SRv6

Technology and Deployment Use-Cases

Clarence Filsfils
Cisco Fellow – cf@cisco.com

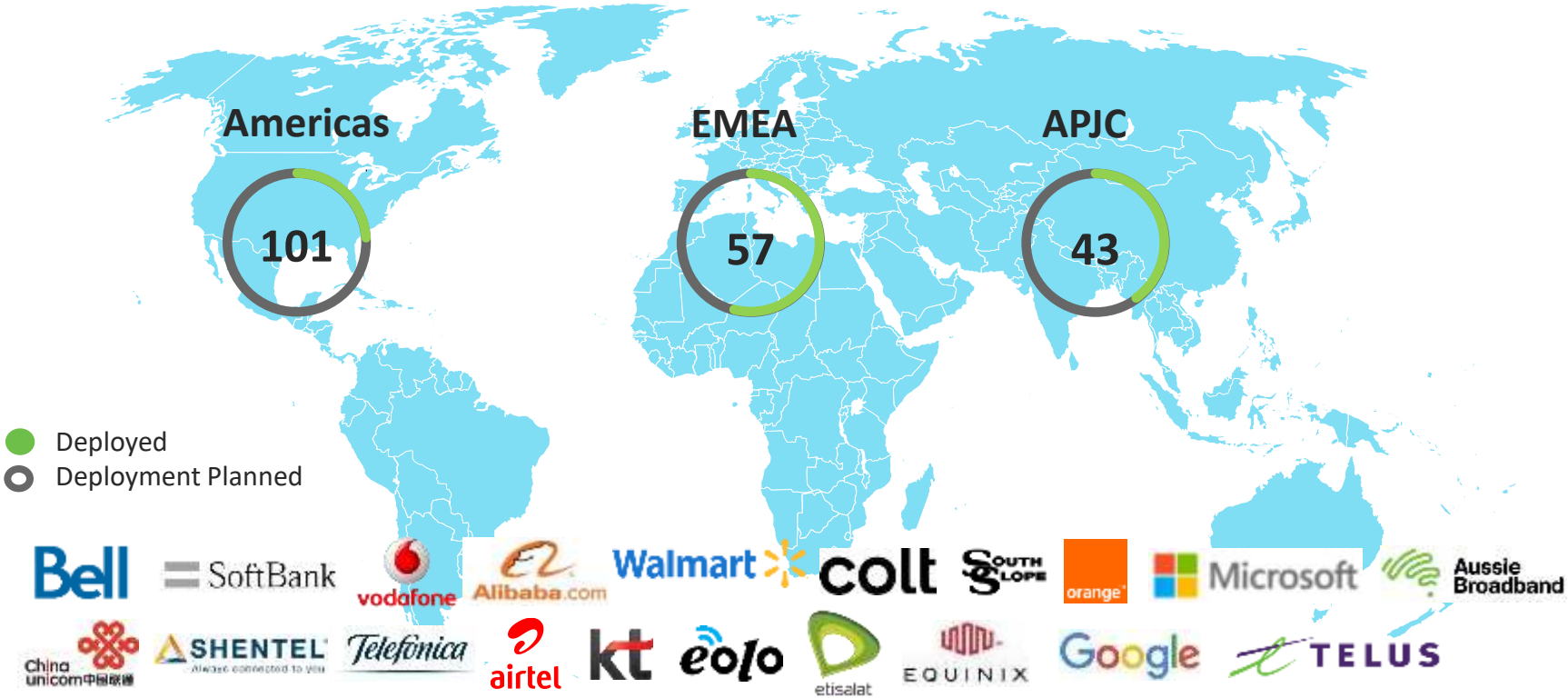
Acknowledgement

- Lead operators / co-development
  - Dennis Cai — Alibaba
  - John Leddy — Akamai
  - Satoru Matsushima — Softbank
  - Sébastien Parisot — Iliad/Free
  - Dan Bernier and Dan Voyer — Bell Canada
- Eco-System Partners
  - Barefoot, Broadcom, Huawei, Intel, Marvell, Mellanox…
- Academic and Open-Source partners
- Cisco SR team

An extended recording
of this presentation is
available at

segment-routing.net/srv6-status/

SR-MPLS: de-facto IPv4 solution



Let’s focus on SRv6: SR for IPv6

Simplicity Always Prevails

~~LDP~~

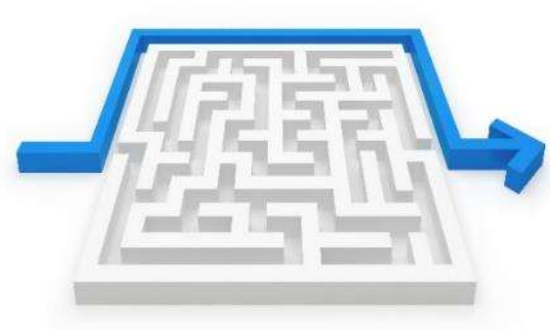
~~RSVP-TE~~

~~Inter-AS Option A/B/C~~

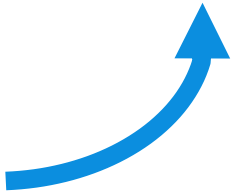
~~MPLS~~

~~UDP/VxLAN~~

~~NSH~~



Furthermore with more scale and functionality



SRv6 Eco-System

At record speed

- In 2019: 8 large-scale commercial deployments
  - Softbank, Iliad, China Telecom, LINE corporation, China Unicom, CERNET2, China Bank and Uganda MTN.
- 18 HW linerate implementations
  - Cisco Systems, Huawei
  - Broadcom, Barefoot, Intel, Marvell, Mellanox
  - Multiple Interop Reports
- 11 open-source platforms/ Applications
  - Linux, FD.io VPP, P4, Wireshark, tcpdump, iptables, nftables, snort, ExaBGP, Contiv-VPP

SoftBank Group

- First commercial SRv6 deployment

2018年3月16日

5Gデータ送信 効率化へ新規格

ソフトバンク、SRv6標準化提案

シンプルで低コスト

材料開発ナノ解析

CTC、クラウドで一貫支援

https://newsroom.cisco.com/press-release-content?type=webcontent&articleId=1969030

iliad

- Nationwide deployment in Italy
- 1000 Cisco NCS 5500
- **1800 Iliad Nodeboxes**

Re: [spring] SPRING SRv6 Deployment Status draft

Sébastien Parisot <sparisot@free-mobile.fr> | Tue, 10 December 2019 09:34 UTC | Show header

Hi Satoru, Zafar,

I would like to provide an update to SRv6 deployment in Iliad's nationwide network in Italy.

As of the end of 2019, the SRv6 network consists of:
- 1000 Cisco NCS 5500 routers
- 1800 Iliad's Nodeboxes
- The network services 4.5 million mobile subscribers (as of Q3 2019)
- The network is carrying 300 Gbps of commercial traffic at peak hours
- It is expected to grow to more than 4000 Nodeboxes in 2020.

The following SRv6 features have been deployed:
- A Segment Routing Header based data plane
- End (PSP), End.X (PSP), End.DT4, T.Encaps.Red, T.Insert.Red functions
- BGP VPN SRv6 extensions
- ISIS SRv6 extensions
- SRH-based Topology Independent (TI-LFA) Fast Reroute mechanisms
- Support for ping and traceroute

Can you please update the SRv6 deployment draft accordingly?

Thanks,
Sébastien

SRv6 Ecosystem

Network Equipment Manufacturers

CISCO

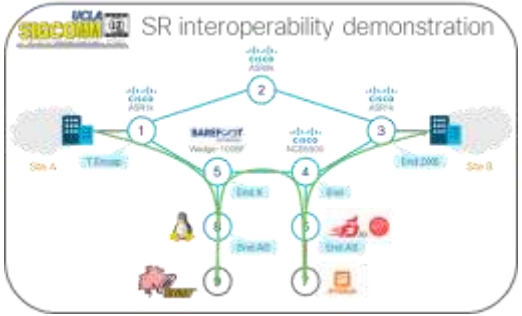
HUAWEI

Merchant Silicon

BAREFOOT NETWORKS

BROADCOM

MARVELL



Open-Source Applications

Pyroute2

SERA IPTABLES

SNORT

TCPDUMP

Contiv

WIRESHARK

exa

Open-Source Networking Stacks

.io

Smart NIC

intel

Mellanox TECHNOLOGIES

NFV Partners

TREND MICRO

ENEA Qosmos Division

SRv6 is a Proposed Standard

- RFC 8402 - Architecture
  - Defines SRv6 with SRH and SRv6 SID’s
- RFC xxxx – SR Extension Header (SRH)
  - Defines the SRv6 dataplane encapsulation
- Last-Call status
  - Net Pgm
  - ISIS
  - OAM

Shipping: Cisco NCS5500, NCS560, NCS540, ASR9k

- ISIS
  - TILFA and uLoop
  - Flex-Algo (Low-Delay Slice) with TILFA
- BGP
  - PIC Core/Edge
  - L3VPN (IPv4)
  - Internet (IPv4)
  - eVPN VPWS
- SRv6-SR-MPLS Gateway
- OAM
  - Ping
  - Trace
  - SID Verification

Shipping: DC – Cisco Nexus 9K GX series

Nexus 9K Platforms

- 16 X 400G
- 28x100G+8x400G
- 64x100G

N9K-C9316D-GX

N9K-C93600CD-GX

N9K-C9364C-GX

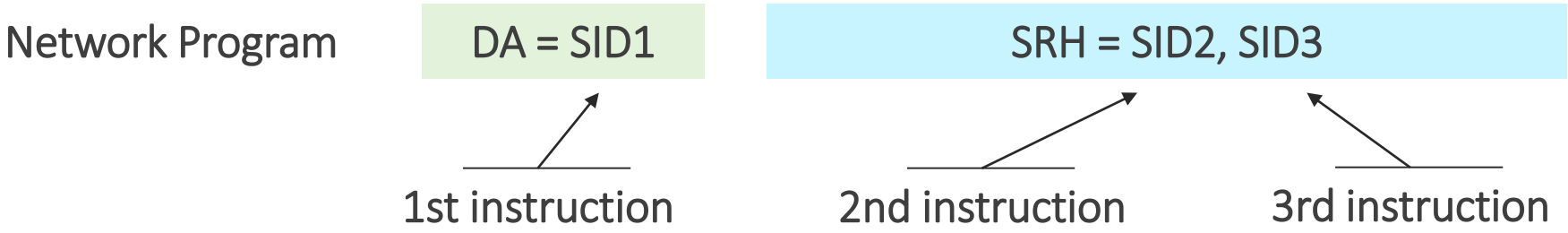
SRv6 forwarding performance

- 400G Line-rate Performance for SRv6
- 6.4 TBPS Packet Processing
- Insert up to 9 SIDs @ line-rate

- IS-IS, OSPFv3
- BGP
  - L3VPN (IPv4, IPv6)
  - Internet (IPv4,IPv6)
- VxLAN – SRv6 gateway
- OAM
  - Ping
  - Trace
  - SID verification

SRv6 - Reminder

Network Programming



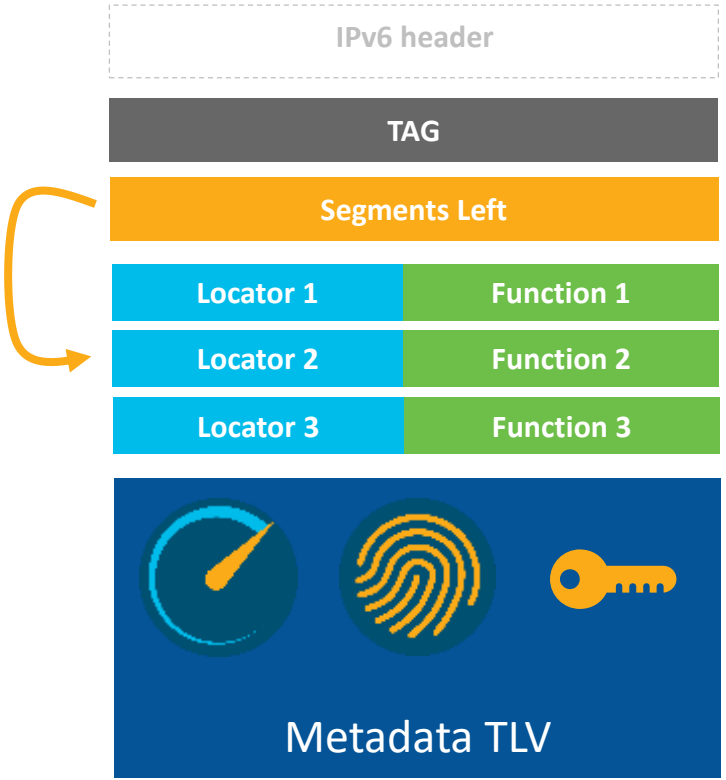
- A network program is a list of instructions (128-bit SRv6 SID)
- An instruction can be bound to any behavior
  - TE/FRR: END, END.X
  - VPN: END.DX, END.DT

SRv6 Header

RFC – Proposed Standard

18 HW linerate implementations

8 deployments



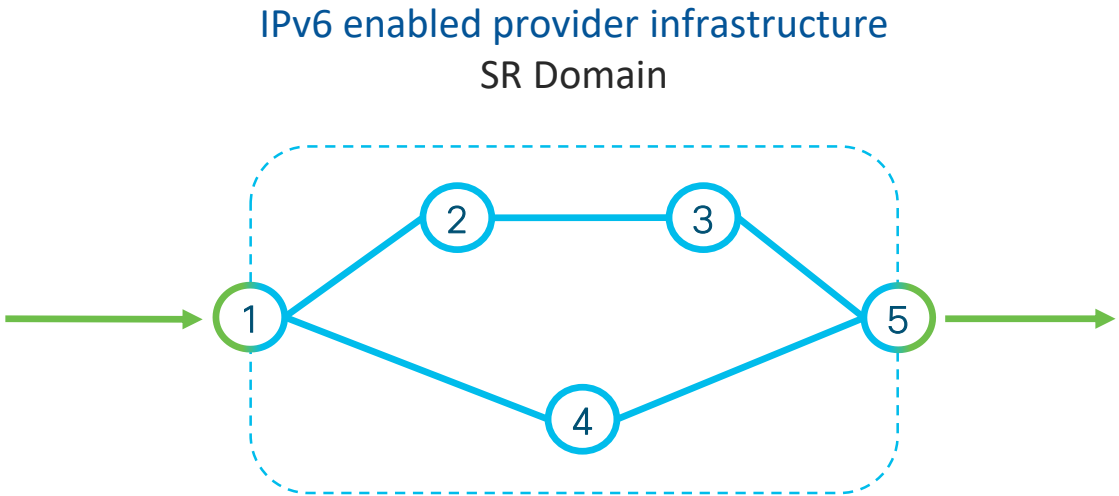
Optimized for HW processing

e.g. FRR, TE & VPN use-cases

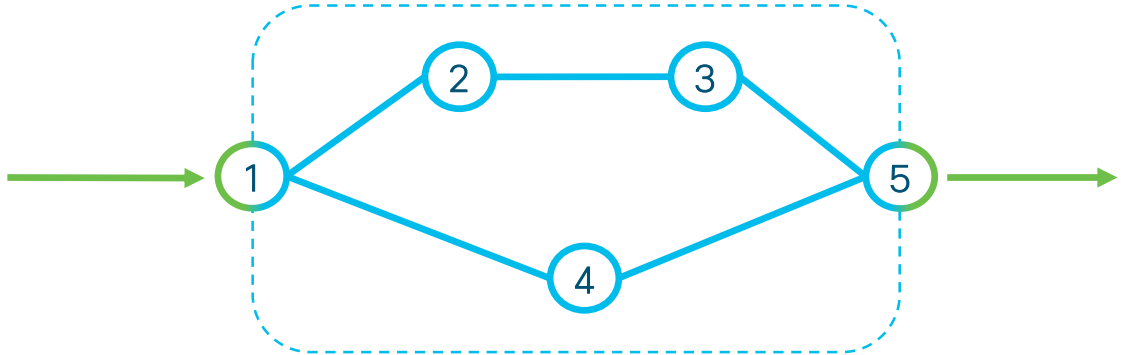
Optimized for SW processing

e.g. NFV, Container, Micro-Service

SRv6 Domain

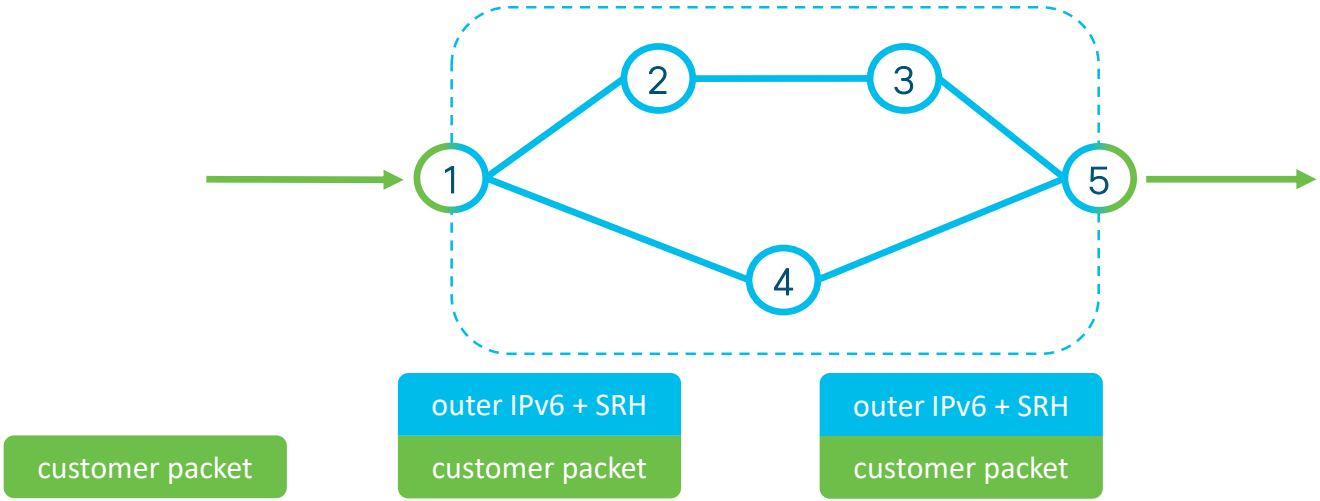


Encapsulation at the Domain ingress



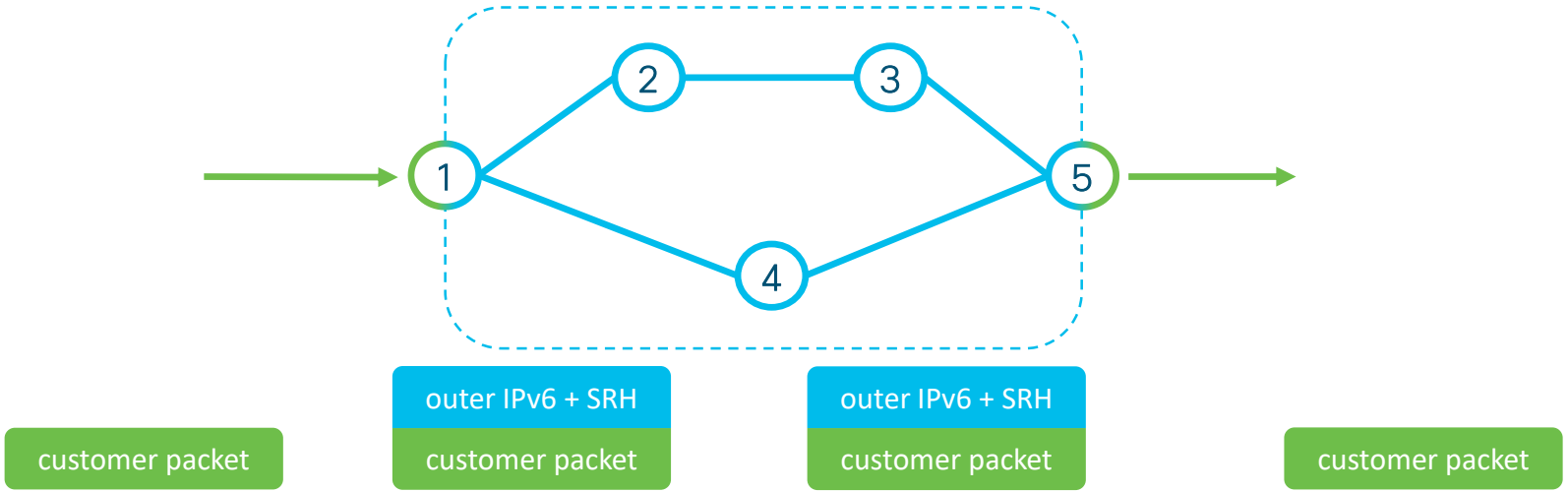
- IPv4, IPv6 or L2 frame is encapsulated within the SR Domain
- Outer IPv6 header includes an SRH with the list of segments

SRH of the outer IPv6 encapsulation



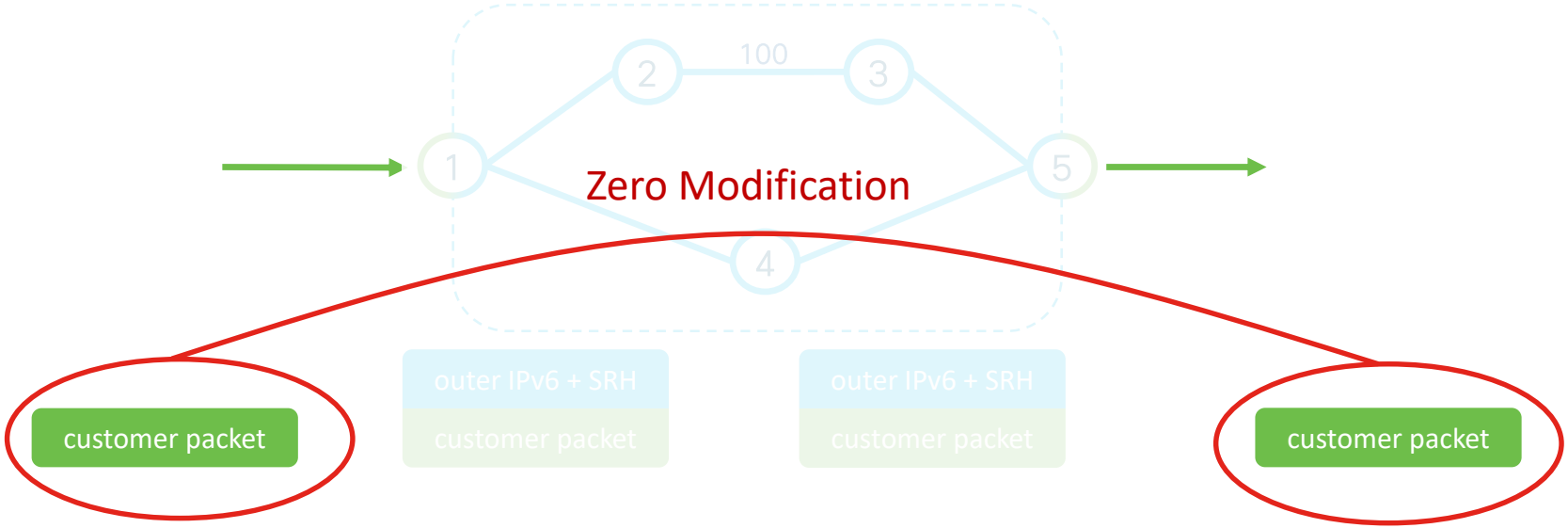
- Domain acts as a giant computer
- The network program in the outer SRH is executed

Decapsulation at Domain Egress



- Egress PE removes the outer IPv6 header as the packet leaves the SR domain

End-to-End Integrity



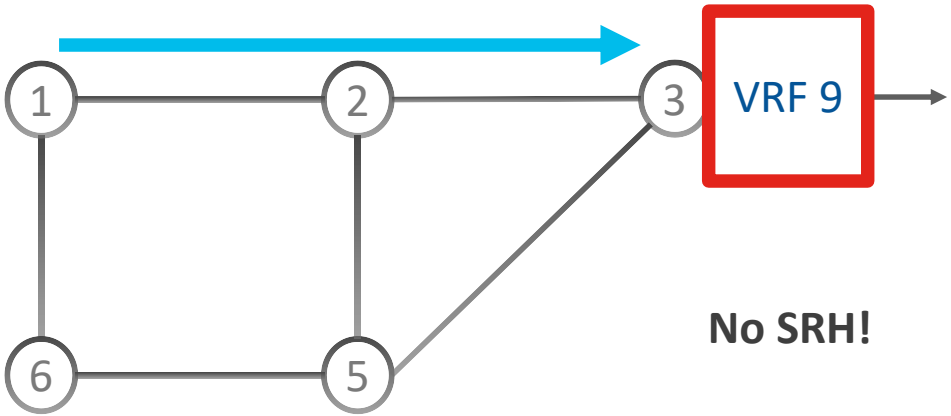
- End-to-end integrity principle is strictly guaranteed
  - Inner packet is unmodified
  - Same as SR-MPLS (MPLS stack is replaced by IPv6 outer header and SRH)

SRv6 Deployed Use-Cases

VPN over Best-Effort 5G Slice

Network Program: B:3:V(9)

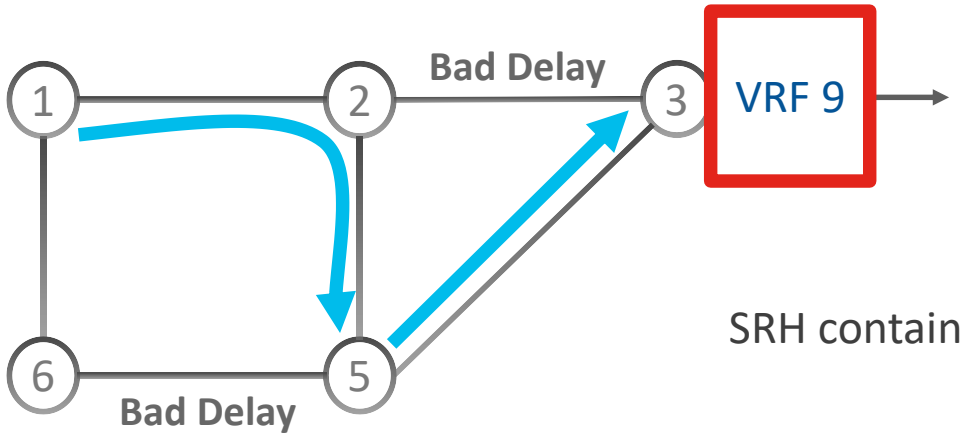
B*: locator block is associated with *ISIS base algo (Low Cost, Best Effort)



VPN with Low-Delay 5G Slice – SR-TE option

Network Program: B:2:C5 then B:3:V(9)

B: locator block is associated with ISIS base algo (Low Cost)

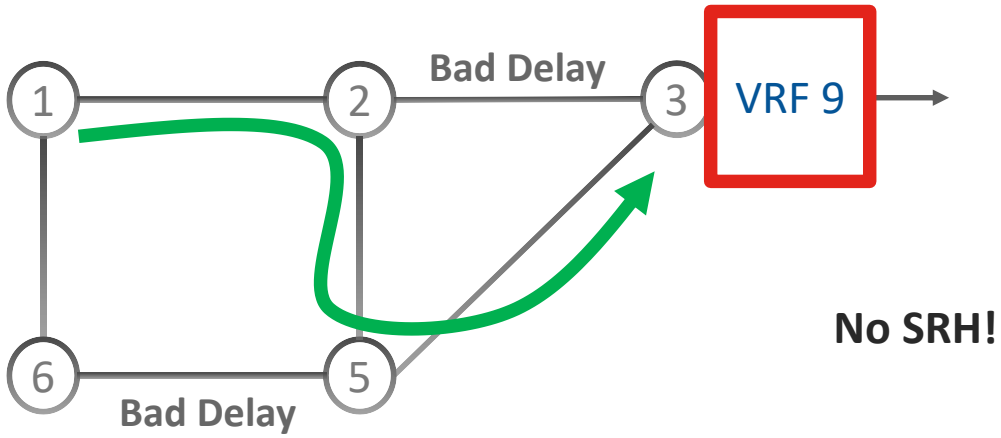


SRH contains 1 single SID

VPN with Low-Delay 5G Slice – Flex-algo option

Network Program: D:3:V(9)

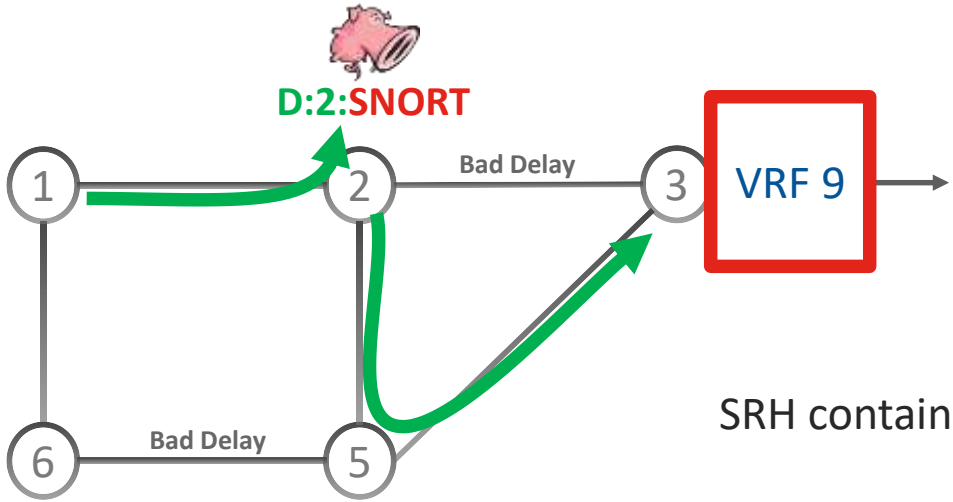
***D*: locator block is associated with *Low Delay* Flex-Algo**



Snort firewall, VPN & Low-Delay Slice

Network Program: D:2:SNORT then D:3:V(9)

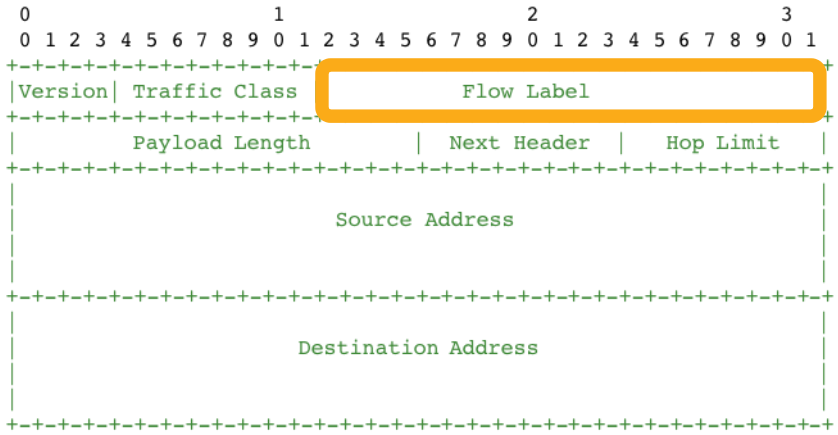
***D**: locator block is associated with **Low Delay** Flex-Algo*



SRH contains 1 Single SID

Load-balancing

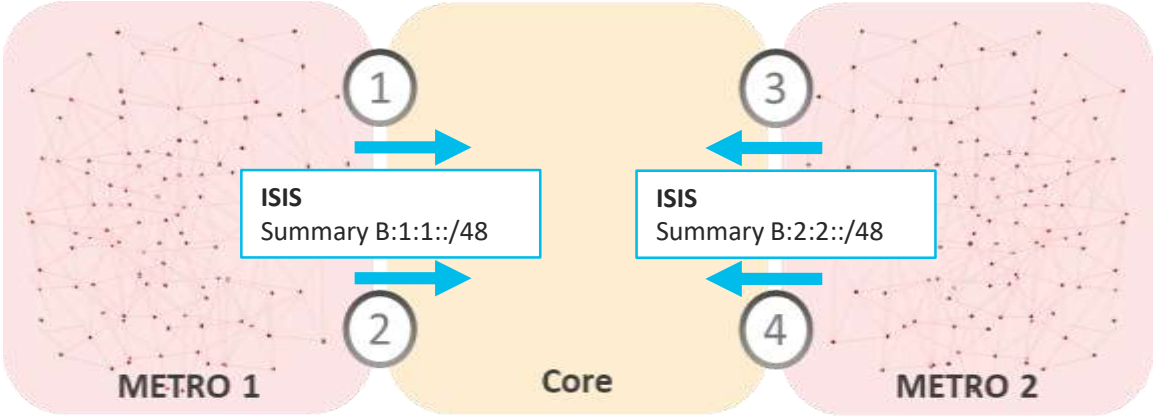
- 20-bit entropy
- No additional protocol
  - infamous mpls entropy label



Seamless Incremental Deployment

- As soon as the network supports plain IPv6 forwarding
  - A new SRv6-VPN service only requires PE upgrade
  - TE objective can be achieved with a few well selected TE waypoints
  - FRR is deployed incrementally

Prefix Summarization



- Back to basic IP routing and summarization
- No BGP inter-AS Option A/B/C

SRv6 has excellent native Scale

- Many use-cases do not even use an SRH ☺
  - Any VPN (L3VPN, PW, eVPN)
  - Egress Peering Engineering
  - Low-Latency or Disjoint Slicing
  - Optimal Load-Balancing
- If SRH is needed, most cases will use 1 or 2 SID’s
- Prefix Summarization gain
- Talk to the operators who deployed, they are happy to share experience

SRv6 NSO Automation

- Address allocation
  - Loopback and interfaces
- SID allocation
  - Algo 0 and Flex-Algos
- Multi-Domain
- ISIS summarization and redistribution between domains

- Latency Measurement
- Flex-Algo – Delay Slice
- TI-LFA
- BFD
- BGP-VPN

Negligible SRv6 SID block allocation - Iliad

As of the end of 2019, the SRv6 network consists of:

- 1000 Cisco NCS 5500 routers.
- 1800 Iliad's Nodeboxes.
- The network services 4.5 million mobile subscribers (as of Q3 2019).
- The network is carrying 300 Gbps of commercial traffic at peak hours.
- It is expected to grow to more than 4000 Nodeboxes in 2020. The SRv6 SIDs are allocated from a /40 sub-block of FC/8.

Less than 1 billionth of the FC/8 space - Negligible

Negligible SRv6 SID block allocation - SBB

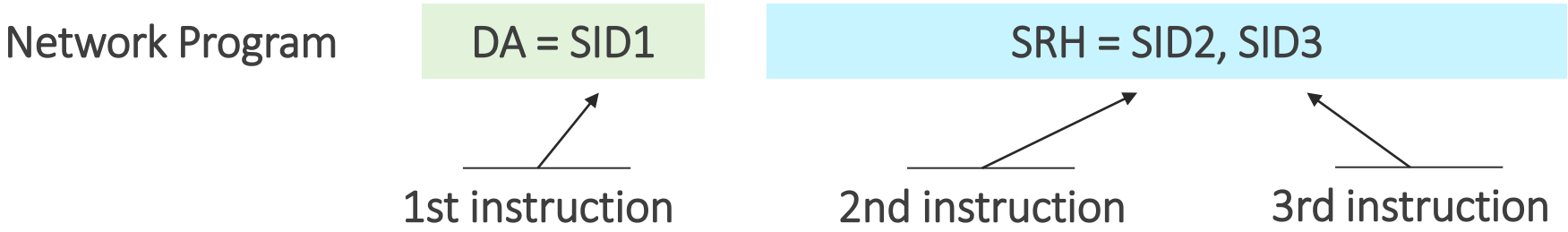
- SBB currently has a /20 public IPv6 space from APNIC
- SBB SRv6 is supported by a /40 sub-block
- This is only 1 millionth of the current SBB allocation

Micro-Program

A new set of (micro)-instructions

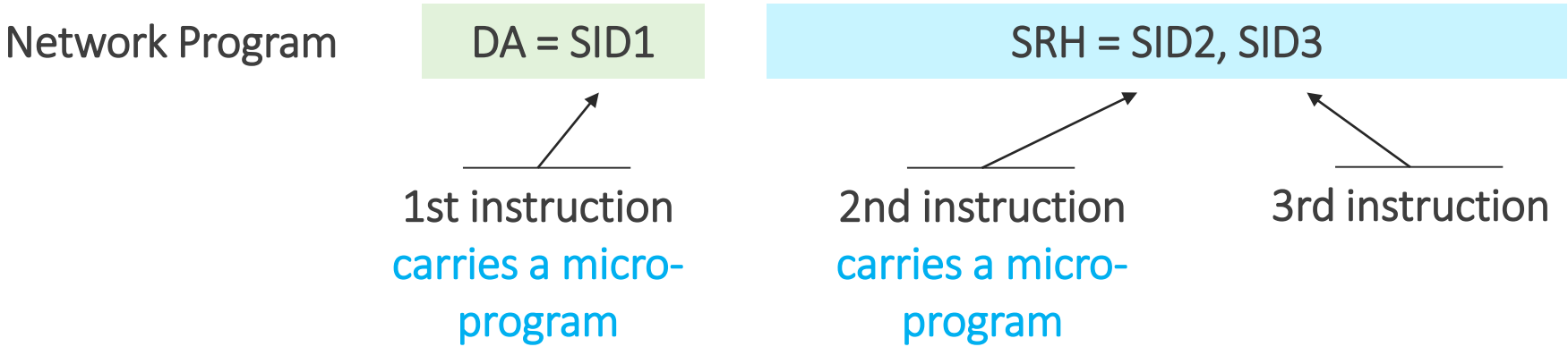
- Full leverage of SRH encapsulation
  - Zero extension
- Full leverage of SRv6 control-plane
  - Zero extension

Network Program

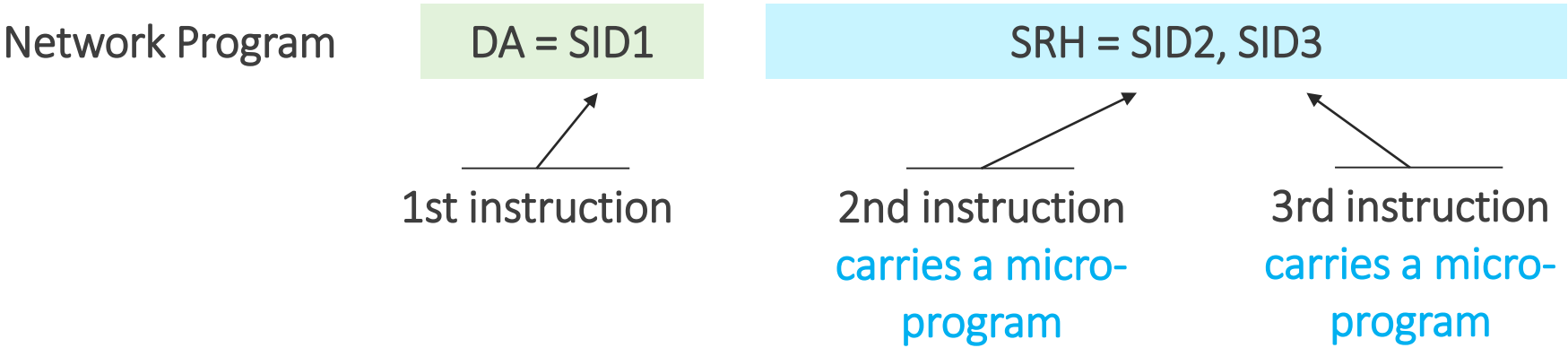


- A network program is a list of instructions (128-bit SRv6 SID)
- An instruction can be bound to any behavior
  - TE/FRR: END, END.X
  - VPN: END.DX, END.DT

Any instruction could hold a micro-program



Any instruction could hold a micro-program



Micro-Program in an SRv6 SID

SRv6 SID = 128 bits = 8 groups of 4 nibbles

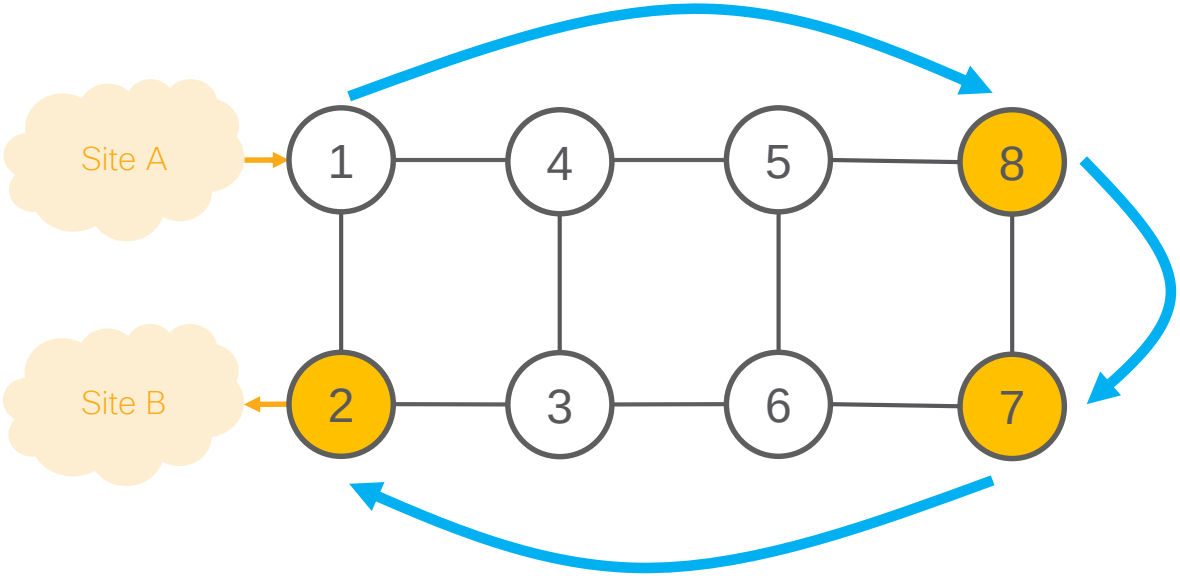
Assuming an allocation block in /32 (B:B::/32)

Assuming a micro-instruction ID in 4 nibbles

B:B:uID1:uID2:uID3:uID4:uID5:uID6

6 micro instructions per SRv6 Instruction

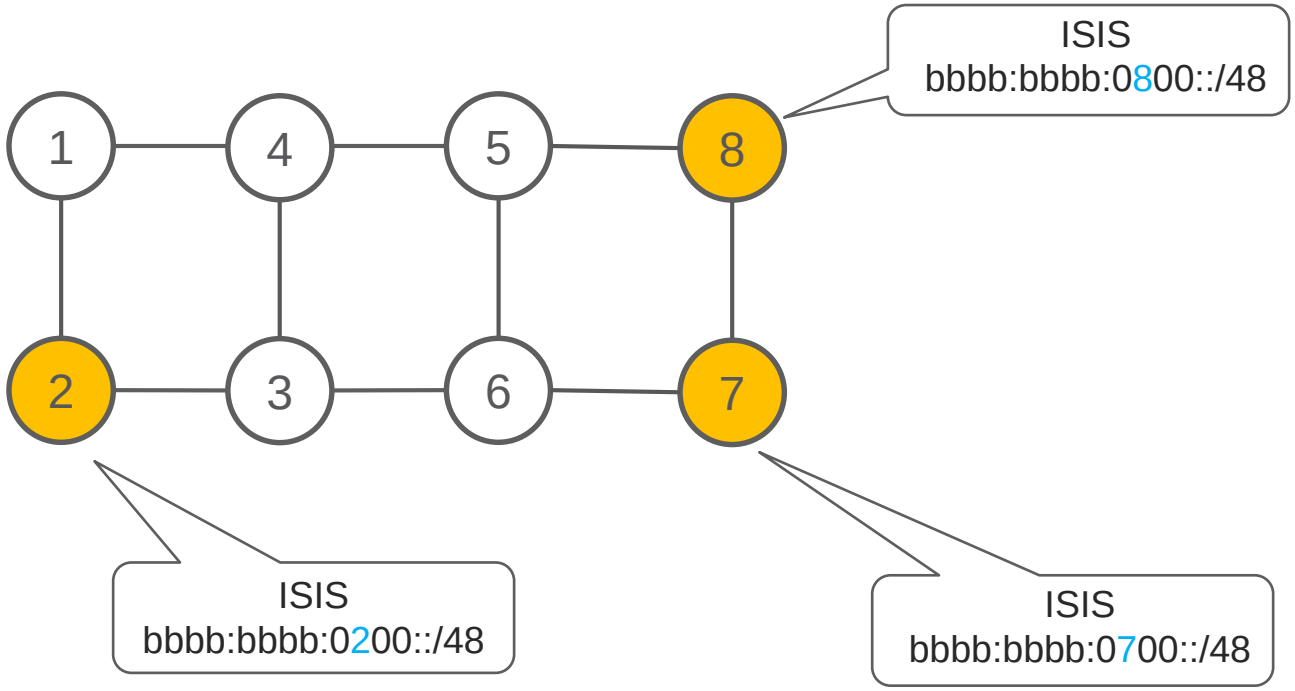
Illustration: go to 8 then 7 then 2 and decaps



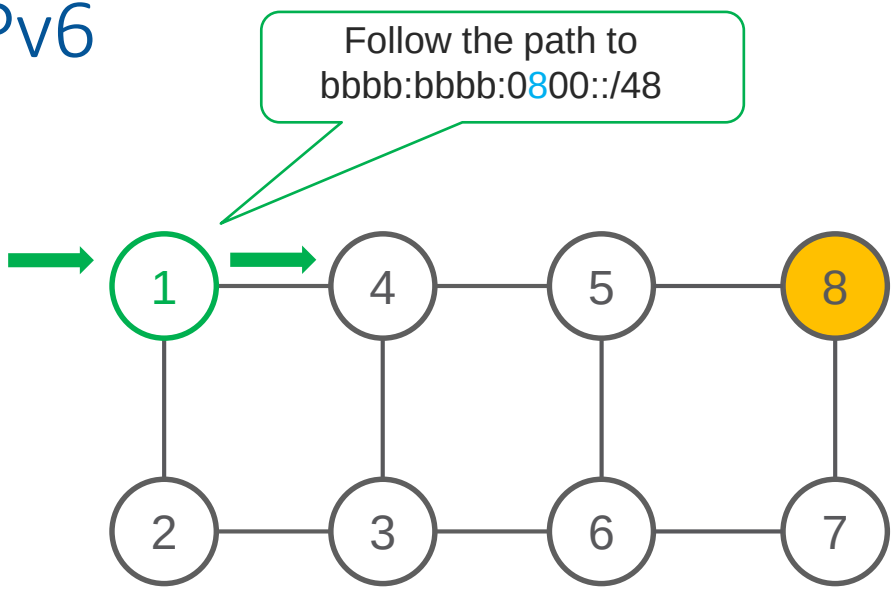
One single micro-program in the DA is enough

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

Basic IP Routing: no new extension

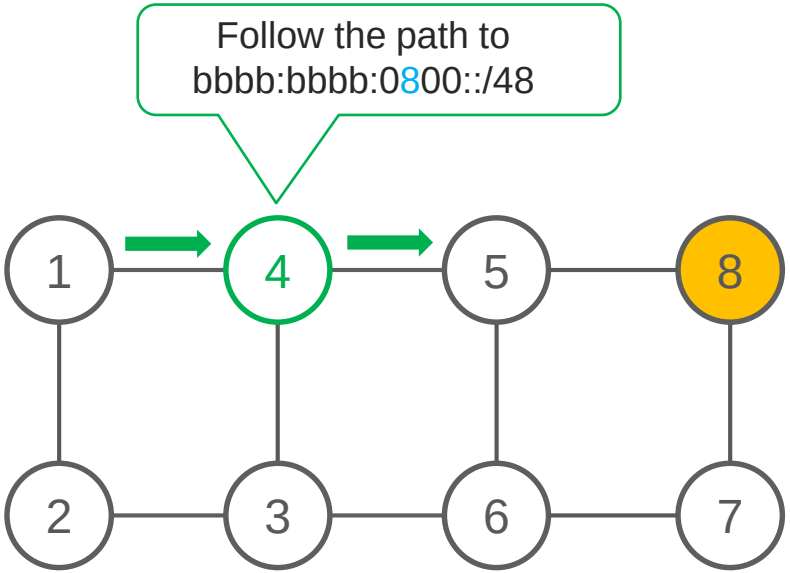


@1: basic IPv6



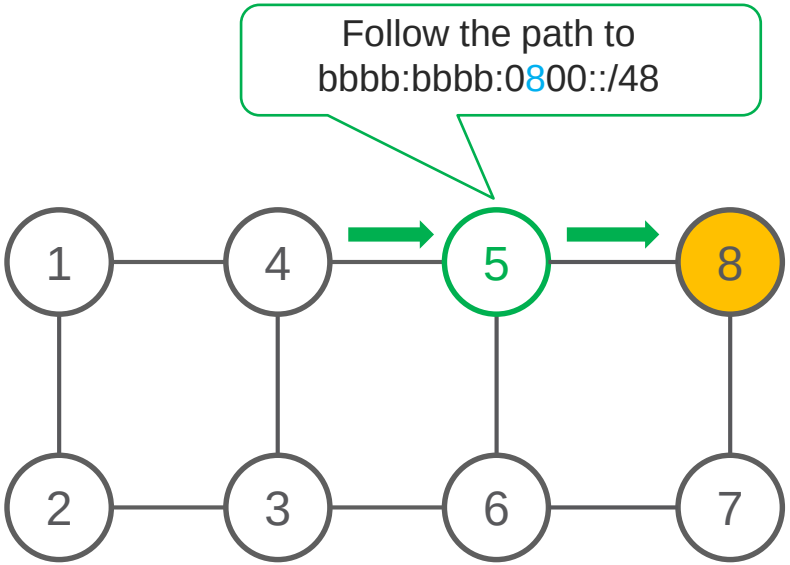
DA = bbbb:bbbb:0800:0700:0200:0000:0000:0000

@4: basic IPv6



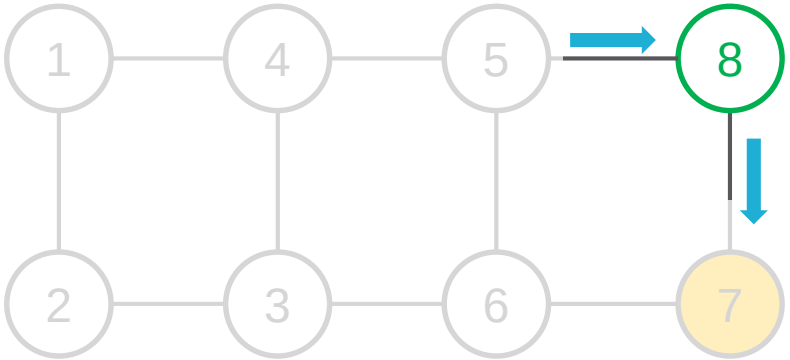
DA = bbbb:bbbb:0800:0700:0200:0000:0000:0000

@5: basic IPv6



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

@8: Shift and Forward



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

SHIFT << 16

Tx'd DA: bbbb:bbbb:0700:

bbbb:bbbb:0700::/48

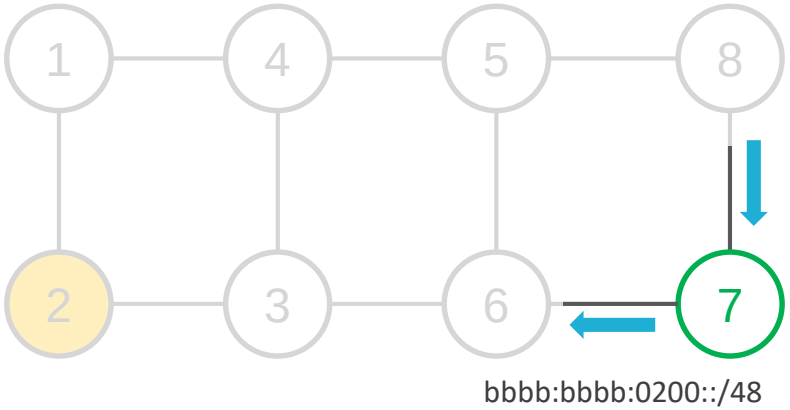
FIB Longest-Match bbbb:bbbb:0800::/48 → SRv6 Instruction:

Shift micro-Program by one micro-Instruction

Set last micro-instruction to "end of micro-program"

Lookup the updated DA and forward

@7: Shift and Forward



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

SHIFT << 16

Tx'd DA: bbbb:bbbb:0200:

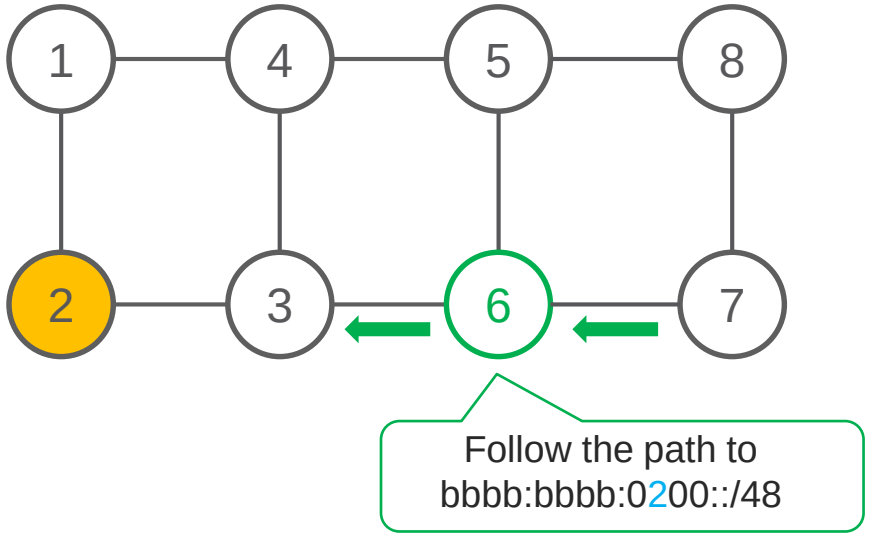
FIB Longest-Match bbbb:bbbb:0700::/48 → SRv6 Instruction:

Shift micro-Program by one micro-Instruction

Set last micro-instruction to "end of micro-program"

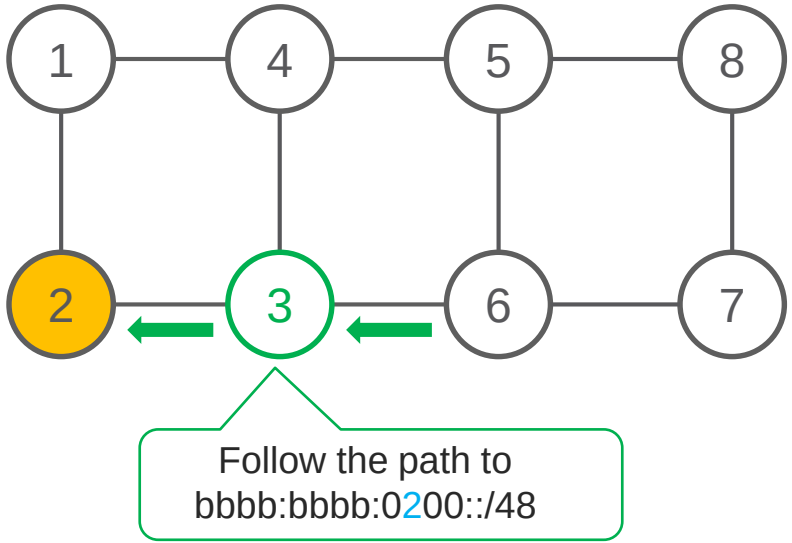
Lookup the updated DA and forward

@6: basic IPv6



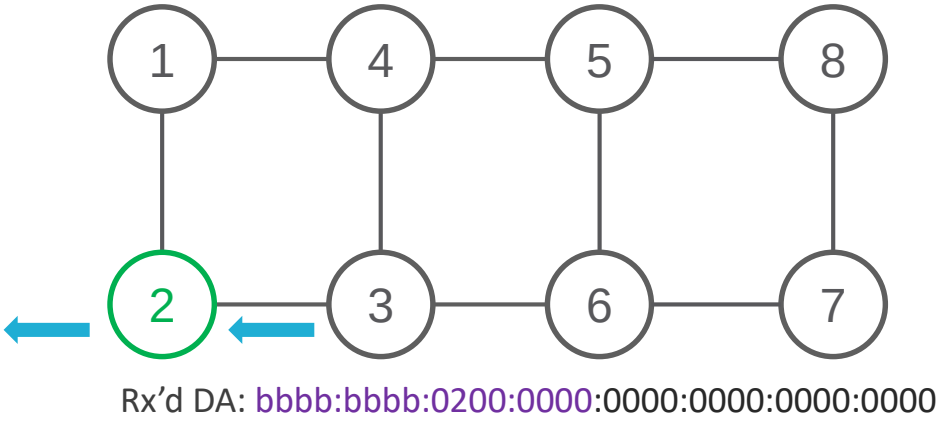
DA = bbbb:bbbb:0200:0000:0000:0000:0000:0000

@3: basic IPv6



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

@2: SRv6 End.DX4 behavior



FIB Longest-Match bbbb:bbbb:0200:0000::/64 → SRv6 Instruction:

Decapsulate and cross-connect inner IPv4 packet to Site B

Benefits

- Ultra-scalable for 5G deployment
  - 18 FRR, TE, NFV and VPN micro-instructions in only 40 byte SRH overhead
- Mathematically the best SRv6 compression solution
- Linerate for multi-Tbps hardware
  - Shift is a basic hardware logic
- Friendly to merchant silicon
  - Proven by endorsement and interop
- Friendly to legacy equipment

No Cost

- Full leverage (zero change) to SRv6
  - Net PGM model
  - SRH encapsulation
  - Control Plane
- Seamless Deployment on IPv6 host

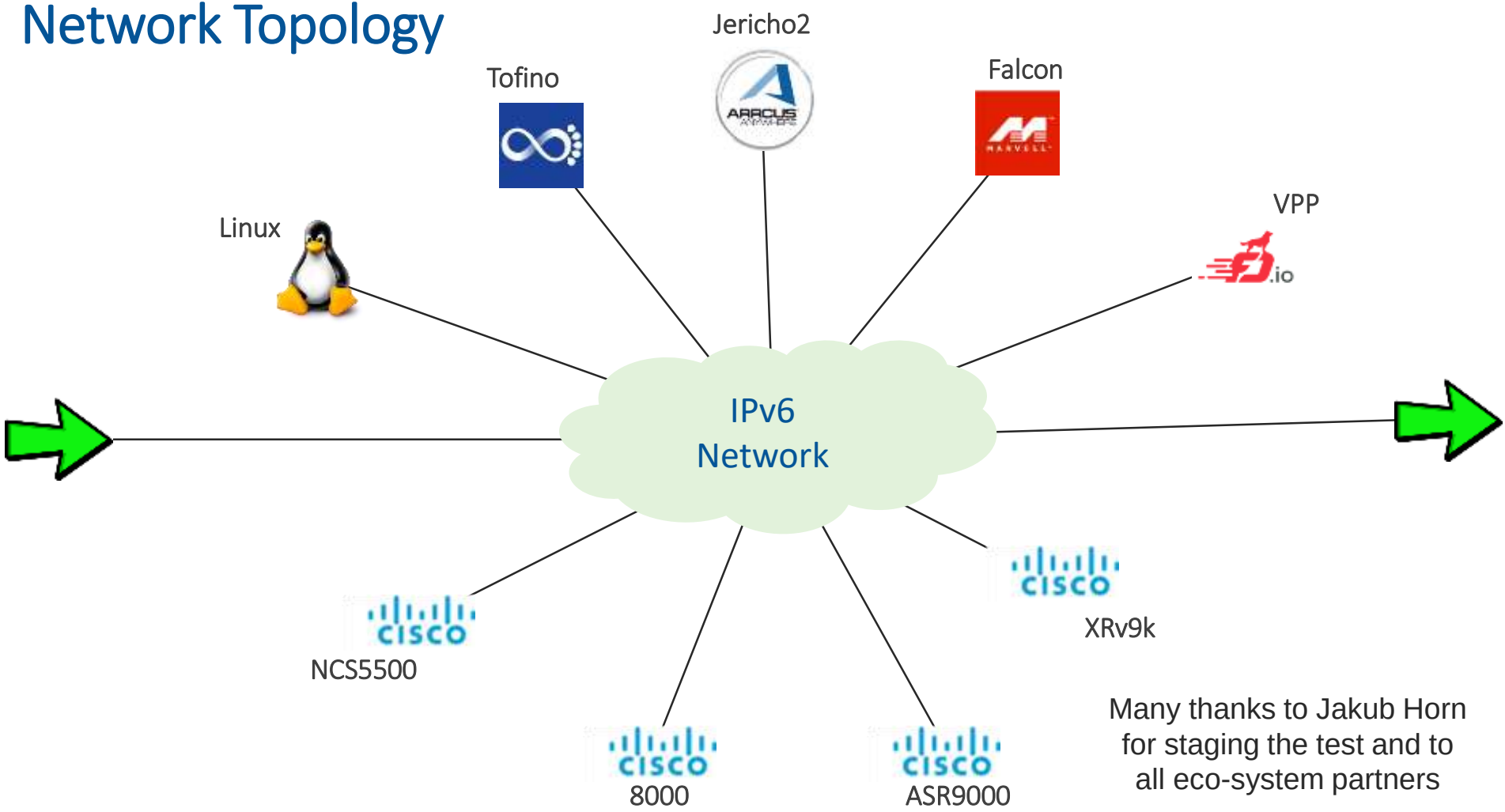
Alibaba

- Dennis Cai reports successful milestone in January 2020
  - SRv6 micro-program linerate hardware verification in Cisco lab
  - Cisco 8000 series (silicon one), NCS-5500, ASR9k
- Use-case
  - Applications are already IPv6 enabled
  - Network is already IPv6-enabled
  - Seamless end-to-end SDN control from Apps through DC, Metro, Backbone

Bell Canada

- Dan Voyer reports successful milestone in January 2020
  - SRv6 micro-program linerate hardware verification
  - Cisco 8000 series (silicon one), NCS-5500, ASR9k, CRS-X
- Use-case: 5G with
  - Ultra Scale
  - Protocol simplification and IPv6 convergence
  - Integrated TE, FRR, Slicing, VPN and NFV for end-to-end value-added service
  - Optimum Load-Balancing
  - Legacy reuse, CRS-X

Network Topology



Many thanks to Jakub Horn for staging the test and to all eco-system partners

Conclusion

Simplicity Always Prevails

~~LDP~~

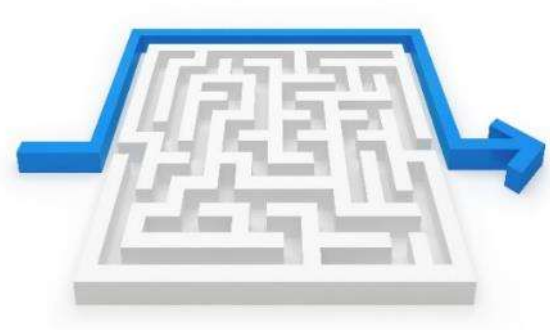
~~RSVP-TE~~

~~Inter-AS Option A/B/C~~

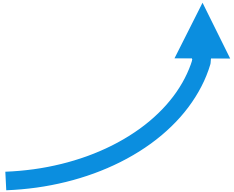
~~MPLS~~

~~UDP/VxLAN~~

~~NSH~~



Furthermore with more scale and functionality



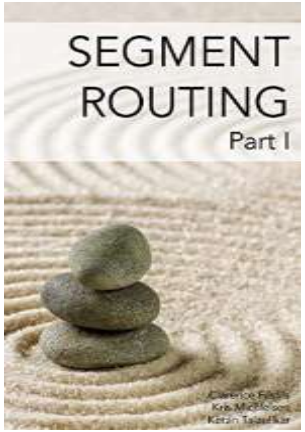
At record speed

- In 2019: 8 large-scale commercial deployments
  - Softbank, Iliad, China Telecom, LINE corporation, China Unicom, CERNET2, China Bank and Uganda MTN.
- 18 HW linerate implementations
  - Cisco Systems, Huawei
  - Broadcom, Barefoot, Intel, Marvell, Mellanox
  - Multiple Interop Reports
- 11 open-source platforms/ Applications
  - Linux, FD.io VPP, P4, Wireshark, tcpdump, iptables, nftables, snort, ExaBGP, Contiv-VPP

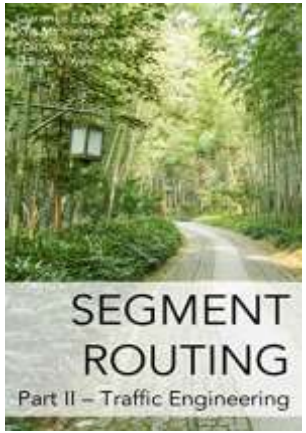
An extended recording
of this presentation is
available at

segment-routing.net/srv6-status/

Stay up-to-date



amzn.com/B01I58LSUO



amazon.com/dp/B07N13RDM9

twitter.com/SegmentRouting

segment-routing.net

facebook.com/SegmentRouting/

linkedin.com/groups/8266623

Thank you!