IGP Segmentů (MPLS)

- Kombinací Globálnich a Lokálních segmentů lze poslat provoz libovolnou cestou v síti
- Cesta je *Segment list* - seznam Segmentů (zde stack MPLS labelů)
- Cesta **není** vysignalizovaná/postavená (LSP)
- Síť **nedrží** „stav“ pro konkrétní flow



All nodes use default SRGB 16.000-23.999

Segment Routing MPLS a SRv6

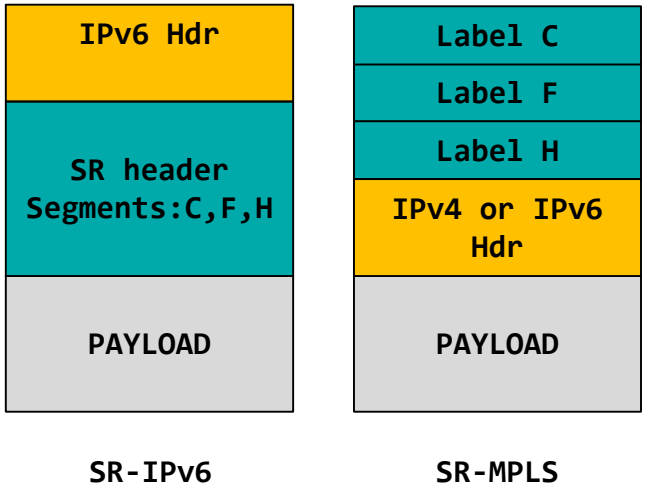
Segment Routing SRv6

IPv6 Data Plane

SR for IPv6 Dataplane

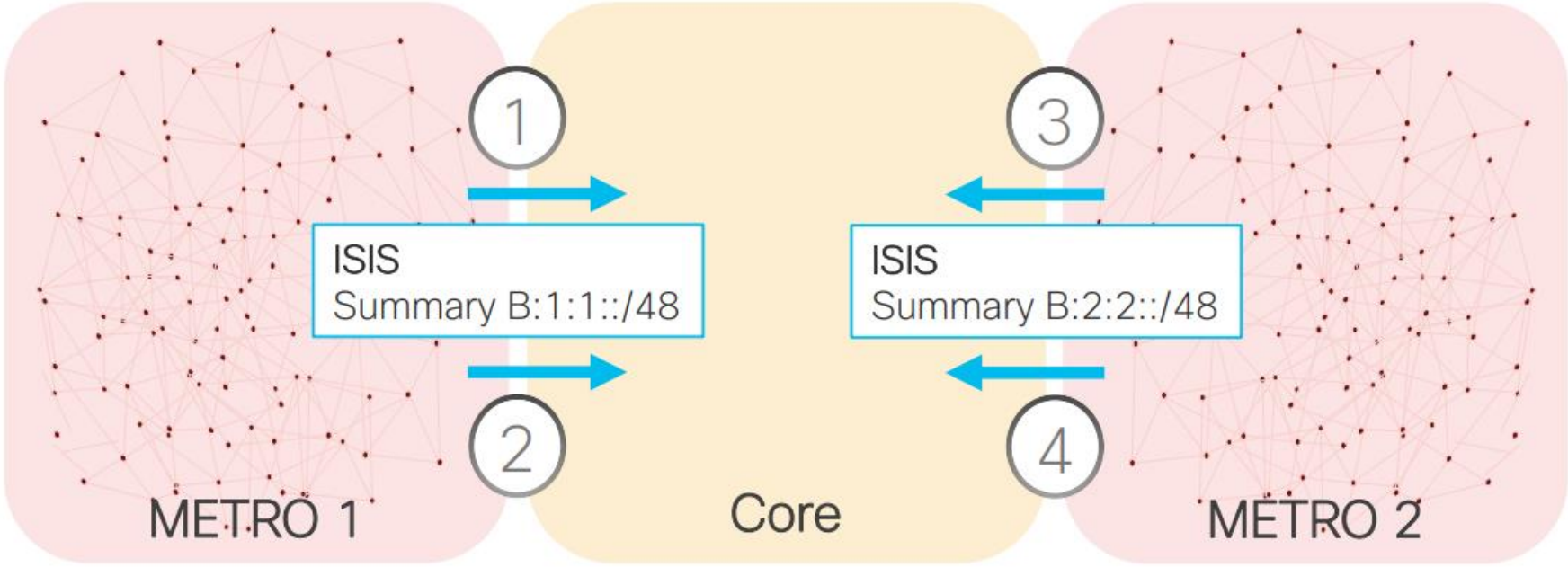
- Segment Routing pro IPv6 Dataplane
  - SID je definován jako 128 bit IPv6 adresa (žádný label, žádný MPLS Data Plane, čisté IPv6)
  - Výrazně jednodušší z pohledu signalizace – propaguje se pouze IPv6 prefix (prefix je SID)

- Segmenty jsou uloženy v IPv6 hlavičce
  - zavádí se nový „*Routing Extensions Header*“ typ
    - Segment Routing Header (SRH)
    - Obsahuje Segment List
    - VELMI blízké předchozímu standardu RFC2460



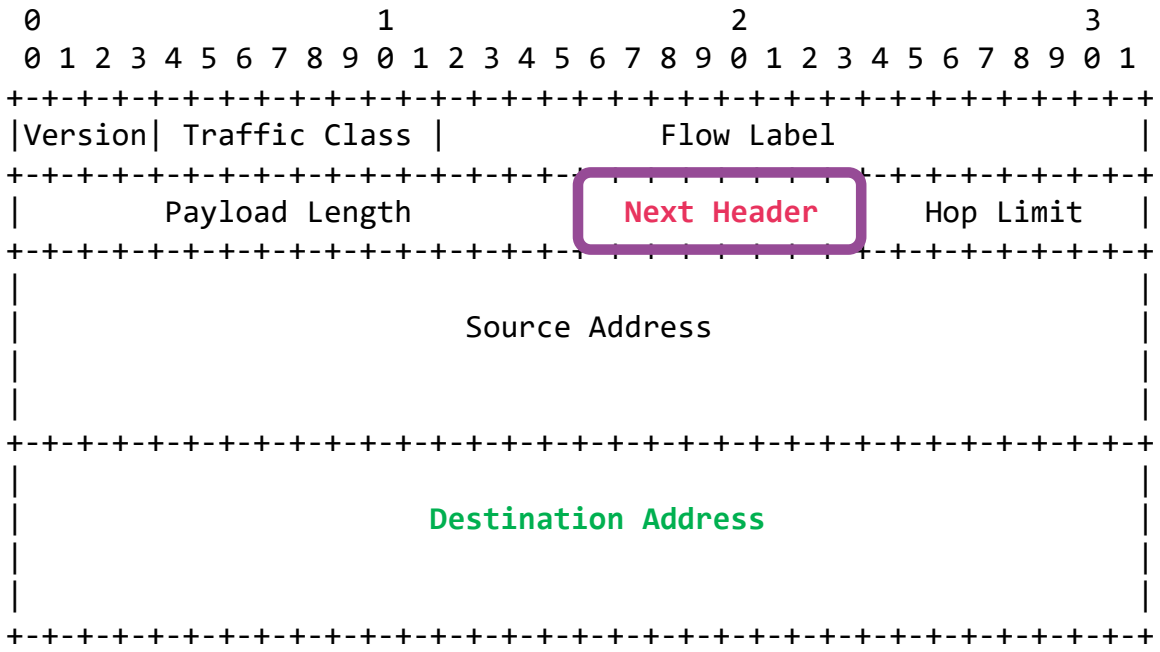
Prefix Summarization

- Návrat s základním principům IP směrování **a sumarizaci**
- Nepotřebujeme BGP inter-AS Option A/B/C, BGP-LU



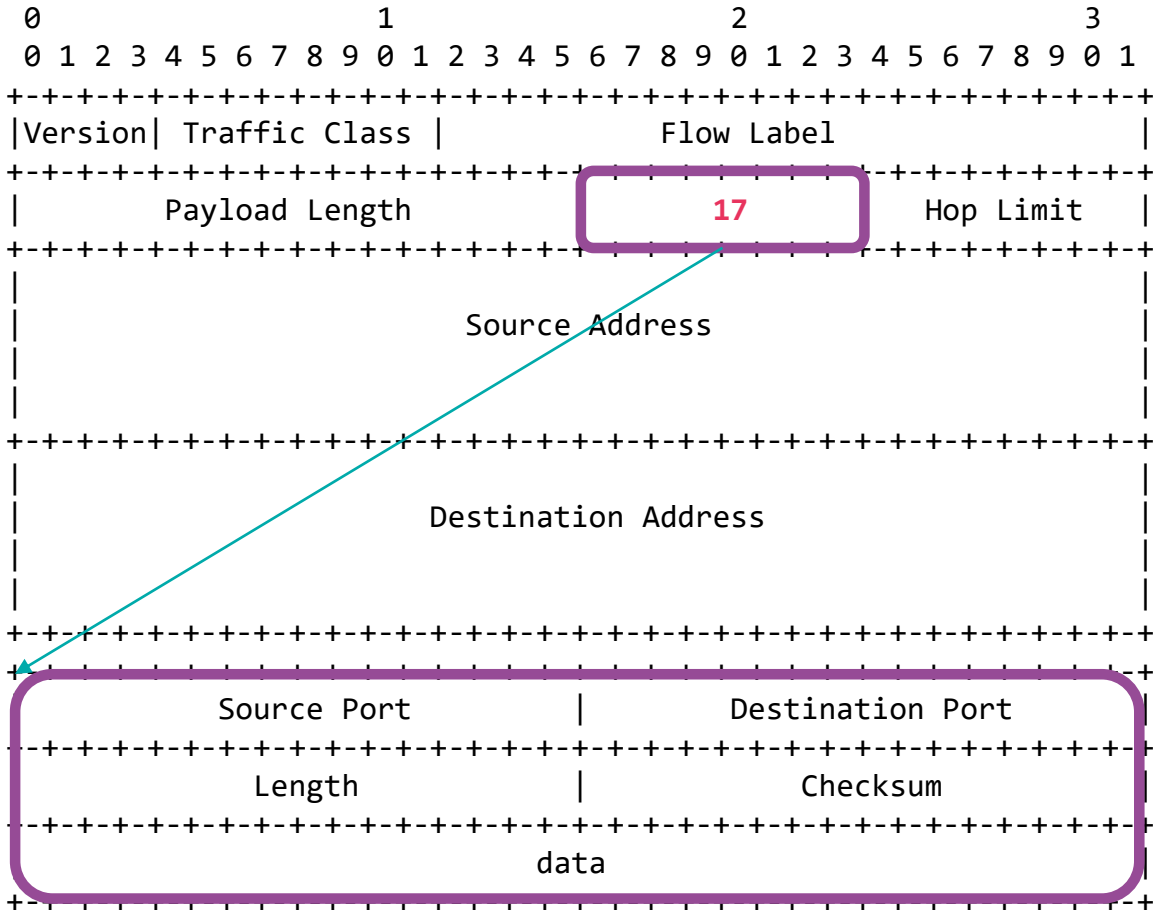
IPv6 SR Header

- IPv6 hlavička
  - **Next header** pole: co následuje v další hlavičce
  - Destination address



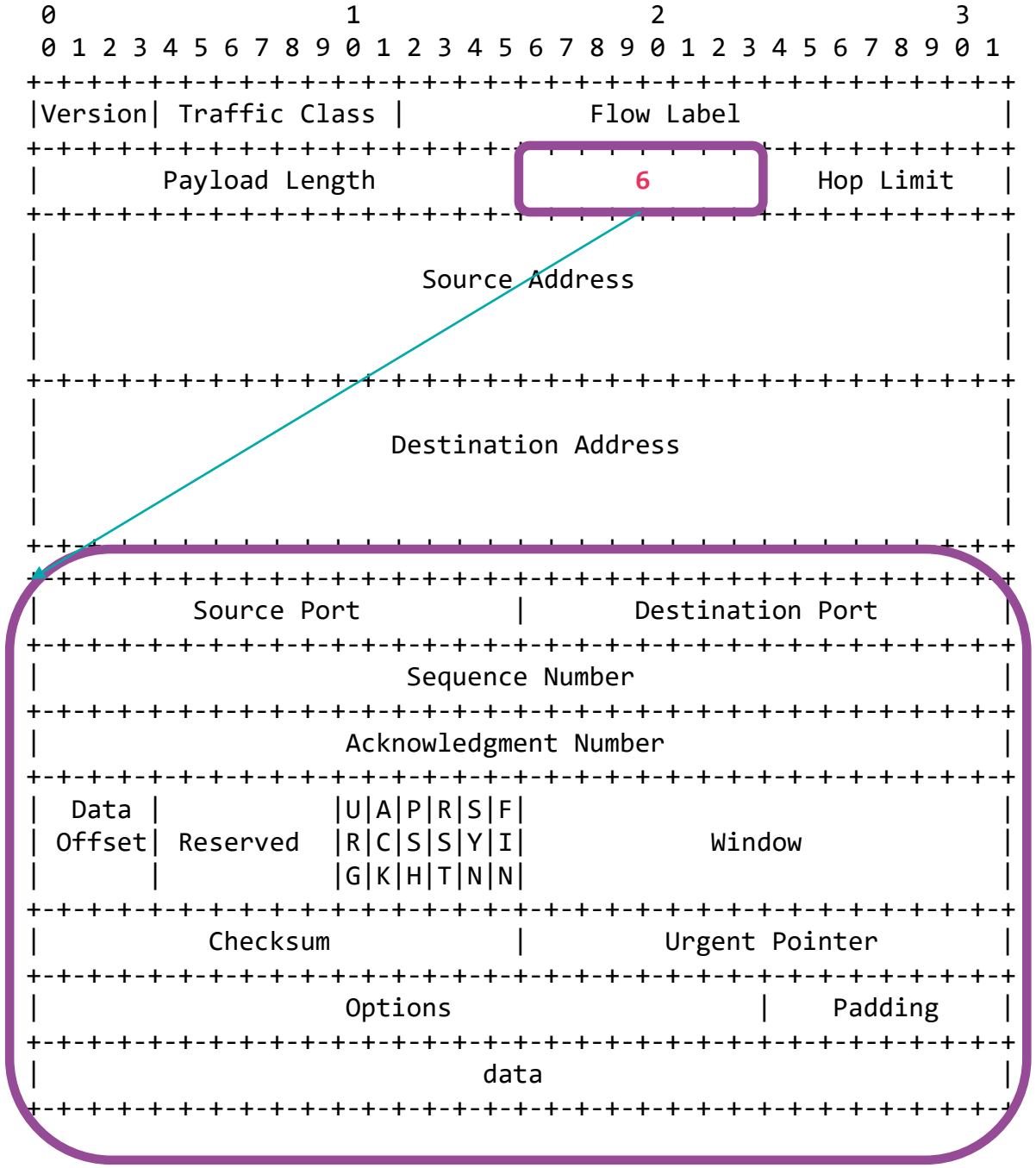
IPv6 SR Header

- IPv6 hlavička
  - **Next header**: 17 -> UDP



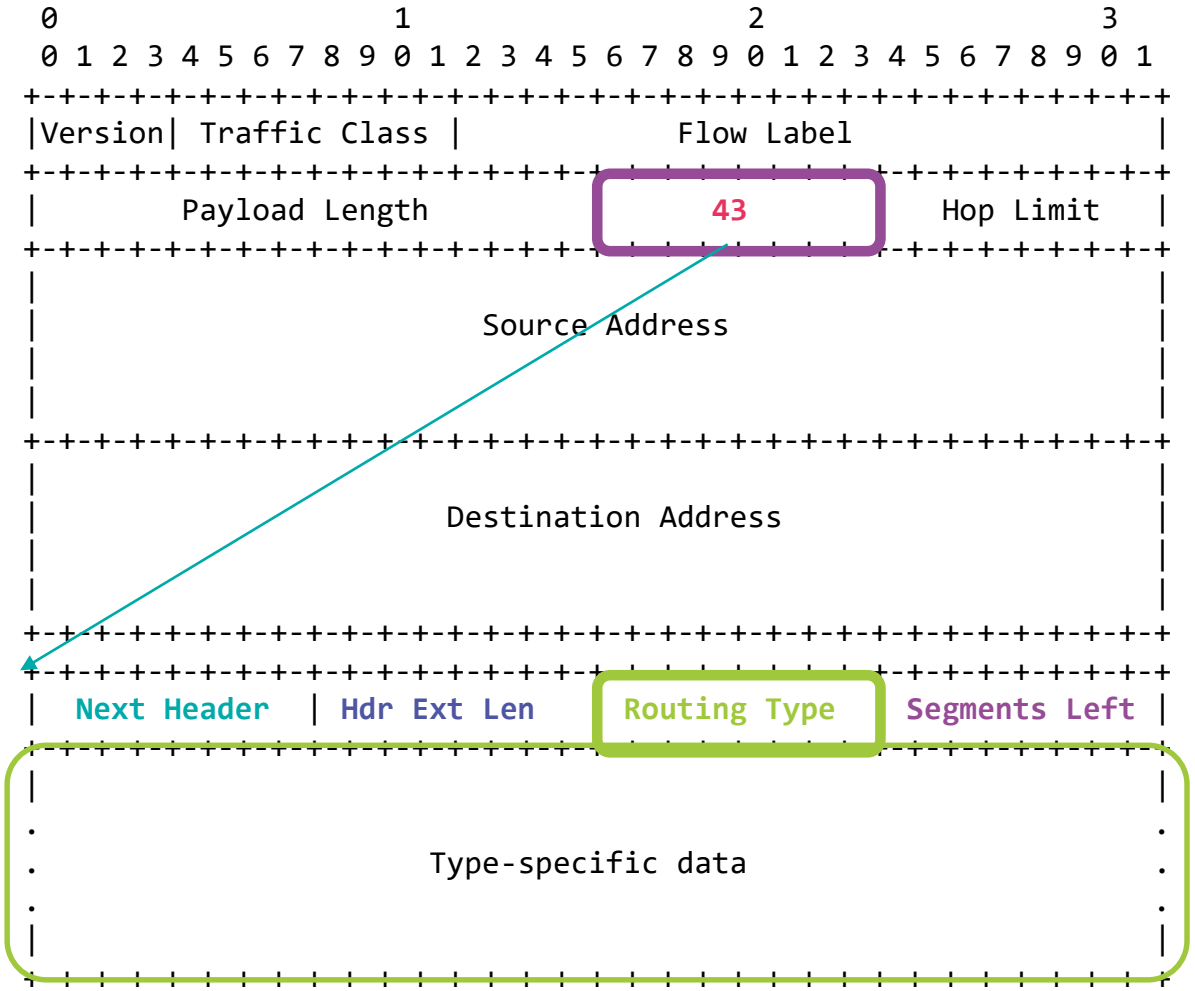
IPv6 SR Header

- IPv6 hlavička
  - **Next header**: 6 -> TCP



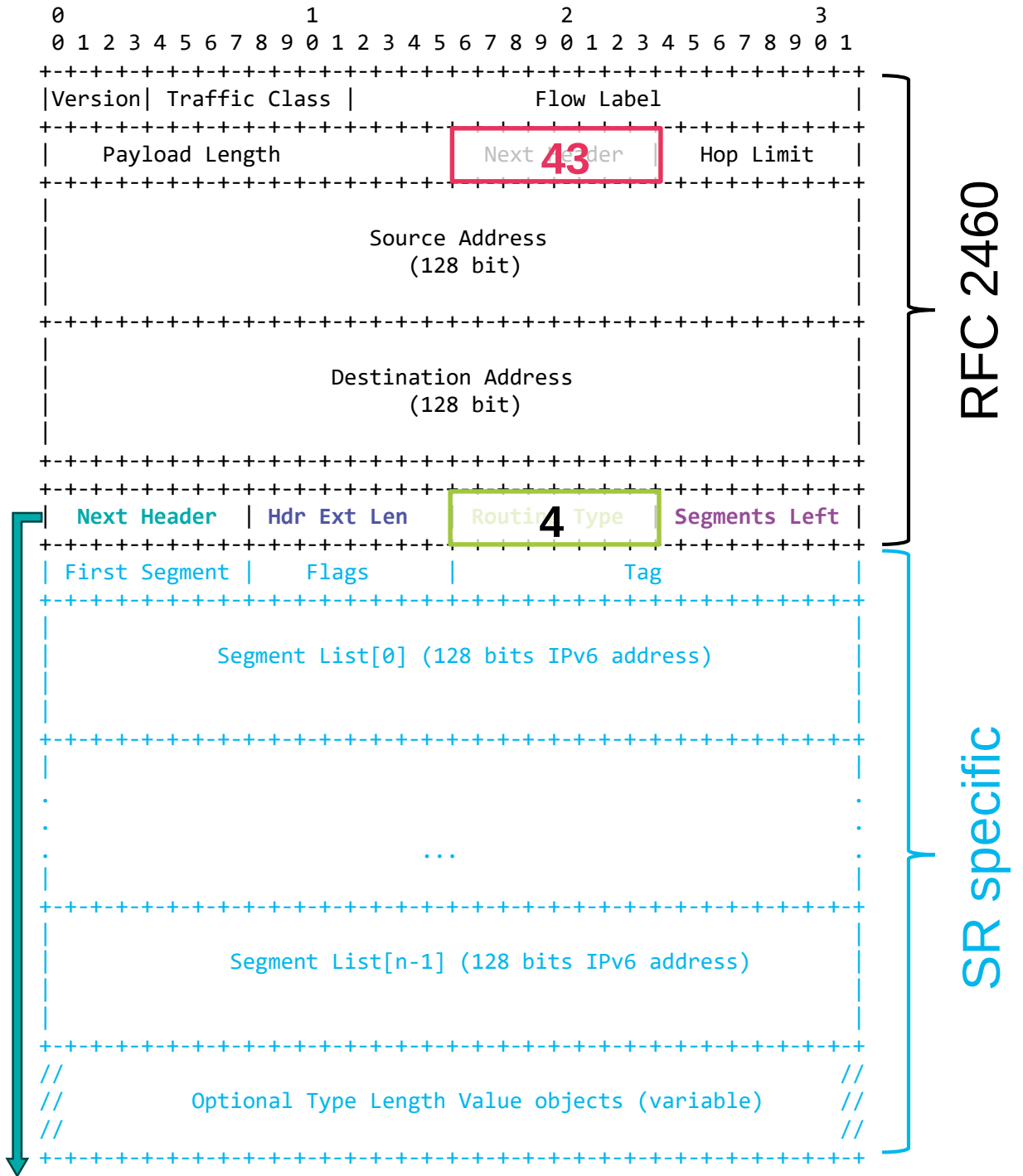
IPv6 SR Header

- IPv6 hlavička
  - **Next header**: 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**



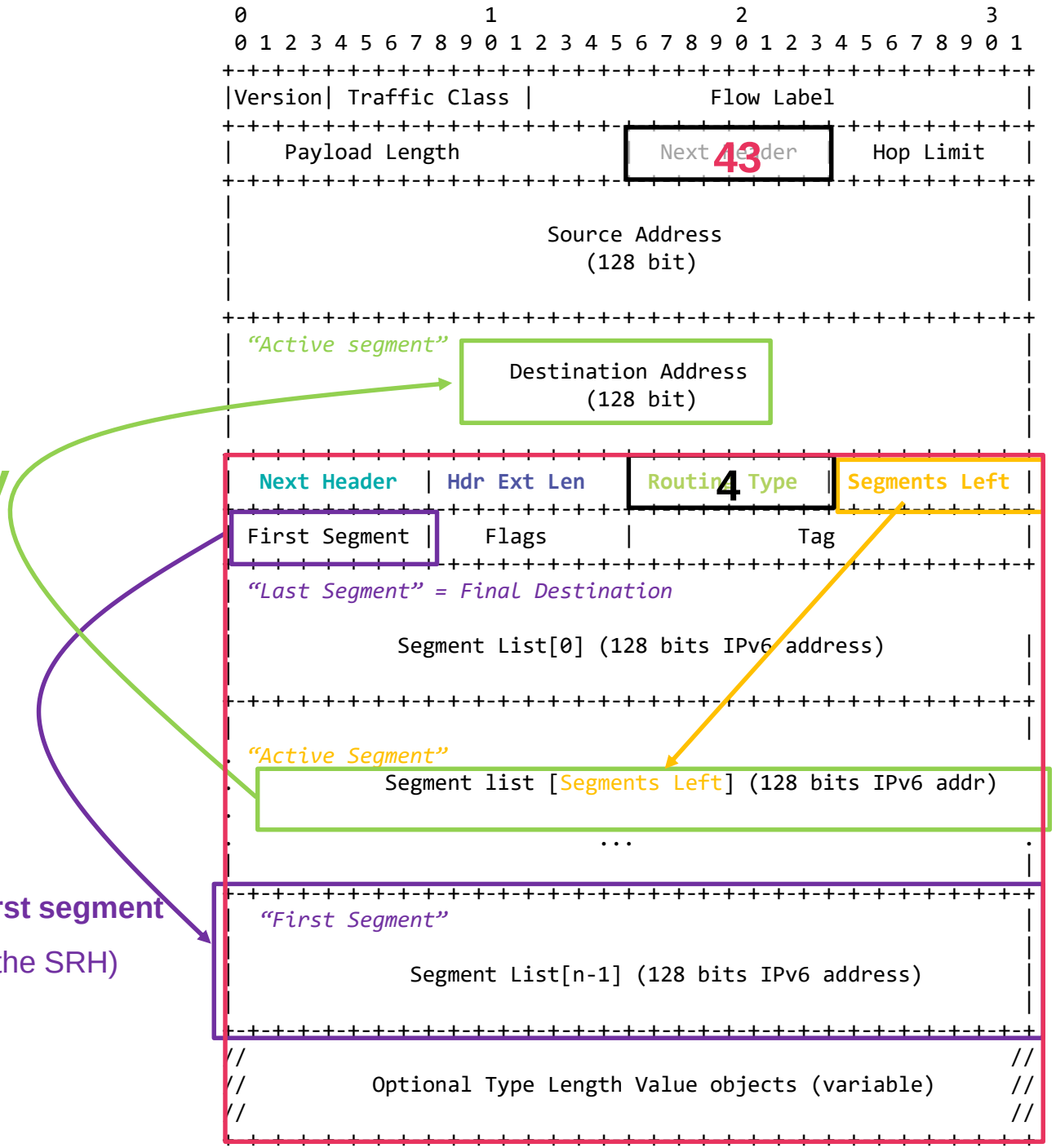
IPv6 SR Header

- IPv6 hlavička
  - **Next header**: 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 (RFC8754)**



IPv6 SR Header

- IPv6 hlavička
  - **Next header**: 43 -> Routing
  - First Segment
    - Pointer to very first SID

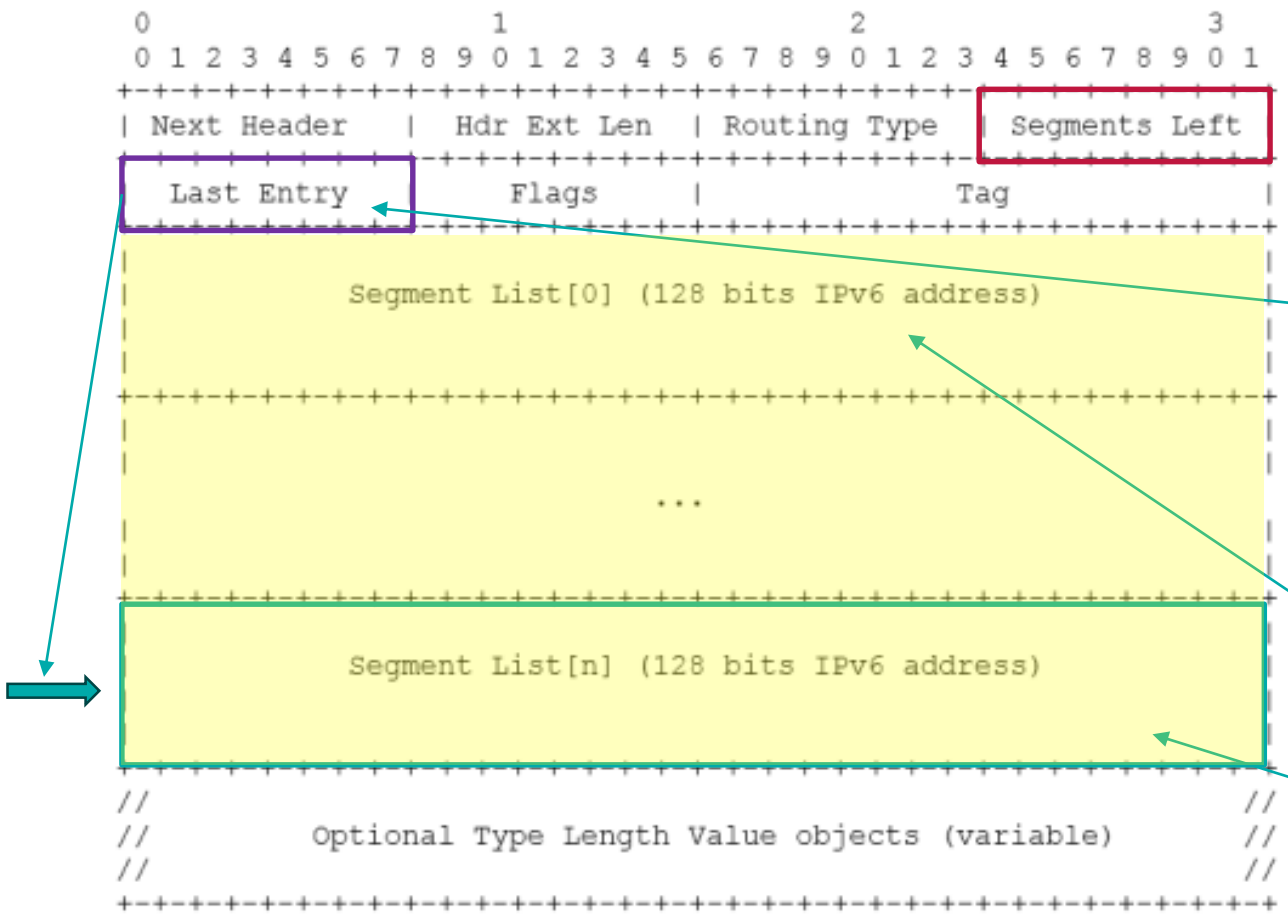


copy

Pointer to the **First segment**

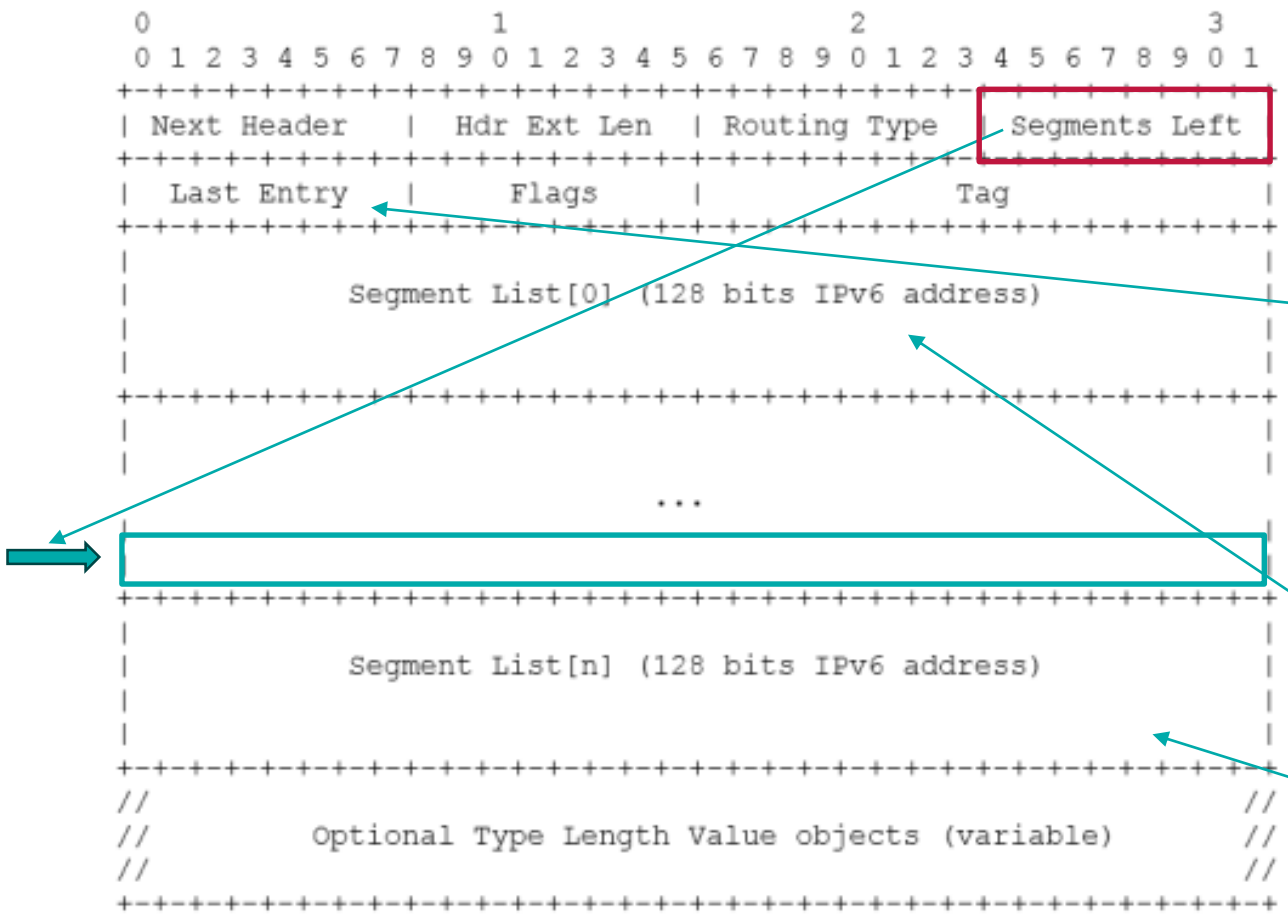
(= last record in the SRH)

SRv6 Dataplane # n



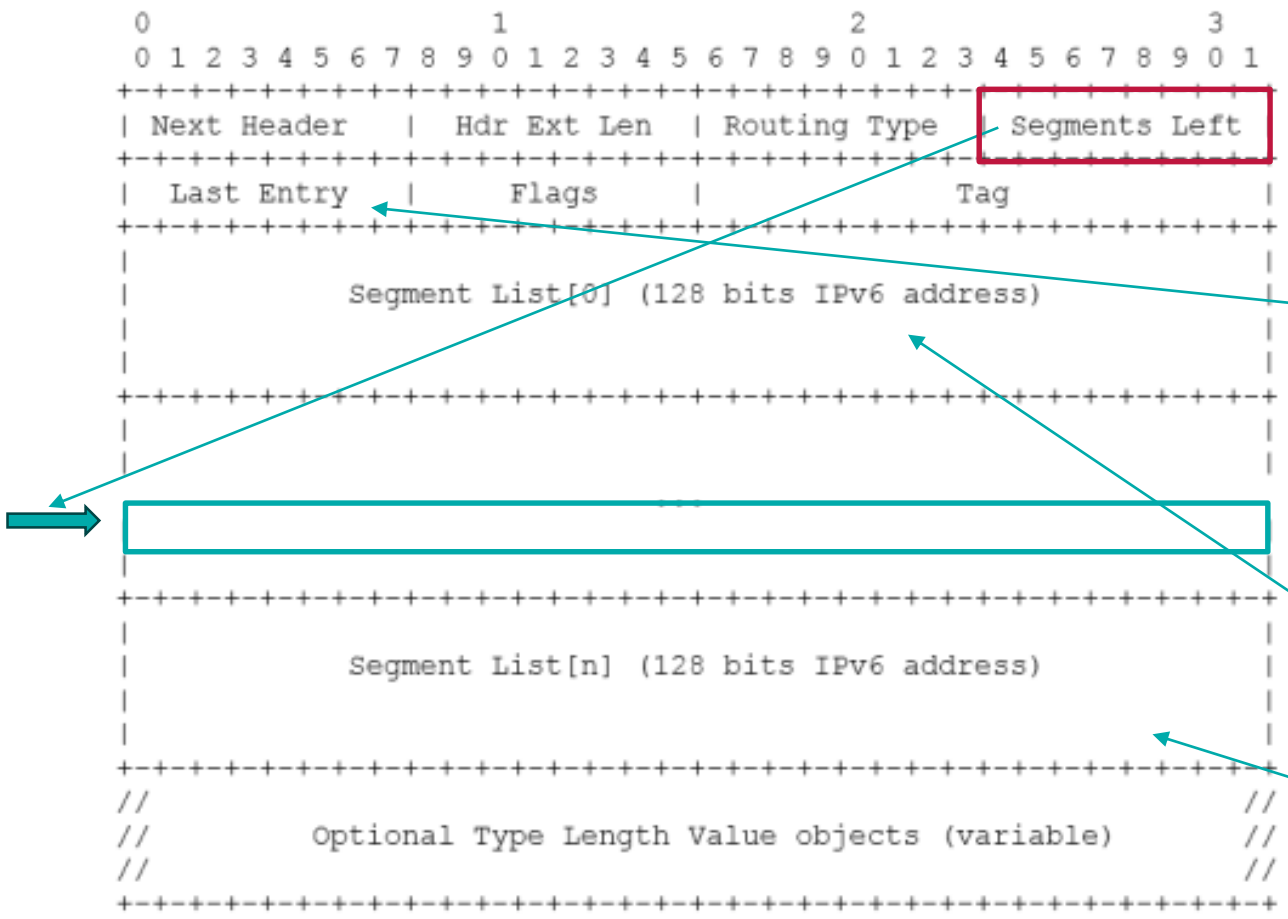
- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

SRv6 Dataplane # n-1



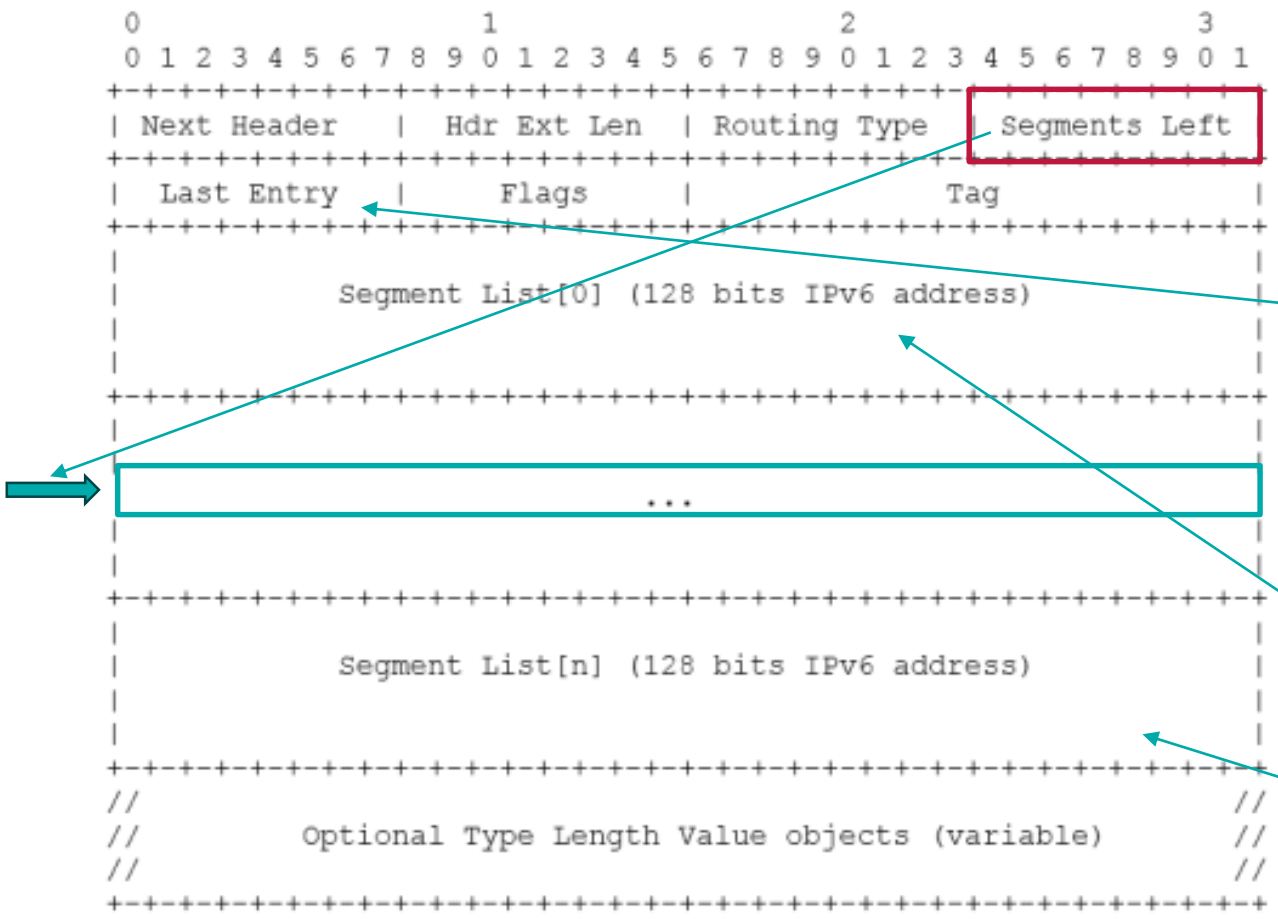
- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

SRv6 Dataplane # n-2 ….



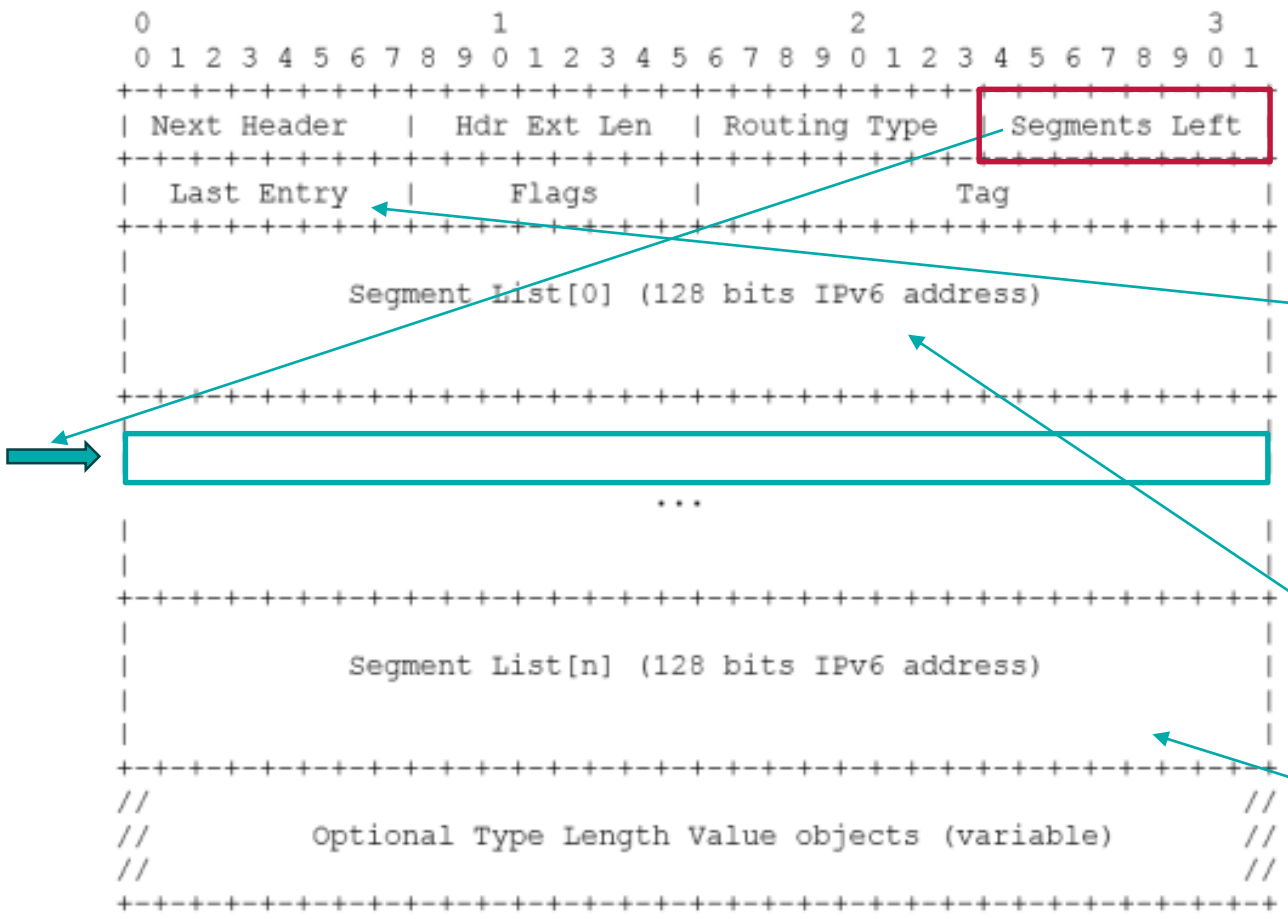
- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

SRv6 Dataplane # 3



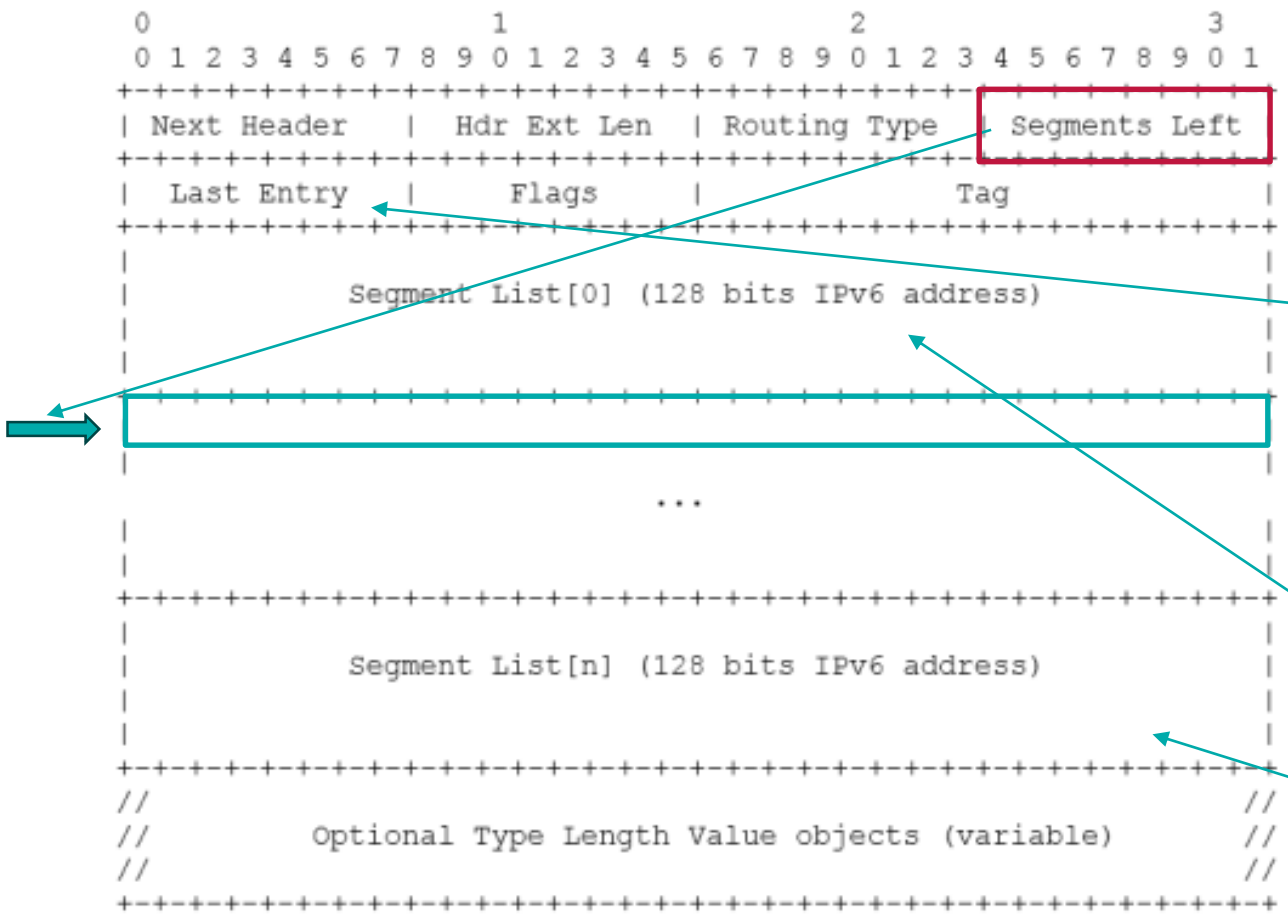
- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

SRv6 Dataplane # 2



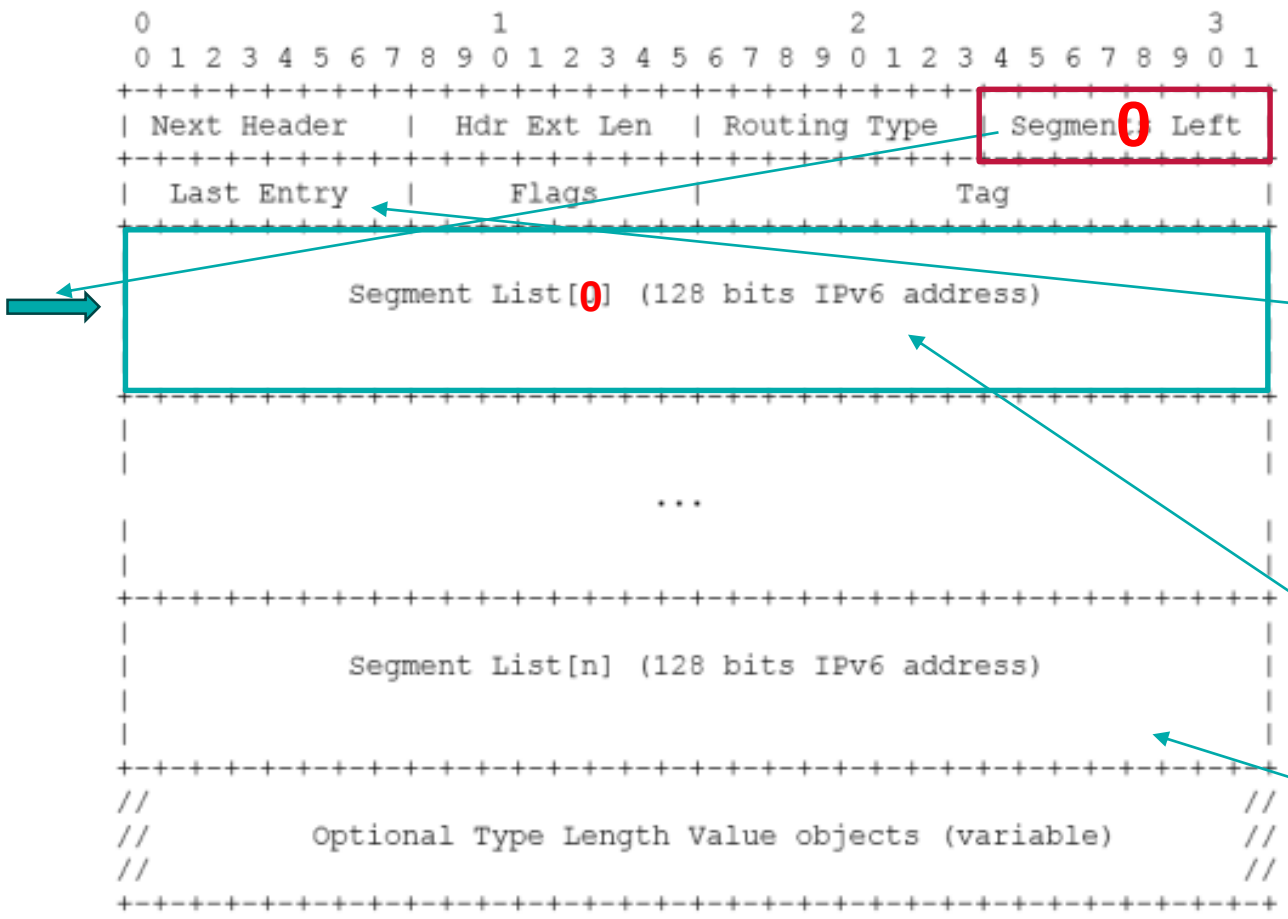
- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

SRv6 Dataplane # 1



- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

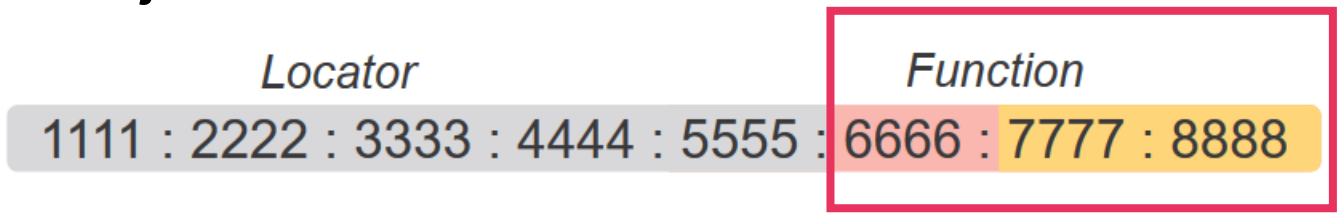
SRv6 Dataplane # 0 (last)



- **Segment Left** – ukazuje na **active Segment** v **Segment Listu**
- Mění se POUZE hodnota ukazatele – tj. řádek Segments List [xxx] se neodstraňuje jako v případě MPLS labelů
- Poslední Segment SR Policy (index = 0)
- První Segment SR Policy (index = *n*)

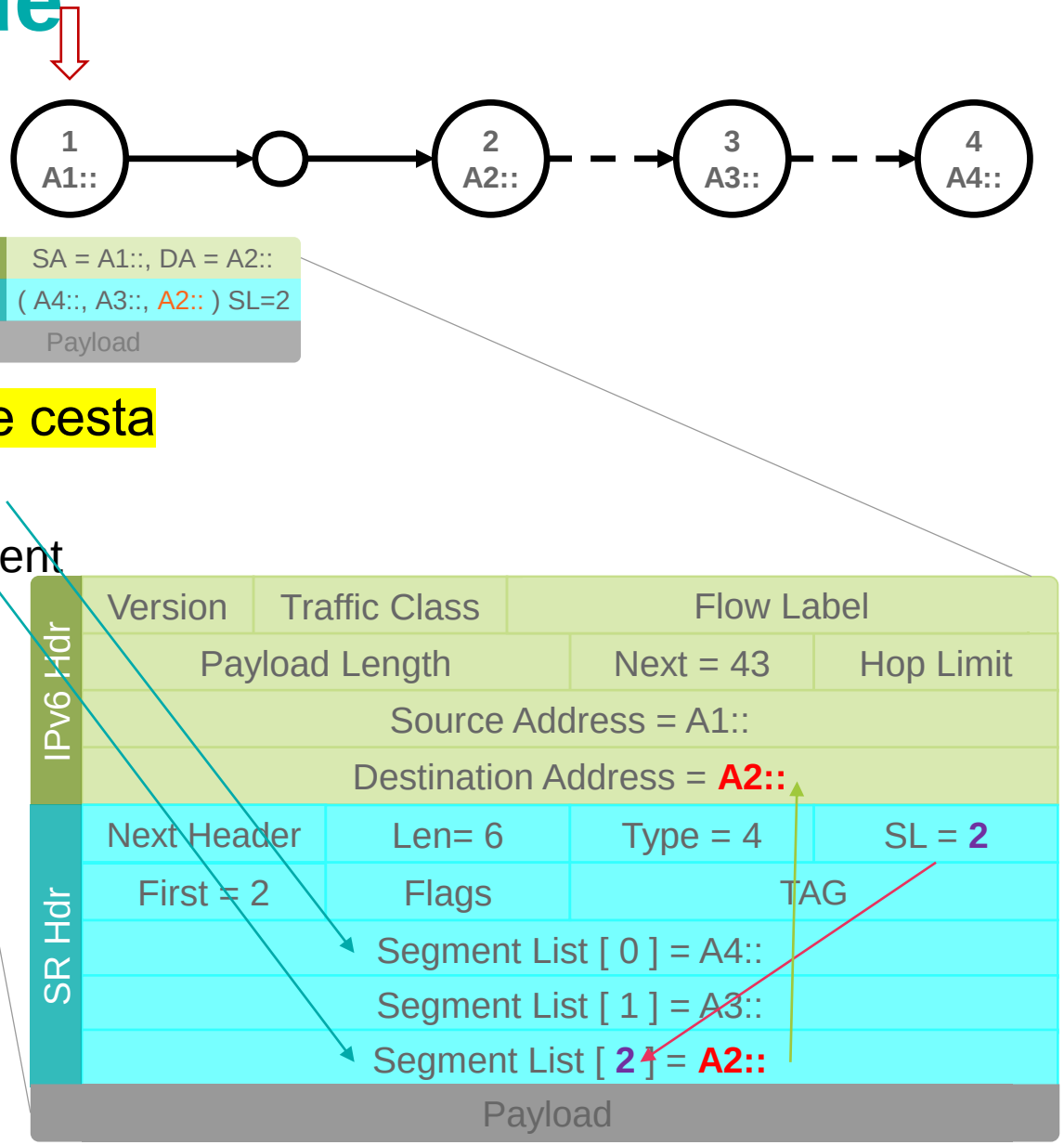
Network instruction

- 128-bit SRv6 SID
  - **Locator** (horní část ze 128b SIDu): směrován k uzlu, který provádí Funkci
  - **Function** (dolní část ze 128b SIDu): libovolná funkce (případně doplněná o argument)
    - vykoná lokálně NPU či aplikace ve VM/Container
  - **Agrument** (nejnižší bity) – volitelná/nepovinná část
- Hranice Locator/Function obecně je flexibilní

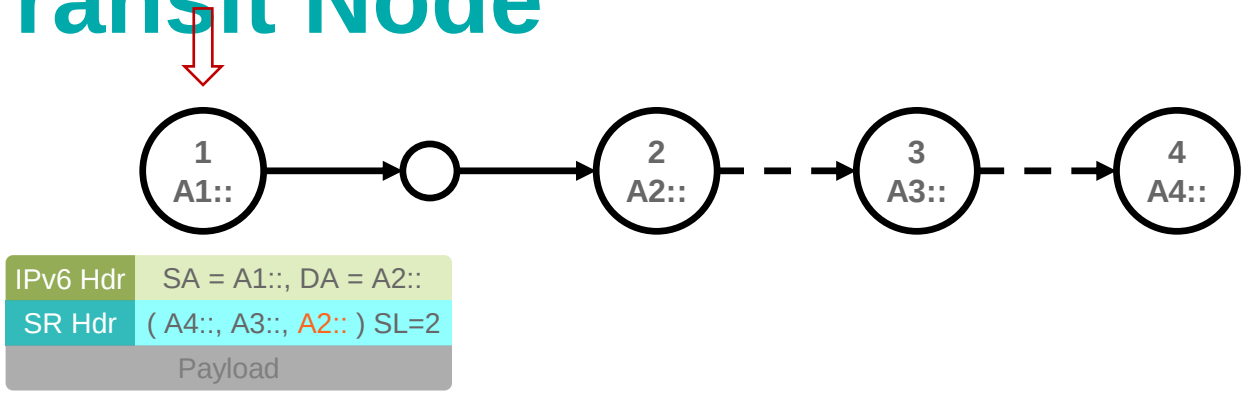


Forwarding – Source Node

- Zdroj je SR-capable
- SR Header (SRH) obsahuje
  - Segment list v obráceném pořadí než je cesta
    - Segment List [0] is the LAST segment
    - Segment List [$n-1$] is the FIRST segment
  - „Segments Left“ nastaven $n-1$
  - „First Segment“ nastaven $n-1$
- IPv6 Dest Adresa je FIRST segment
- Packet je odeslán k IPv6 Dest. Adrese
  - Normální IPv6 forwarding



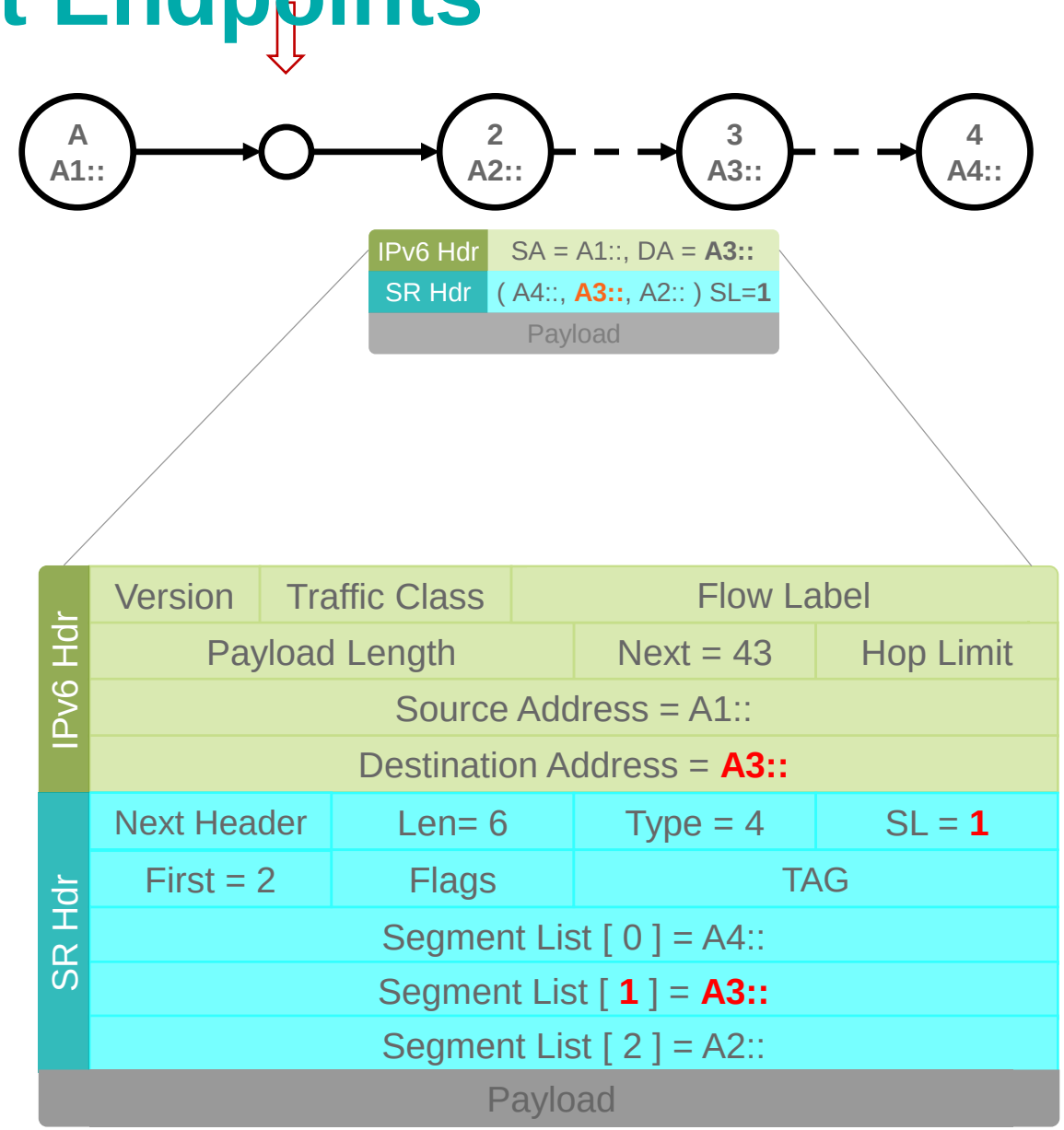
Forwarding - Non-SR Transit Node



- **Čistý IPv6 forwarding**
  - Výhradně podle IPv6 dest. adresy
  - Neprovádí se žádná inspekce ani modifikace SRH

Forwarding - SR Segment Endpoints

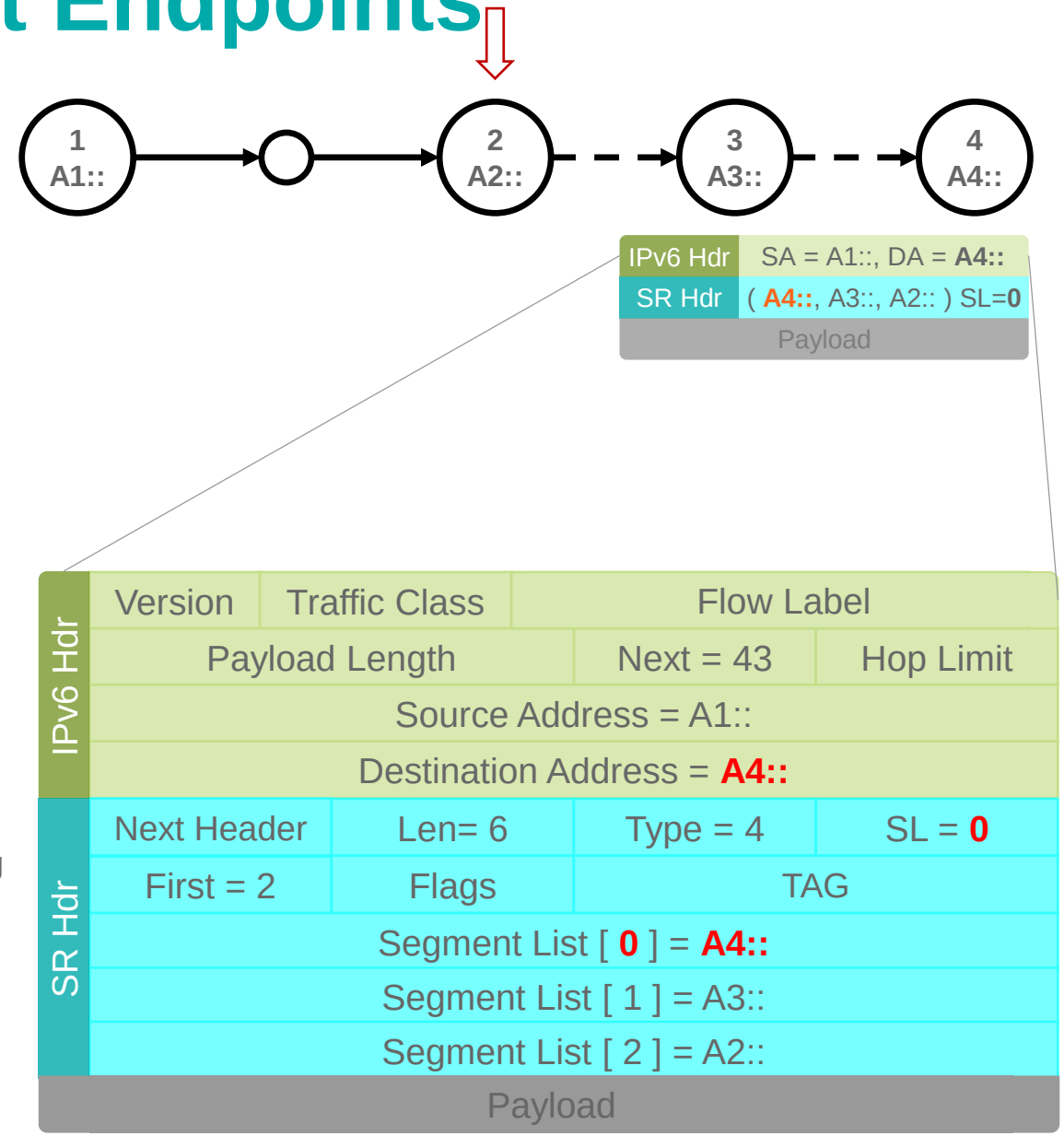
- SR Endpoints: SR-capable uzel jehož adresa je dst. adresou v paketu
- SR Endpoints zpracuje SRH a:
  - IF Segments Left > 0, THEN
    - Decrement Segments Left (-1)
    - Update DA with Segment List [Segments Left]
    - Forward according to the new IP DA



Forwarding - SR Segment Endpoints

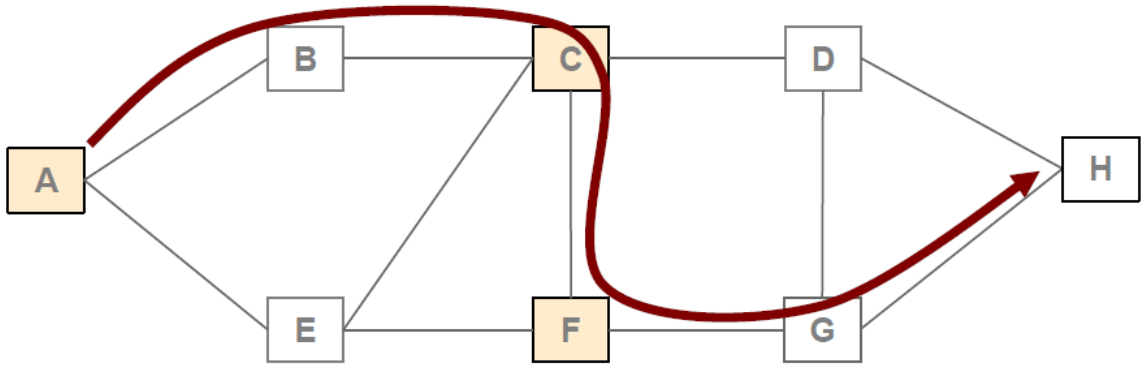
- SR Endpoints: SR-capable nodes whose address is in the IP DA
- SR Endpoints inspect the SRH and do:
  - IF Segments Left > 0, THEN
    - Decrement Segments Left (-1)
    - Update DA with Segment List [Segments Left]
    - Forward according to the new IP DA
  - ELSE (Segments Left = 0)
    - Remove the IP and SR header
    - Process the payload:
      - Inner IP: Lookup DA and forward
      - TCP / UDP: Send to socket
      - …

Standard IPv6 processing
The final destination does not have to be SR-capable.



Interoperability

- **Není třeba aby všechny uzly přes které protéká provoz „uměly“ SRv6**
- Příklad:
  - Traffic Engineering z A do H cestou „ABCFGH“
  - Uzly **A**, **C** a **F** rozumí SRv6
  - Uzly **B**, **D**, **E**, **G** a **H** zajišťují čistý IPv6 forwarding



Full SID příklad konfigurace – IOS-XR

```
!
segment-routing
 srv6
  logging locator status
  locators
   locator MyFullLoc01
    prefix fc00:ab00:2::/64
!
router isis 1
 is-type level-2-only
 net 49.0001.0000.0000.0002.00
!
 address-family ipv6 unicast
  metric-style wide
  segment-routing srv6
   locator MyFullLoc01
    level 2
   !
  !
 !
 interface Loopback0
  passive
  circuit-type level-2-only
  address-family ipv6 unicast
  !
 !
 interface GigabitEthernet0/0/0/0
  circuit-type level-2-only
  point-to-point
  address-family ipv6 unicast
  !
 !
```

Locator name

Locator value in IPv6 address format

Routing protocol for IPv6 routing

Full SID show commands příklad – IOS-XR

SIDs associated with *MyFullLoc01*

Configured Locators

Locator | Function (Arg)

```
RP/0/RP0/CPU0:R02#sh segment-routing srv6 locator MyFullLoc01
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           3        0     fc00:ab00:2::/64          Up
RP/0/RP0/CPU0:R02#sh segment-routing srv6 locator MyFullLoc01 detail
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           3        0     fc00:ab00:2::/64          Up
  Interface:
    Name: srv6-MyFullLoc01
    IFH : 0x0000002c
    IPv6 address: fc00:ab00:2::/64
  Number of SIDs: 3
  Created: Nov  5 14:12:55.129 (1d21h ago)
RP/0/RP0/CPU0:R02#

RP/0/RP0/CPU0:R02#sh segment-routing srv6 locator MyFullLoc01 sid
SID                         Behavior          Context                          Owner               State  RW
--------------------------  ----------------  -------------------------------  ------------------  -----  --
fc00:ab00:2:0:1::           End (PSP/USD)     'default':1                      sidmgr              InUse  Y
fc00:ab00:2:0:40::          End.X (PSP/USD)   [Gi0/0/0/0, Link-Local]           isis-1              InUse  Y
fc00:ab00:2:0:41::          End.X (PSP/USD)   [Gi0/0/0/2, Link-Local]           isis-1              InUse  Y
fc00:ab00:2:0:42::          End.X (PSP/USD)   [Gi0/0/0/4, Link-Local]           isis-1              InUse  Y
RP/0/RP0/CPU0:R02#
```

```
RP/0/RP0/CPU0:R02#sh segment-routing srv6 locator MyFullLoc01 internal
Thu Nov  7 11:26:11.810 UTC
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           3        0     fc00:ab00:2::/64          Up
  Index: 3
  Flags: 0x0
  Format: base <40B/24N/16F/8A> (1)
  FSM state: UP
  Global OOR State: Resource Available
  Local SID OOR State: Resource Available
  Interface:
    Name: srv6-MyFullLoc01
    IFH : 0x0000002c
    IPv6 address: fc00:ab00:2::/64 (IM notified: fc00:ab00:2::/64)
    Algo: 0 (IM notified: 0)
  Number of SIDs: 3
  Chkpt Obj ID: 0x2f60
  Number of SID CTX: 4
  Number of Auto SIDs: 1
  TI Object:
    Type: Entry
    Ptr: 0x139839370908128, Producer ID: 0
    Flags:
      Generic:  0x0 ()
      Specific: 0x0 ()
    Modified: Tue Nov 05 14:12:54 UTC 2024 (1d21h ago)
  Created: Nov  5 14:12:55.129 (1d21h ago)
  Event history:
    -- cut --
```

Segment Routing MPLS a SRv6

Mikro SID

RFC draft

SRv6 Network Programming

- SRv6 Network Programming (RFC8986)
  - zdroj zapíše „Network Program“ jako seznam instrukcí
  - První instrukce je v IPv6 Destination Address
  - Zbývající instrukce jsou v Segment Routing Header (SRH)
- Instrukce je nazývána **SID**
- **Container SID** může obsahovat až 6 mikro-instrukcí které jsou nazývány **uSID (Micro-SID)**
  - The IETF term is “NEXT-CSID” (Compressed SID)

Container pro 6 uSID’s

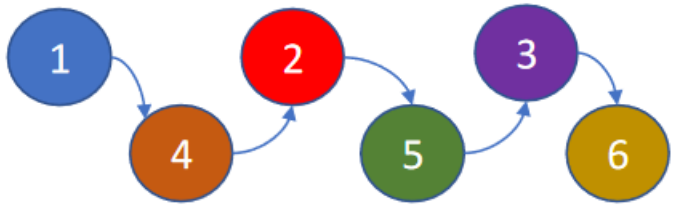
FC00:0000:1111:2222:3333:4444:5555:6666

Min-Cost Block

- Doporučená alokace **uSID** z privátního bloku IPv6 adres (FC/8)
  - FC/8: neroutovatelné mimo vlastní síť - bezpečnější
  - Lze použít i Public adresy
- Nejčastěji použit /32 Block
  - lze použít /16 či /48

uSID Format – Quick review

- Compressed SID. Can be of length 16/32 bits
- Configured on the boxes in the form SRv6 locators
- SRv6 Locator consists of two portions: uSID block, and a uSID.
  - SRv6 Locator→ [uSID Block]:[uSID]::/**L**
- Four different formats defining the length of the uSID block & uSID:
  - F1616: 16-bit uSID block, 16-bit uSID → **L** = 32
  - F3216: 32-bit uSID block, 16-bit uSID → **L** = 48 “Industry Standard”
  - F4816: 48-bit uSID block, 16-bit uSID → **L** = 64
  - F3232: 32-bit uSID block, 32-bit uSID → **L** = 64
- **L** represents SRv6 locator length
- **F3216**: uSID Container with a **32 bit uSID Block + 16 bit uSIDs** making up a TE path



fccc:cc00:0001:0004:0002:0005:0003:0006/128

- Max. of 6 uSIDs per container. More than 6 uSIDs, use of SRH required.

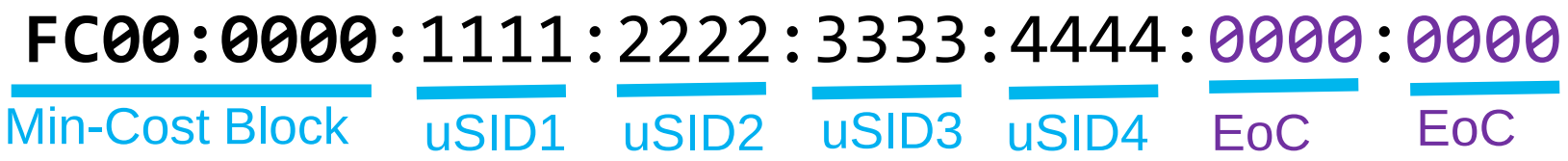
Container of 6 uSID's

FC00:0000:1111:2222:3333:4444:5555:6666

Min-Cost Block | uSID1 | uSID2 | uSID3 | uSID4 | uSID5 | uSID6

- uSID "default size" je 16 bitů ":WXYZ:"
  - lze současně použít uSID-y různé délky
  - např. 32-bits uSID pro „ultra-scale" service (mnoho L2VPN, ...)
- Až 6x uSID v IPv6 dst. adrese 
  - uSID program „čteme" přirozeně z leva do prava
    - jdi na 1111 potom 2222 potom 3333 potom ....

Méně než 6 uSID's v DA



- Nepoužité uSIDy v mikro-programu jsou vyplněny ":0000:"
  - ":0000:" znamená "End of Container" (EoC)

Je-li třeba více než 6 uSID

Outer DA: FC00:0000:0001:0002:0003:0004:0005:0006

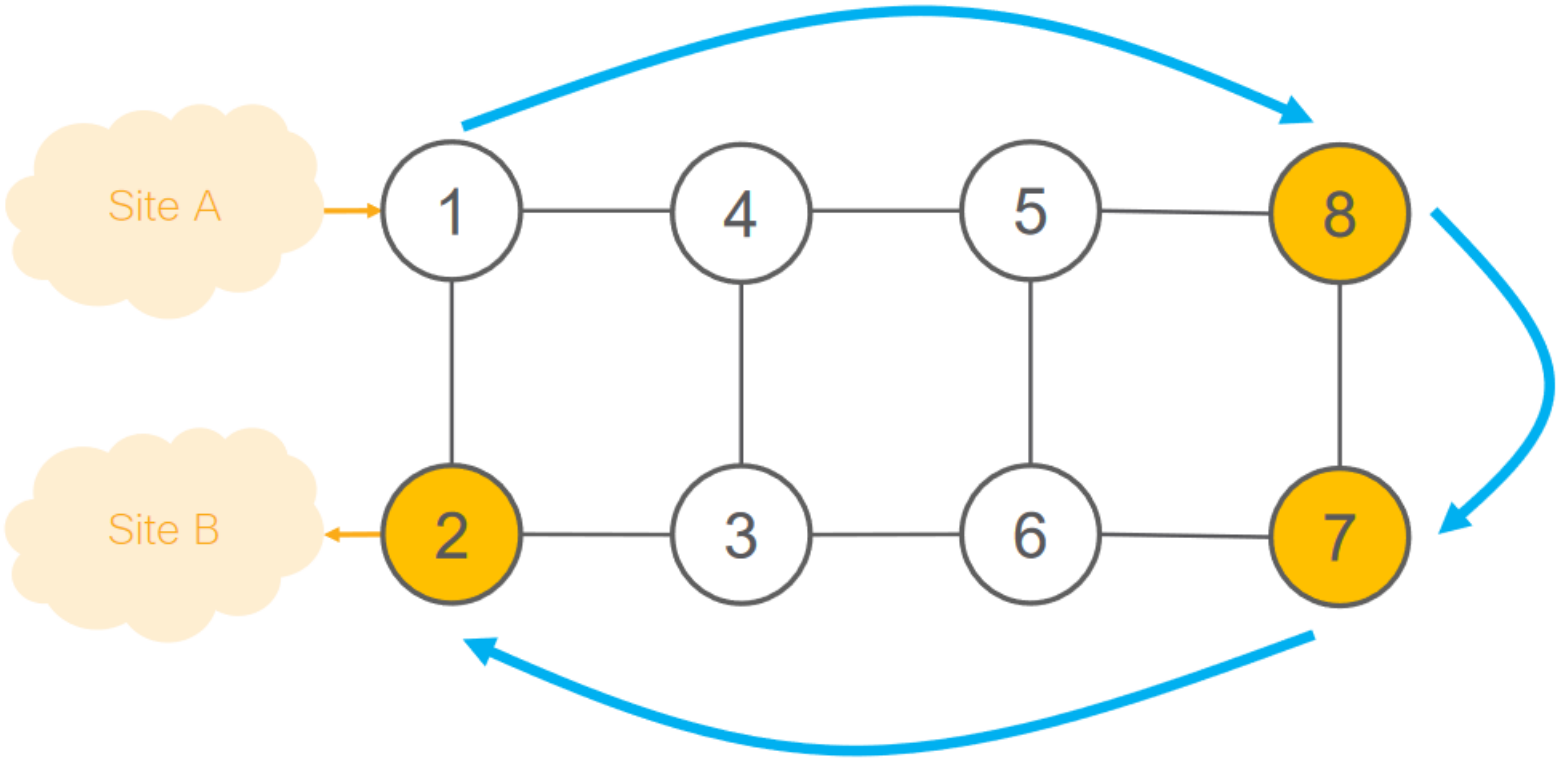
uSID1 uSID2 uSID3 uSID4 uSID5 uSID6

Outer SRH: FC00:0000:0007:0008:0009:0010:0011:0012

uSID7 uSID8 uSID9 uSID10 uSID11 uSID12

- 12 uSID pouze s jedním SRH Container-SID
  - 6x v IPv6 Dst Addr + 6x v SRH
  - pouze 24Bytes MTU overheadu
    - (8B pevná část SRH + 16Bytes za každý 128bit SID)
  - 50% ve srovnání s Full-SID
    - (12*4B = 48B)

Příklad: jdi na 8 potom na 7 potom na 2

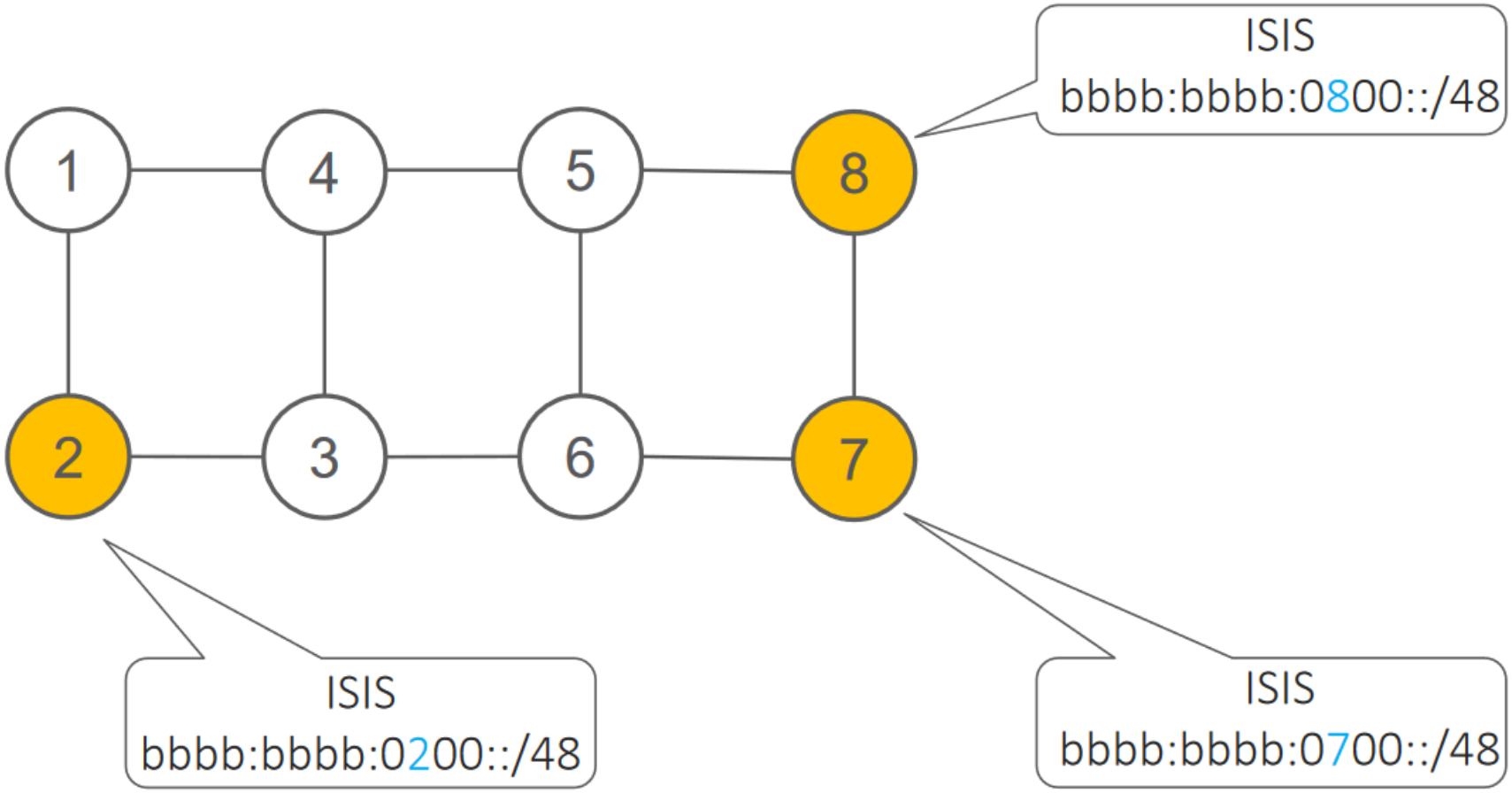


One single micro-program in the DA is enough

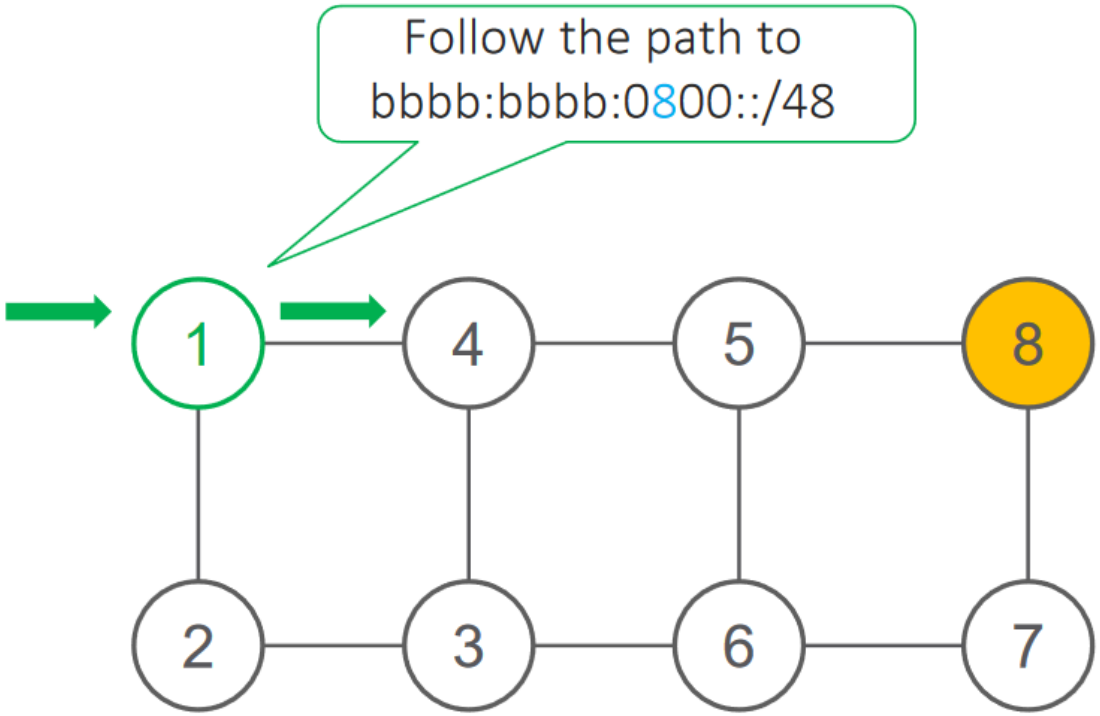
DA = bbbb:bbbb:0800:0700:0200:FDT4:0000:0000

ALEF

Základní IP Routing:

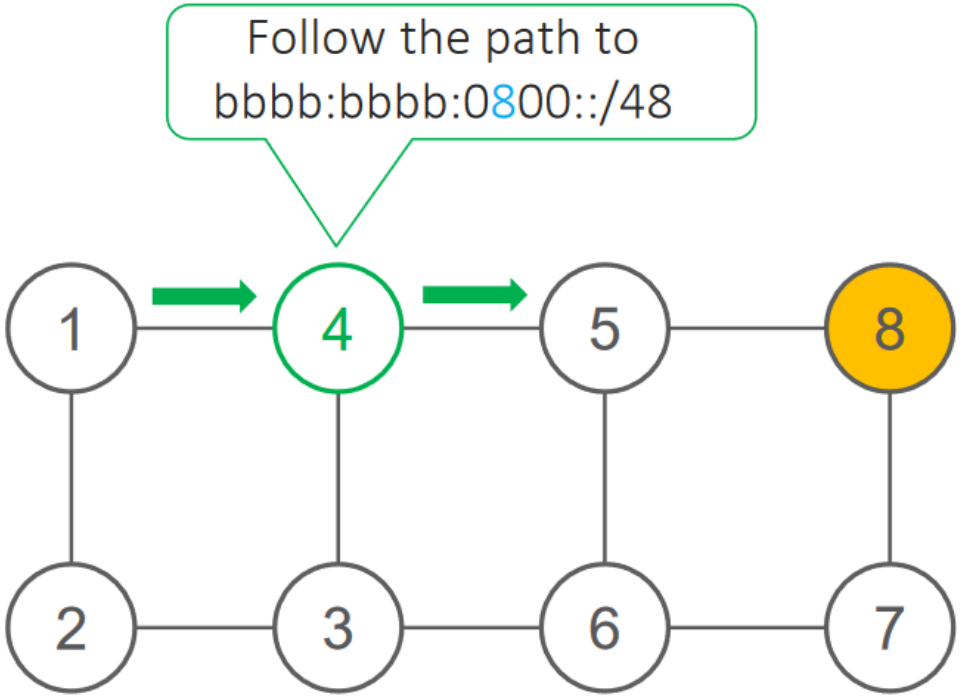


Node 1: klasické IPv6



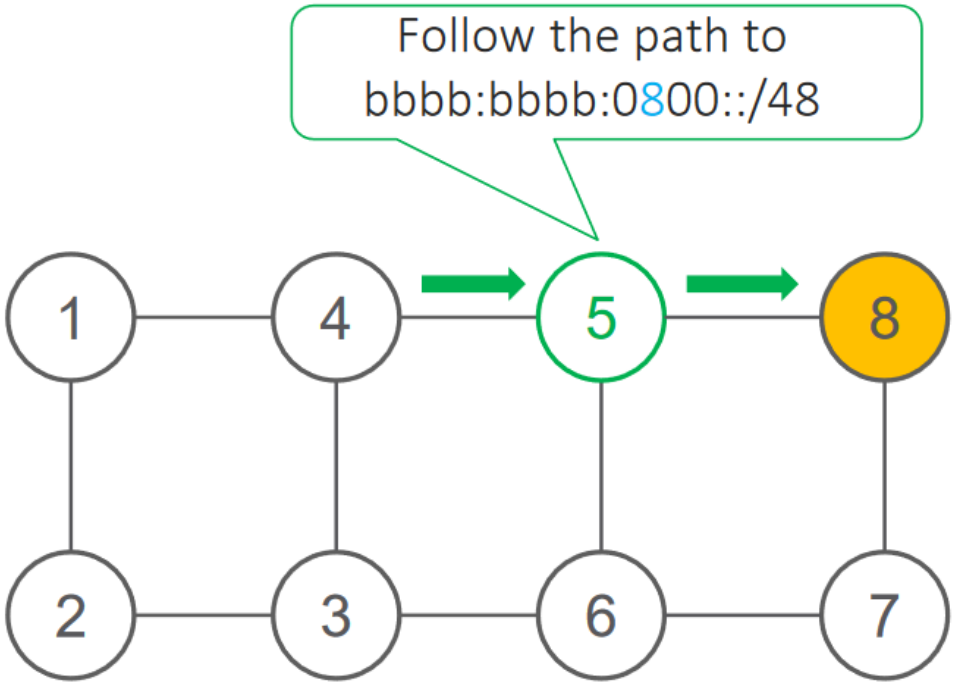
DA = bbbb:bbbb:0800:0700:0200:FDT4:0000:0000

Node 4: klasické IPv6



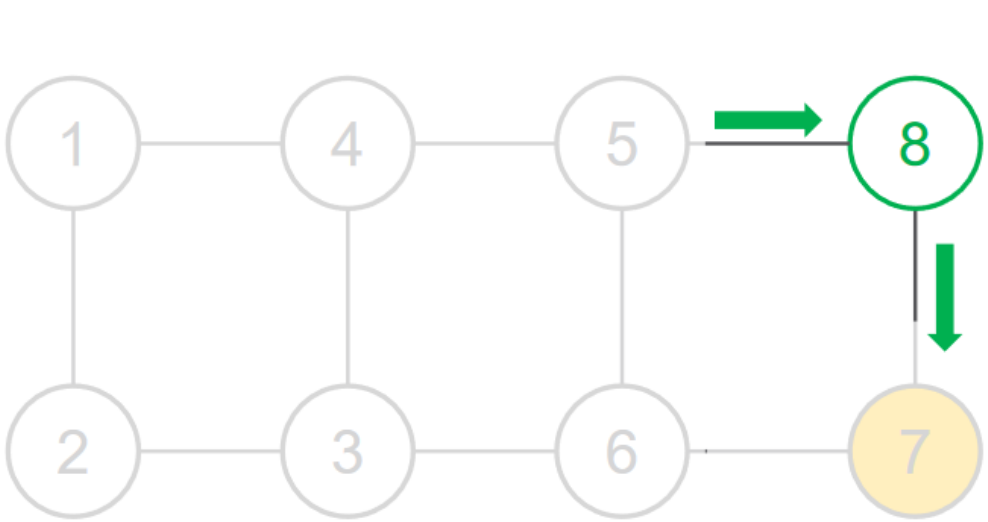
DA = bbbb:bbbb:0800:0700:0200:FDT4:0000:0000

Node 5: klasické IPv6



DA = bbbb:bbbb:0800:0700:0200:FDT4:0000:0000

Node 8: Shift and Forward



Rx'd DA: bbbb:bbbb:0800:0700:0200:FDT4:0000:0000

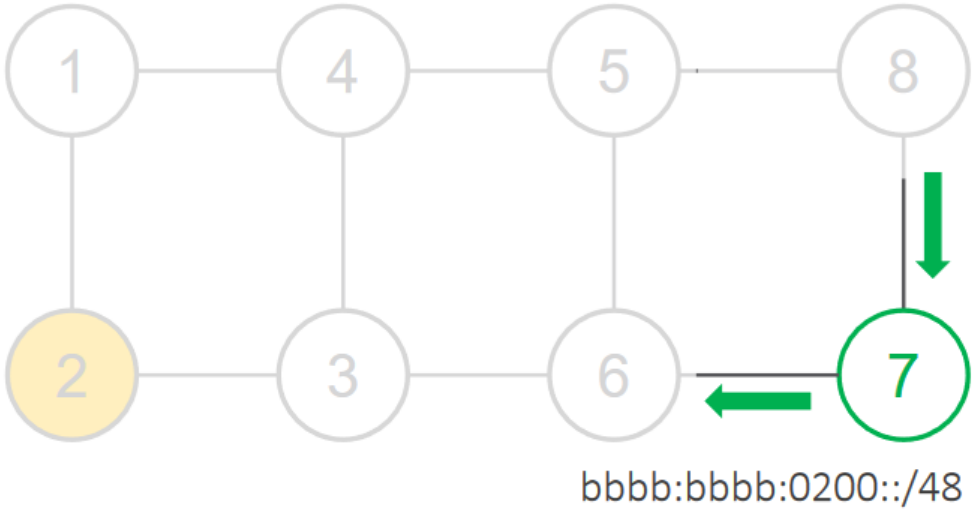
SHIFT << 16

Tx'd DA: bbbb:bbbb:0700:0200:FDT4:0000:0000:0000

bbbb:bbbb:0700::/48

```
FIB Longest-Match bbbb:bbbb:0800::/48 → SRv6 Instruction:
        Shift micro-program by one micro-instruction
        Set last micro-instruction to "End-of-Container"
        Lookup the updated DA and forward
```

Node 7: Shift and Forward



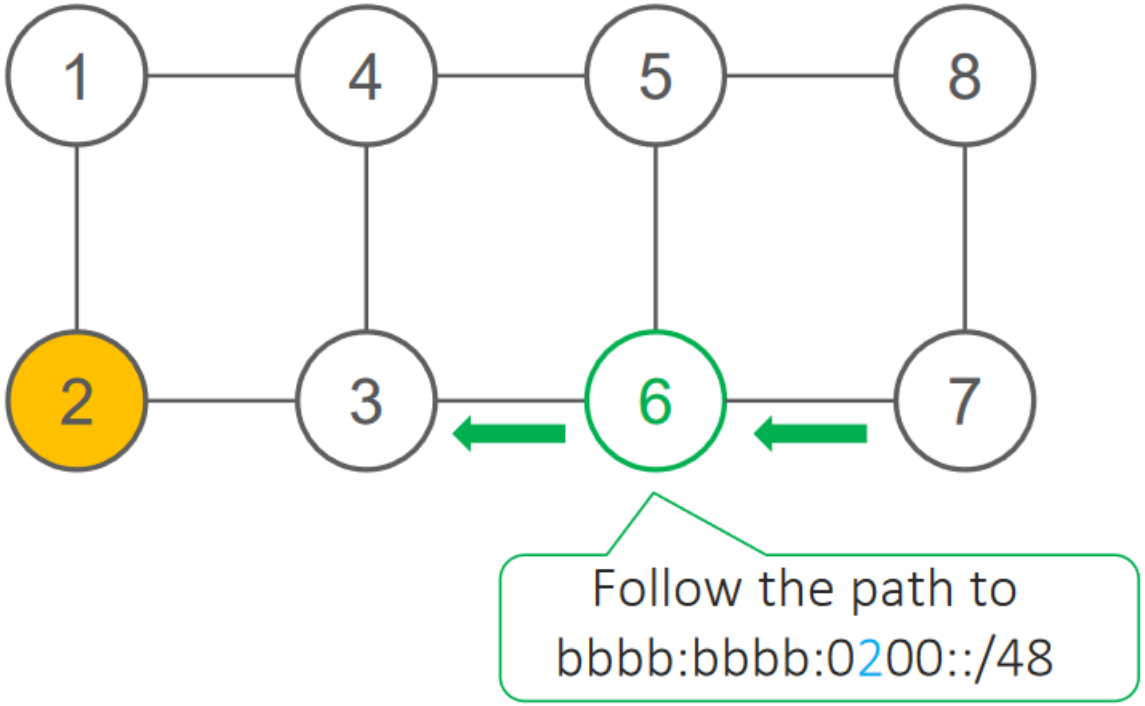
Rx'd DA: bbbb:bbbb:0700:0200:FDT4:0000:0000:0000

SHIFT << 16

Tx'd DA: bbbb:bbbb:0200:FDT4:0000:0000:0000:0000

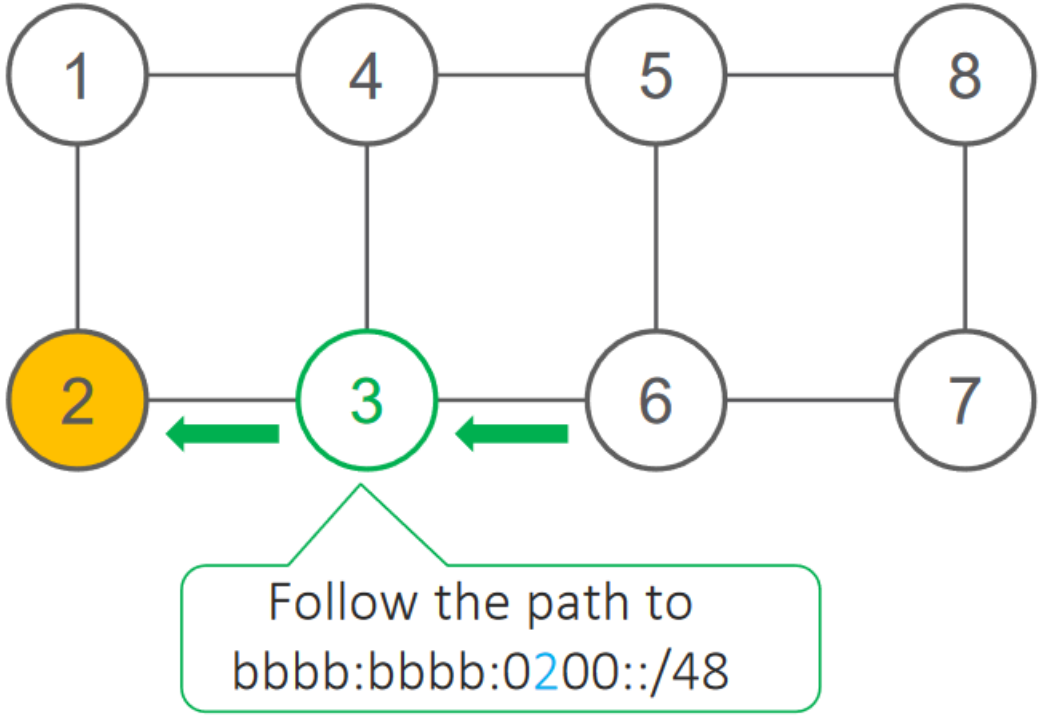
```
FIB Longest-Match bbbb:bbbb:0700::/48 → SRv6 Instruction:
        Shift micro-program by one micro-instruction
        Set last micro-instruction to "End-of-Container"
        Lookup the updated DA and forward
```

Node 6: klasické IPv6



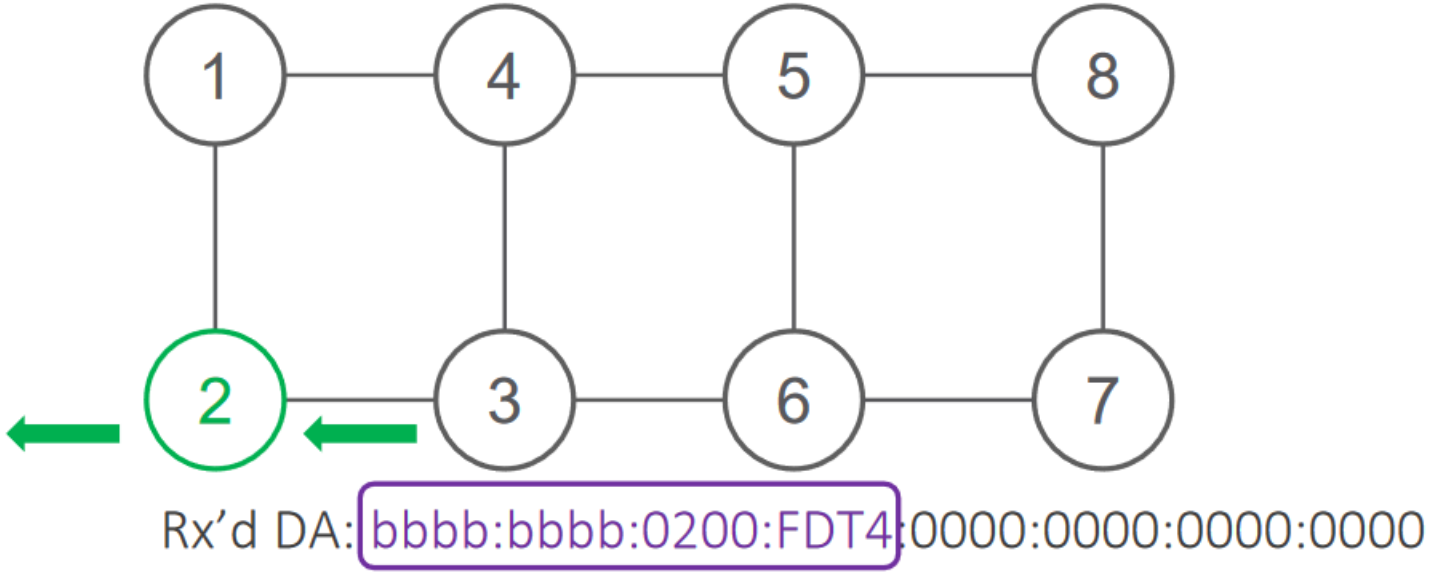
DA = bbbb:bbbb:0200:FDT4:0000:0000:0000:0000

Node 3: klasické IPv6



DA = bbbb:bbbb:0200:FDT4:0000:0000:0000:0000

Node 2: SRv6 End.DT4 behavior




```
FIB Longest-Match bbbb:bbbb:0200:FDT4::/64 → SRv6 Instruction:
  Decapsulate and forward inner IPv4 packet to Site B
```

Micro SID configuration example – IOS-XR

```
!
segment-routing
 srv6
  logging locator status
  locators
   locator Loc_uSID_Algo_0
    micro-segment behavior unode psp-usd
    prefix fc00:ab00:1::/48
   !
!
router isis 1
 is-type level-2-only
 net 49.0001.0000.0000.0002.00
!
 address-family ipv6 unicast
  metric-style wide
  segment-routing srv6
   locator Loc_uSID_Algo_0
    level 2
  !
 !
 interface Loopback0
  passive
  circuit-type level-2-only
  address-family ipv6 unicast
  !
 !
 interface GigabitEthernet0/0/0/0
  circuit-type level-2-only
  point-to-point
  address-family ipv6 unicast
  !
 !
```

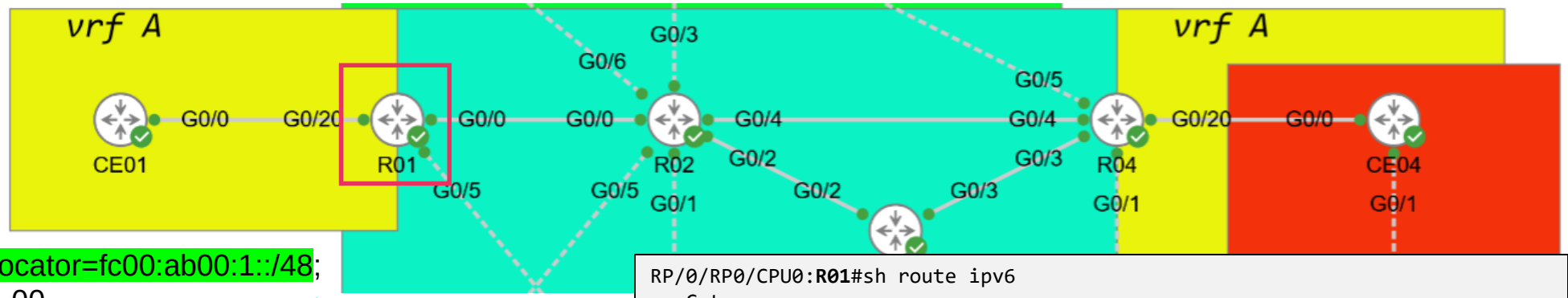
Locator name

Micro-SID is used

Locator value in IPv6 address format

Routing protocol for IPv6 routing

uSID



- R01: Lo0=2001:9999::1/128; Locator=fc00:ab00:1::/48;
  NET=49.0001.0000.0000.0001.00
- R02: Lo0=2001:9999::2/128; Locator=fc00:ab00:2::/48;
  NET=49.0001.0000.0000.0002.00
- R03: Lo0=2001:9999::3/128; Locator=fc00:ab00:3::/48;
  NET=49.0001.0000.0000.0003.00
- R04: Lo0=2001:9999::4/128; Locator=fc00:ab00:4::/48;
  NET=49.0001.0000.0000.0004.00

L3 VPN = vrf A

L2 VPN = EVPN VPWS

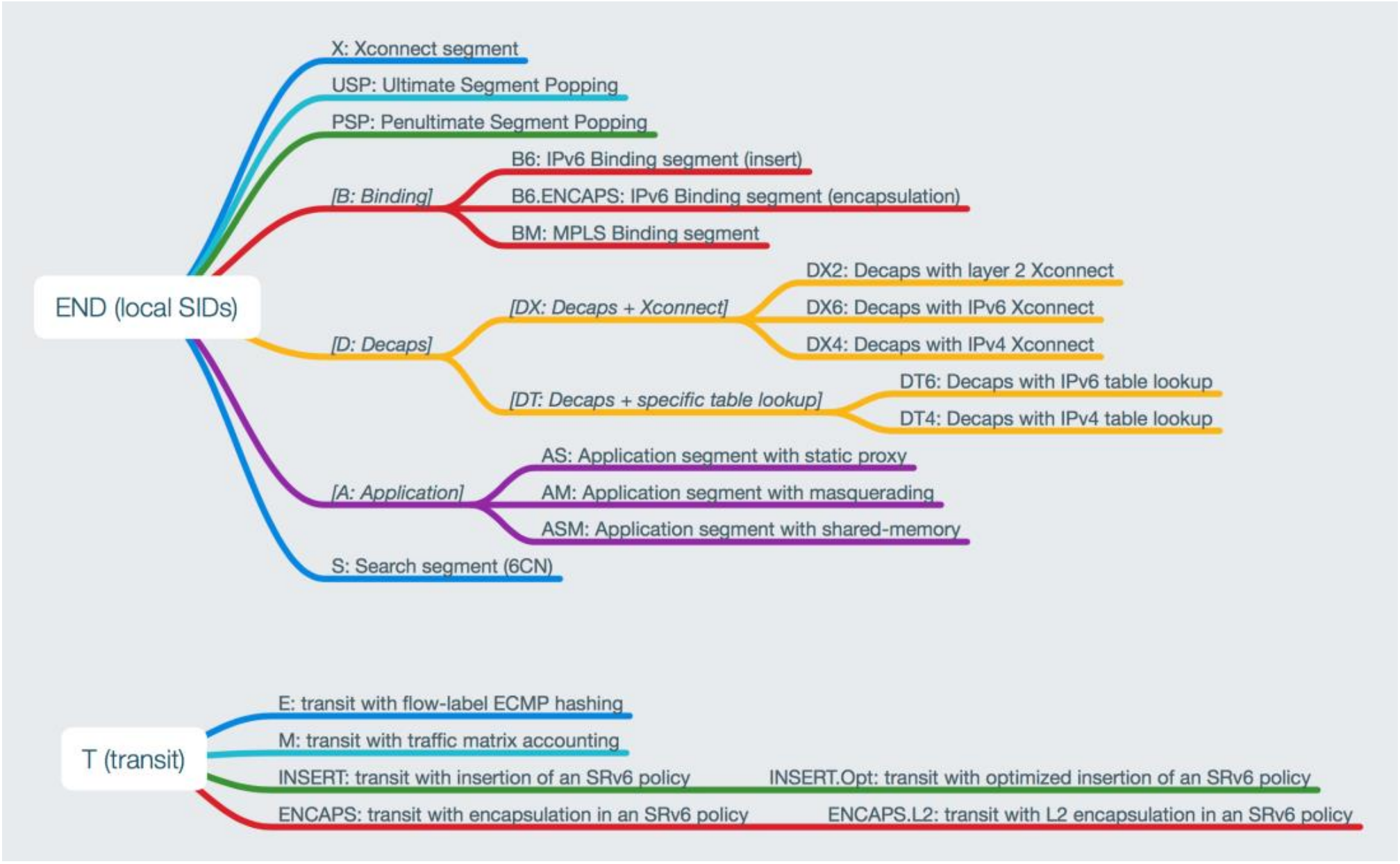
```
RP/0/RP0/CPU0:R01#show segment-routing srv6 locator
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
Loc_uSID_Algo_0       4        0     fc00:ab00:1::/48          Up       U
RP/0/RP0/CPU0:R01#show segment-routing srv6 sid all

*** Locator: 'Loc_uSID_Algo_0' ***

SID                         Behavior          Context                          Owner               Stat
--------------------------  ----------------  -------------------------------  ------------------  ----
fc00:ab00:1::               uN (PSP/USD)      'default':1                      sidmgr              InUs
fc00:ab00:1:e000::          uA (PSP/USD)      [Gi0/0/0/0, Link-Local]:0        isis-1              InUs
fc00:ab00:1:e001::          uDT4              'A'                              bgp-65000           InUs
fc00:ab00:1:e002::          uDX2              14:140104                        l2vpn_srv6          InUs
RP/0/RP0/CPU0:R01#
```

```
RP/0/RP0/CPU0:R01#sh route ipv6
-- Cut --
L    ::ffff:127.0.0.0/104 [0/0] via ::, 1d02h
L    2001:9999::1/128 is directly connected, 1d02h, Loopback0
i L2 2001:9999::2/128
      [115/10] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
i L2 2001:9999::3/128
      [115/20] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
i L2 2001:9999::4/128
      [115/20] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
L    fc00:ab00:1::/48, SRv6 Endpoint uN (shift)
      [0/0] via ::, 1d02h
L    fc00:ab00:1::/64, SRv6 Endpoint uN (PSP/USD)
      [0/0] via ::, 1d02h
L    fc00:ab00:1:e000::/64, SRv6 Endpoint uA (shift)
      [0/0] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
L    fc00:ab00:1:e000::/80, SRv6 Endpoint uA (PSP/USD)
      [0/0] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
L    fc00:ab00:1:e001::/64, SRv6 Endpoint uDT4
      [0/0] via ::ffff:0.0.0.0 (nexthop in vrf A), 22:38:46
L    fc00:ab00:1:e002::/64, SRv6 Endpoint uDX2
      [0/0] via ::, 03:18:07, GigabitEthernet0/0/0/20
i L2 fc00:ab00:2::/48
      [115/11] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
i L2 fc00:ab00:3::/48
      [115/21] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
i L2 fc00:ab00:4::/48
      [115/21] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
L    fc00:ab00:e000::/48, SRv6 Endpoint uA (shift)
      [0/0] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
L    fc00:ab00:e000::/64, SRv6 Endpoint uA (PSP/USD)
      [0/0] via fe80::5054:ff:fe17:af75, 22:40:11, GigabitEthernet0/0/0/0
-- Cut --
RP/0/RP0/CPU0:R01#
```

Funkce a její pojmenování …



SRv6 funkce - příklad

Codename		Behavior	Note
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.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]

https://www.iana.org/assignments/segment-routing/segment-routing.xhtml

uSID - výhody

- Data Plane
  - Menší MTU overhead (6 uSIDs bez nutnosti použít SRH)
  - Hardware-friendly (je „linerate“ na merchant silicon NPU)
- Control Plane
  - Dostatečné množství globálně unikátních uSID v doméně
- IP Power
  - zisk výkonem IP sumarizace a využitím „longest prefix match“
  - 2-3x efektivnější z pohledu HW zdrojů ve FIB než klasické MPLS
  - díky využití *IPv6 Flow label* výrazně efektivnější loadbalancing

Podpora SRv6 uSID

Network Equipment Manufacturers

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

Merchant Silicon

Broadcom, Marvell, Barefoot Networks (Intel)

SR v6
Segment Routing

Open-Source Networking Stacks

Linux, fd.io, P4, OPEN Compute Project, Kubernetes, SONiC, FRRouting

Smart NIC

Intel, Mellanox Technologies

Open-Source Applications

Pyroute2, IPTABLES SERA, GoBGP, exaBGP, eBPF, cilium, TCPDUMP, WIRESHARK, SNORT, GoBMP

Partners

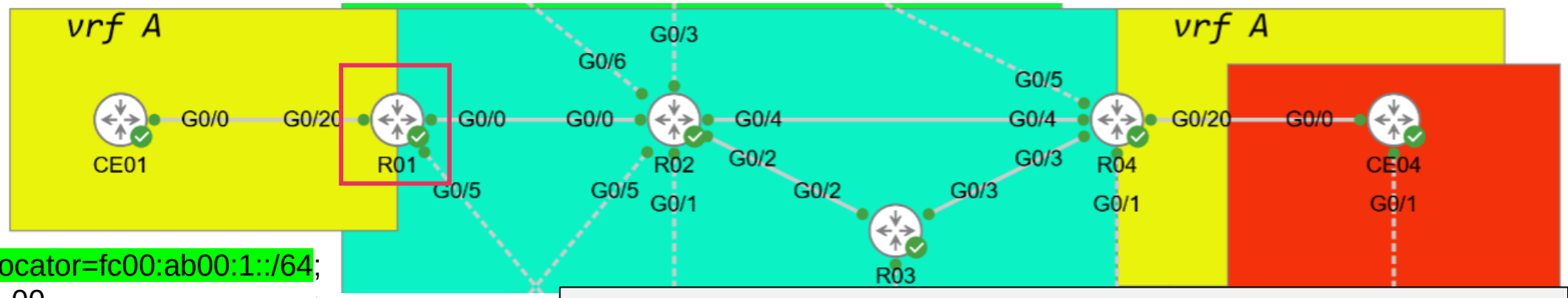
kaloom, Trend Micro, ENEA Qosmos Division, Keysight Technologies, NoviFlow, ixia A Keysight Business, Spirent Promise Assured

ALEF

Segment Routing MPLS a SRv6

Příklady a výpisy

Full SID



- R01: Lo0=2001:9999::1/128; Locator=fc00:ab00:1::/64;
  NET=49.0001.0000.0000.0001.00
- R02: Lo0=2001:9999::2/128; Locator=fc00:ab00:2::/64;
  NET=49.0001.0000.0000.0002.00
- R03: Lo0=2001:9999::3/128; Locator=fc00:ab00:3::/64;
  NET=49.0001.0000.0000.0003.00
- R04: Lo0=2001:9999::4/128; Locator=fc00:ab00:4::/64;
  NET=49.0001.0000.0000.0004.00

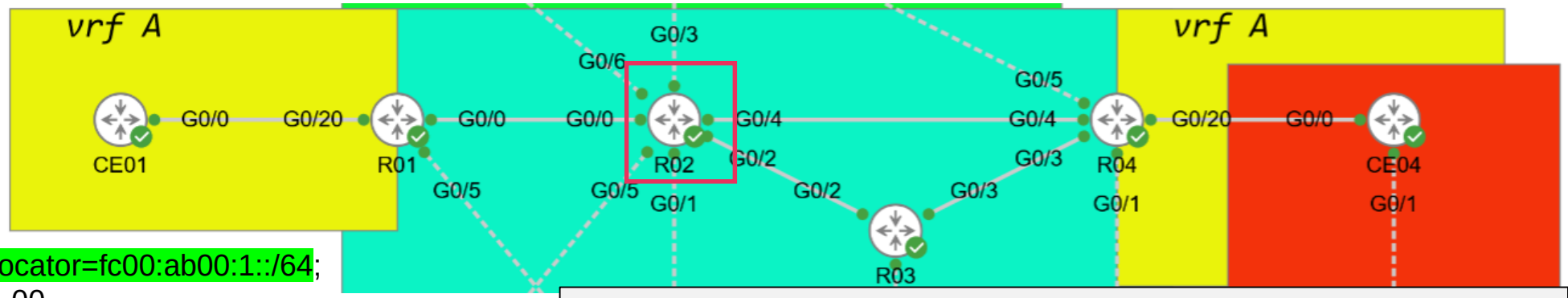
```
RP/0/RP0/CPU0:R01#sh route ipv6
-- Cut --
L     ::ffff:127.0.0.0/104
       [0/0] via ::, 02:46:07
L     2001:9999::1/128 is directly connected,
       01:54:38, Loopback0
i L2 2001:9999::2/128
       [115/10] via fe80::5054:ff:fe17:af75, 01:53:07, GigabitEthernet0/0/0/0
i L2 2001:9999::3/128
       [115/20] via fe80::5054:ff:fe17:af75, 01:51:57, GigabitEthernet0/0/0/0
i L2 2001:9999::4/128
       [115/20] via fe80::5054:ff:fe17:af75, 01:50:30, GigabitEthernet0/0/0/0
L     fc00:ab00:1:0:1::/128, SRv6 Endpoint End (PSP/USD)
       [0/0] via ::, 02:48:04
L     fc00:ab00:1:0:40::/128, SRv6 Endpoint End.X (PSP/USD)
       [0/0] via fe80::5054:ff:fe17:af75, 02:45:19, GigabitEthernet0/0/0/0
i L2 fc00:ab00:2::/64
       [115/11] via fe80::5054:ff:fe17:af75, 01:53:07, GigabitEthernet0/0/0/0
i L2 fc00:ab00:3::/64
       [115/21] via fe80::5054:ff:fe17:af75, 01:51:57, GigabitEthernet0/0/0/0
i L2 fc00:ab00:4::/64
       [115/21] via fe80::5054:ff:fe17:af75, 01:50:30, GigabitEthernet0/0/0/0
-- Cut --
RP/0/RP0/CPU0:R01#
```

```
RP/0/RP0/CPU0:R01#show segment-routing srv6 locator
Name                  ID       Algo  Prefix                    Status    Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           1        0     fc00:ab00:1::/64          Up
RP/0/RP0/CPU0:R01#show segment-routing srv6 sid all

*** Locator: 'MyFullLoc01' ***

SID                         Behavior          Context                          Owner               State  RW
--------------------------  ----------------  -------------------------------  ------------------  -----  --
fc00:ab00:1:0:1::           End (PSP/USD)     'default':1                      sidmgr              InUse  Y
fc00:ab00:1:0:40::          End.X (PSP/USD)   [Gi0/0/0/0, Link-Local]          isis-1              InUse  Y
RP/0/RP0/CPU0:R01#
```

Full SID



- R01: Lo0=2001:9999::1/128; Locator=fc00:ab00:1::/64;
  NET=49.0001.0000.0000.0001.00
- R02: Lo0=2001:9999::2/128; Locator=fc00:ab00:2::/64;
  NET=49.0001.0000.0000.0002.00
- R03: Lo0=2001:9999::3/128; Locator=fc00:ab00:3::/64;
  NET=49.0001.0000.0000.0003.00
- R04: Lo0=2001:9999::4/128; Locator=fc00:ab00:4::/64;
  NET=49.0001.0000.0000.0004.00

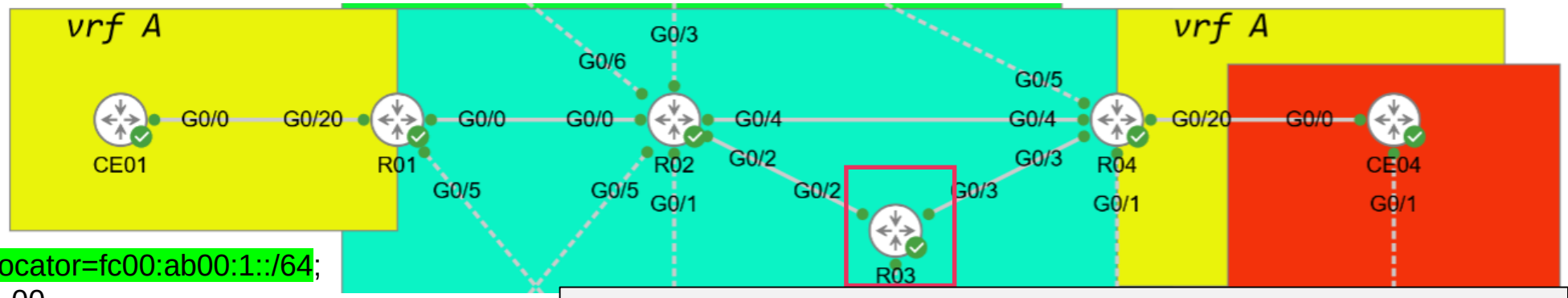
```
RP/0/RP0/CPU0:R02#show segment-routing srv6 locator
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           1        0     fc00:ab00:2::/64          Up
RP/0/RP0/CPU0:R02#show segment-routing srv6 sid all

*** Locator: 'MyFullLoc01' ***

SID                         Behavior          Context                          Owner
--------------------------  ----------------  -------------------------------  --------
fc00:ab00:2:0:1::           End (PSP/USD)     'default':1                      sidmgr
fc00:ab00:2:0:40::          End.X (PSP/USD)   [Gi0/0/0/2, Link-Local]          isis-1
fc00:ab00:2:0:41::          End.X (PSP/USD)   [Gi0/0/0/0, Link-Local]          isis-1
fc00:ab00:2:0:42::          End.X (PSP/USD)   [Gi0/0/0/4, Link-Local]          isis-1
RP/0/RP0/CPU0:R02#
```

```
RP/0/RP0/CPU0:R02#sh route ipv6
-- Cut --
L     ::ffff:127.0.0.0/104
      [0/0] via ::, 03:21:51
i L2 2001:9999::1/128
      [115/10] via fe80::5054:ff:fe14:94d4, 02:29:44, GigabitEthernet0/0/0/0
L     2001:9999::2/128 is directly connected,
      02:28:14, Loopback0
i L2 2001:9999::3/128
      [115/10] via fe80::5054:ff:fe1c:96f4, 02:27:03, GigabitEthernet0/0/0/2
i L2 2001:9999::4/128
      [115/10] via fe80::5054:ff:fe09:6ab0, 02:25:36, GigabitEthernet0/0/0/4
i L2 fc00:ab00:1::/64
      [115/11] via fe80::5054:ff:fe14:94d4, 02:29:44, GigabitEthernet0/0/0/0
L     fc00:ab00:2:0:1::/128, SRv6 Endpoint End (PSP/USD)
      [0/0] via ::, 03:23:20
L     fc00:ab00:2:0:40::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe1c:96f4, 03:20:25, GigabitEthernet0/0/0/2
L     fc00:ab00:2:0:41::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe14:94d4, 03:20:25, GigabitEthernet0/0/0/0
L     fc00:ab00:2:0:42::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe09:6ab0, 03:17:50, GigabitEthernet0/0/0/4
i L2 fc00:ab00:3::/64
      [115/11] via fe80::5054:ff:fe1c:96f4, 02:27:03, GigabitEthernet0/0/0/2
i L2 fc00:ab00:4::/64
      [115/11] via fe80::5054:ff:fe09:6ab0, 02:25:36, GigabitEthernet0/0/0/4
-- Cut --
RP/0/RP0/CPU0:R02#
```

Full SID



- R01: Lo0=2001:9999::1/128; Locator=fc00:ab00:1::/64;
  NET=49.0001.0000.0000.0001.00
- R02: Lo0=2001:9999::2/128; Locator=fc00:ab00:2::/64;
  NET=49.0001.0000.0000.0002.00
- R03: Lo0=2001:9999::3/128; Locator=fc00:ab00:3::/64;
  NET=49.0001.0000.0000.0003.00
- R04: Lo0=2001:9999::4/128; Locator=fc00:ab00:4::/64;
  NET=49.0001.0000.0000.0004.00

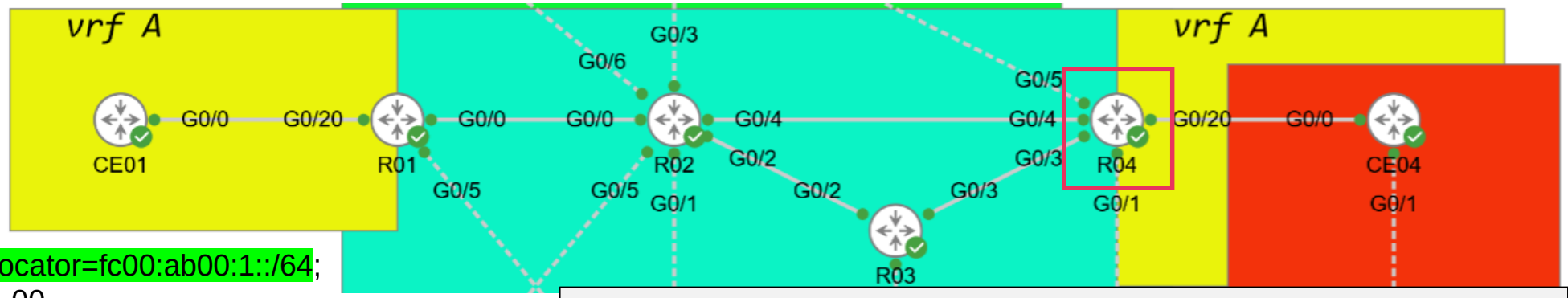
```
RP/0/RP0/CPU0:R03#sh route ipv6
-- Cut --
L     ::ffff:127.0.0.0/104
      [0/0] via ::, 03:47:48
i L2 2001:9999::1/128
      [115/20] via fe80::5054:ff:fe03:dd9a, 02:55:45, GigabitEthernet0/0/0/2
i L2 2001:9999::2/128
      [115/10] via fe80::5054:ff:fe03:dd9a, 02:54:15, GigabitEthernet0/0/0/2
L     2001:9999::3/128 is directly connected,
      02:53:05, Loopback0
i L2 2001:9999::4/128
      [115/10] via fe80::5054:ff:fe11:2f5e, 02:51:37, GigabitEthernet0/0/0/3
i L2 fc00:ab00:1::/64
      [115/21] via fe80::5054:ff:fe03:dd9a, 02:55:45, GigabitEthernet0/0/0/2
i L2 fc00:ab00:2::/64
      [115/11] via fe80::5054:ff:fe03:dd9a, 02:54:15, GigabitEthernet0/0/0/2
L     fc00:ab00:3:0:1::/128, SRv6 Endpoint End (PSP/USD)
      [0/0] via ::, 03:49:21
L     fc00:ab00:3:0:40::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe03:dd9a, 03:46:26, GigabitEthernet0/0/0/2
L     fc00:ab00:3:0:41::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe11:2f5e, 03:43:51, GigabitEthernet0/0/0/3
i L2 fc00:ab00:4::/64
      [115/11] via fe80::5054:ff:fe11:2f5e, 02:51:37, GigabitEthernet0/0/0/3
-- Cut --
RP/0/RP0/CPU0:R03#
```

```
RP/0/RP0/CPU0:R03#show segment-routing srv6 locator
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           1        0     fc00:ab00:3::/64          Up
RP/0/RP0/CPU0:R03#show segment-routing srv6 sid all

*** Locator: 'MyFullLoc01' ***

SID                         Behavior          Context                          Owner               State  RW
--------------------------  ----------------  -------------------------------  ------------------  -----  --
fc00:ab00:3:0:1::           End (PSP/USD)     'default':1                      sidmgr              InUse  Y
fc00:ab00:3:0:40::          End.X (PSP/USD)   [Gi0/0/0/2, Link-Local]           isis-1              InUse  Y
fc00:ab00:3:0:41::          End.X (PSP/USD)   [Gi0/0/0/3, Link-Local]           isis-1              InUse  Y
RP/0/RP0/CPU0:R03#
```

Full SID



- R01: Lo0=2001:9999::1/128; Locator=fc00:ab00:1::/64;
  NET=49.0001.0000.0000.0001.00
- R02: Lo0=2001:9999::2/128; Locator=fc00:ab00:2::/64;
  NET=49.0001.0000.0000.0002.00
- R03: Lo0=2001:9999::3/128; Locator=fc00:ab00:3::/64;
  NET=49.0001.0000.0000.0003.00
- R04: Lo0=2001:9999::4/128; Locator=fc00:ab00:4::/64;
  NET=49.0001.0000.0000.0004.00

```
RP/0/RP0/CPU0:R04#sh route ipv6
-- Cut --
L     ::ffff:127.0.0.0/104
      [0/0] via ::, 04:39:49
i L2 2001:9999::1/128
      [115/20] via fe80::5054:ff:fe05:ee55, 03:50:33, GigabitEthernet0/0/0/4
i L2 2001:9999::2/128
      [115/10] via fe80::5054:ff:fe05:ee55, 03:49:03, GigabitEthernet0/0/0/4
i L2 2001:9999::3/128
      [115/10] via fe80::5054:ff:fe1f:b0f5, 03:47:53, GigabitEthernet0/0/0/3
L     2001:9999::4/128 is directly connected,
      03:46:26, Loopback0
i L2 fc00:ab00:1::/64
      [115/21] via fe80::5054:ff:fe05:ee55, 03:50:33, GigabitEthernet0/0/0/4
i L2 fc00:ab00:2::/64
      [115/11] via fe80::5054:ff:fe05:ee55, 03:49:03, GigabitEthernet0/0/0/4
i L2 fc00:ab00:3::/64
      [115/11] via fe80::5054:ff:fe1f:b0f5, 03:47:53, GigabitEthernet0/0/0/3
L     fc00:ab00:4:0:1::/128, SRv6 Endpoint End (PSP/USD)
      [0/0] via ::, 04:43:52
L     fc00:ab00:4:0:40::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe05:ee55, 04:38:40, GigabitEthernet0/0/0/4
L     fc00:ab00:4:0:41::/128, SRv6 Endpoint End.X (PSP/USD)
      [0/0] via fe80::5054:ff:fe1f:b0f5, 04:38:40, GigabitEthernet0/0/0/3
-- Cut --
RP/0/RP0/CPU0:R04#
```

```
RP/0/RP0/CPU0:R04#sh segment-routing srv6 locator
Wed Oct  2 13:11:57.420 UTC
Name                  ID       Algo  Prefix                    Status   Flags
--------------------  -------  ----  ------------------------  -------  --------
MyFullLoc01           1        0     fc00:ab00:4::/64          Up
RP/0/RP0/CPU0:R04#sh segment-routing srv6 sid all
Wed Oct  2 13:12:02.846 UTC

*** Locator: 'MyFullLoc01' ***

SID                         Behavior          Context                          Owner               State  RW
--------------------------  ----------------  -------------------------------  ------------------  -----  --
fc00:ab00:4:0:1::           End (PSP/USD)     'default':1                      sidmgr              InUse  Y
fc00:ab00:4:0:40::          End.X (PSP/USD)   [Gi0/0/0/4, Link-Local]           isis-1              InUse  Y
fc00:ab00:4:0:41::          End.X (PSP/USD)   [Gi0/0/0/3, Link-Local]           isis-1              InUse  Y
RP/0/RP0/CPU0:R04#
```

L3VPN přes SRv6 (Cisco IOS-XR)

- BGP Control Plane - VPNv4 (VPNv6) AF (stejné jako u MPLS)
  - Internal BGP, BGP RR
  - IPv4 or IPv6 transport (BGP session)
- VRF config
  - stejný jako pro MPLS/SR based L3 VPN
- PE-CE routing
  - stejný jako pro MPLS/SR based L3 VPN
- VRF config v části BGP

```
router bgp 100
 vrf A
  rd 65000:1
  address-family ipv4 unicast
   segment-routing srv6
    locator My_locator
    alloc mode per-vrf
   !
   redistribute connected
```

L2VPN přes SRv6 (Cisco IOS-XR)

- Podpora pouze pro EVPN
  - Pseudowire potřebuje MPLS dataplane
  - VPLS potřebuje Pseudowire a bridge-domain
- BGP Control Plane – `l2vpn evpn` address-family
- EVPN config

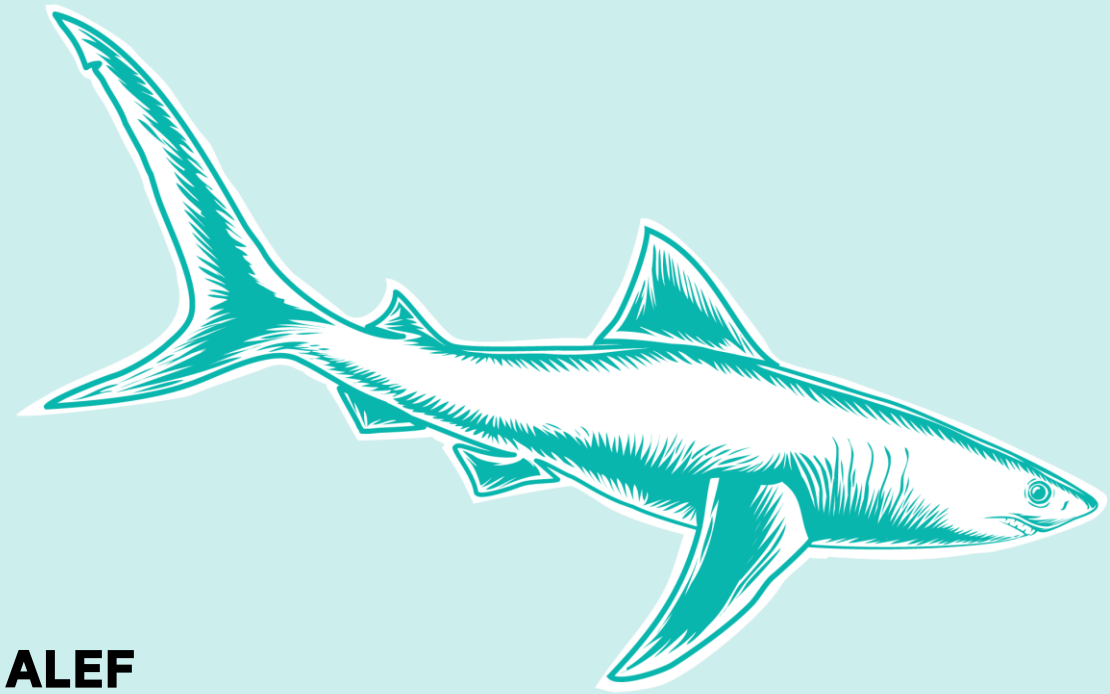
```
evpn
 interface TenGigE0/0/0/0
 !
 segment-routing srv6
  locator My_Locator
 !
!
```

```
router bgp 100
 bgp router-id 4.4.4.4
 address-family l2vpn evpn
 !
 neighbor fcbb:bb00:1::1
  remote-as 100
  update-source Loopback0
 address-family l2vpn evpn
  !
```

- L2VPN config

```
l2vpn
 xconnect group 2
  p2p 2
   interface TenGigE0/0/0/0.2
   neighbor evpn evi 2 service 2 segment-routing srv6
   !
```

Děkuji za pozornost



ALEF