SRv6 Technology Update

Pablo Camarillo

pcamaril@cisco.com

February 2019

Segment Routing

- Source Routing
  - the topological and service (NFV) path is encoded in packet header
- Scalability
  - the network fabric does not hold any per-flow state for TE or NFV
- Simplicity
  - automation: TILFA sub-50msec FRR
  - protocol elimination: LDP, RSVP-TE, VxLAN, NSH, GTP, ...
- End-to-End
  - DC, Metro, WAN

Two dataplane instantiations

Segment Routing

MPLS

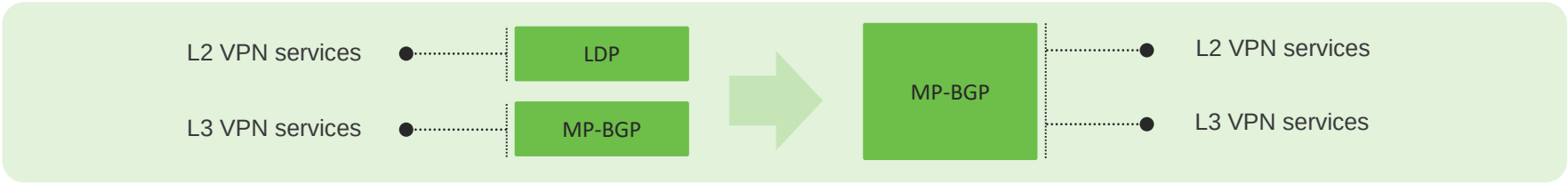
- leverage the mature MPLS HW with only SW upgrade
- 1 segment = 1 label
- a segment list = a label stack

IPv6

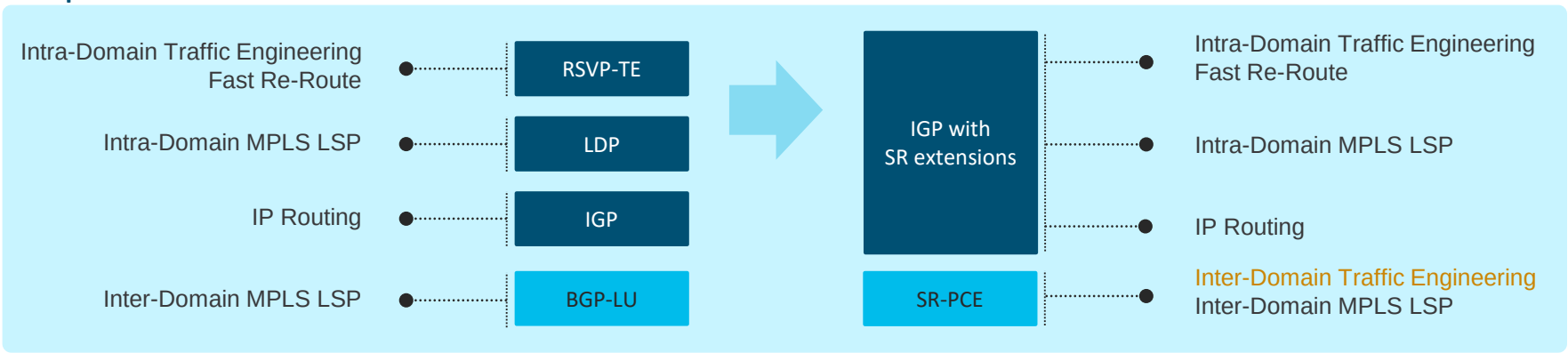
- leverages RFC8200 provision for source routing extension header
- 1 segment = 1 address
- a segment list = an address list in the SRH

Network Evolution

Service Protocols

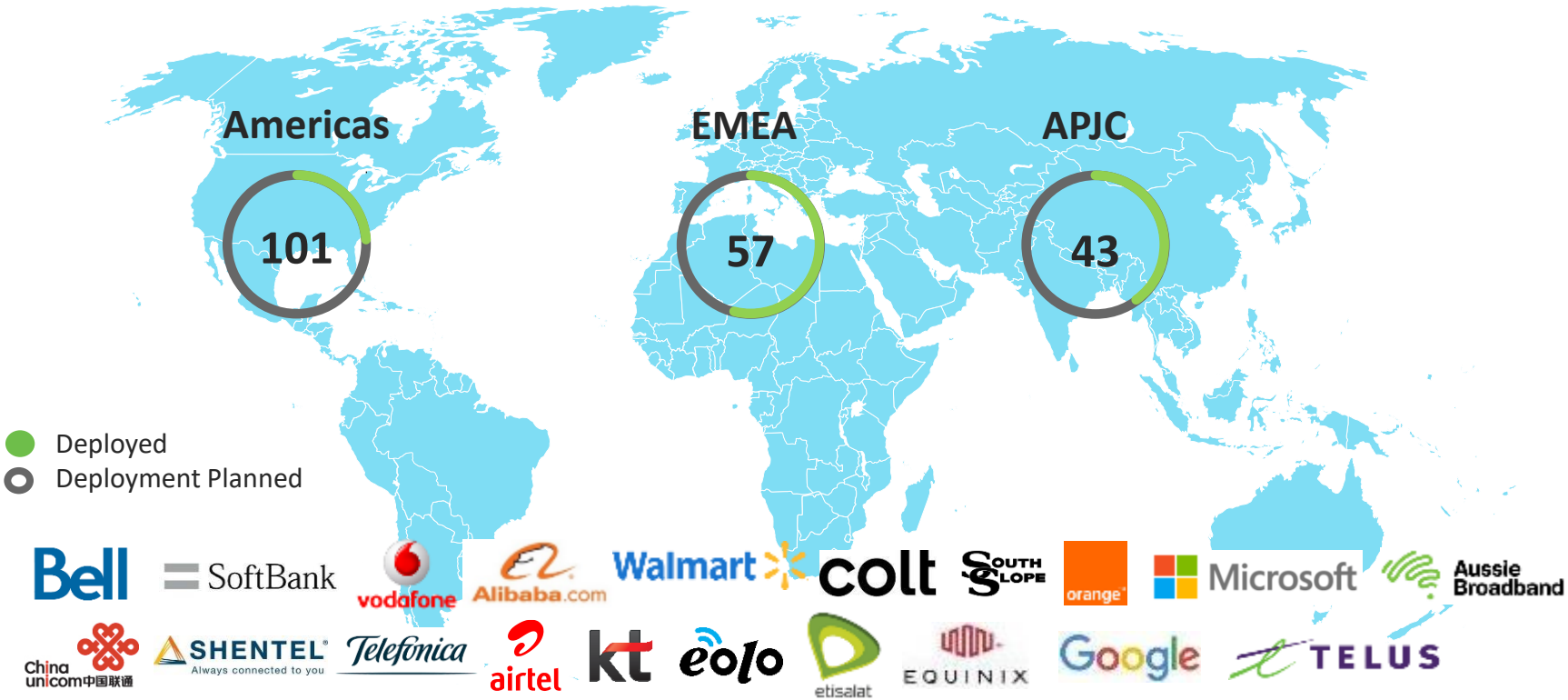


Transport Protocols



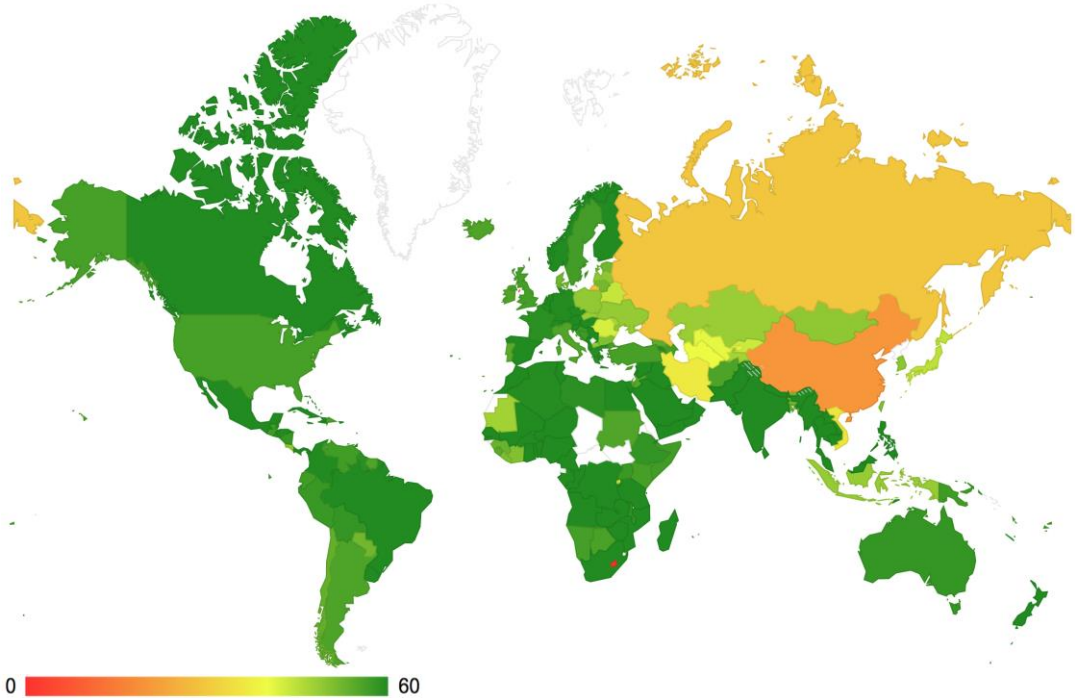
LDP: Label Distribution Protocol, **MP-BGP**: Multi-protocol BGP, **BGP-LU**: BGP Labeled-Unicast, **PCE**: Path Computation Element, **RSVP-TE**: Reservation Protocol Traffic Engineering

SR-MPLS: de-facto IPv4 solution



Let’s focus on SRv6: SR for IPv6

IPv6 adoption is a reality



% Web pages available over IPv6

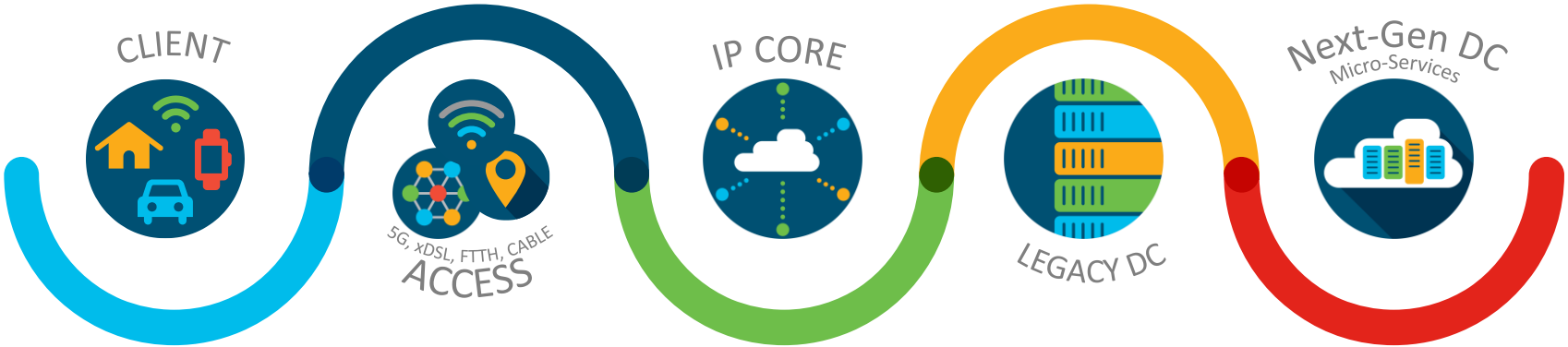
Sources: 6lab.cisco.com – Web content
Cisco VNI Global IP Traffic Forecast, 2017-2022

Global IPv6 traffic **grew 226%** in 2017

Globally IPv6 traffic **will grow 18-fold** from 2017 to 2022

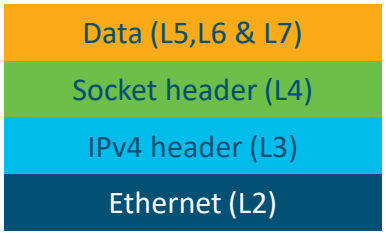
IPv6 **will be 38%** of total Internet traffic in 2022

IPv6 provides reachability



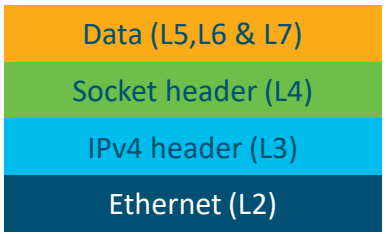
IPv4 limitations & work-arounds

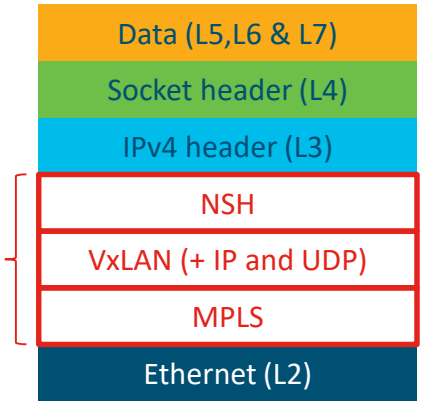
- × Limited address space
- × No engineered Load Balancing
- × No VPN
- × No Traffic Engineering
- × No Service Chaining



IPv4 limitations & work-arounds

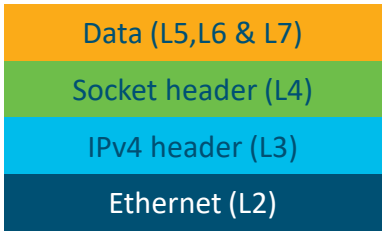
- × Limited address space → NAT
- × No engineered Load Balancing → MPLS Entropy Label, VxLAN UDP
- × No VPN → MPLS VPN's, VxLAN
- × No Traffic Engineering → RSVP-TE, SR-TE MPLS
- × No Service Chaining → NSH



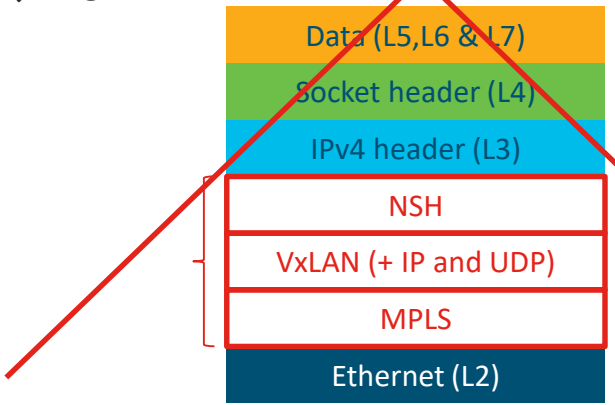


IPv4 limitations & work-arounds

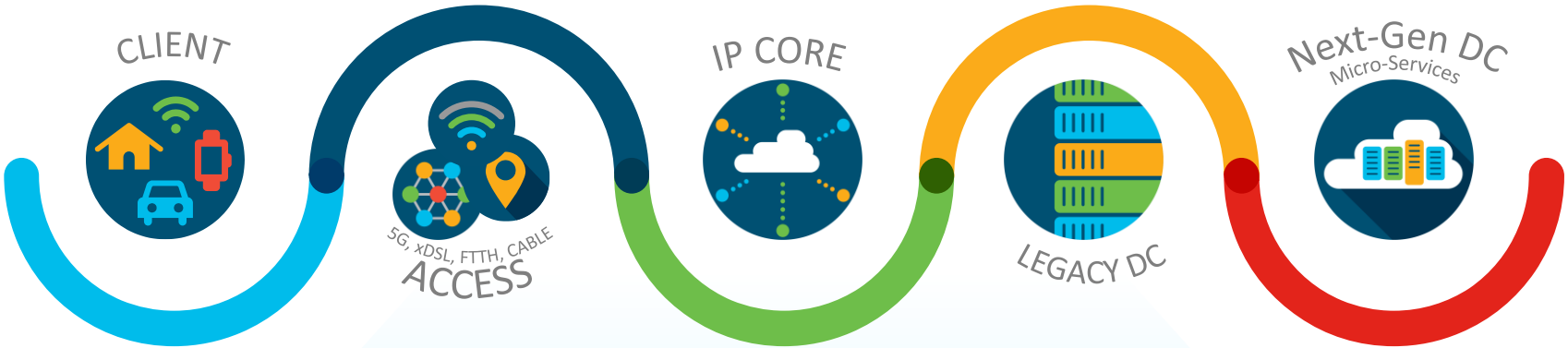
- × Limited address space → NAT
- × No engineered Load Balancing → MPLS Entropy Label, VxLAN UDP
- × No VPN → MPLS VPN's, VxLAN
- × No Traffic Engineering → RSVP-TE, SR-TE MPLS
- × No Service Chaining → NSH



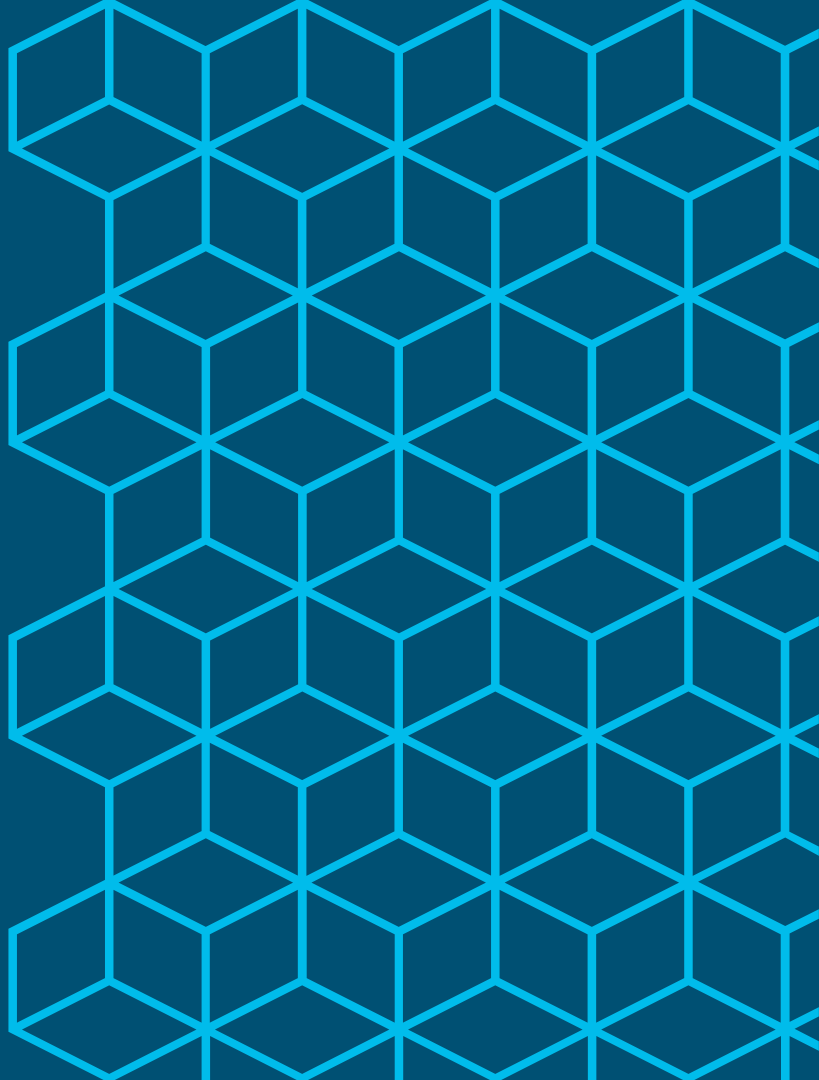
Simplicity (back to the OSI model)



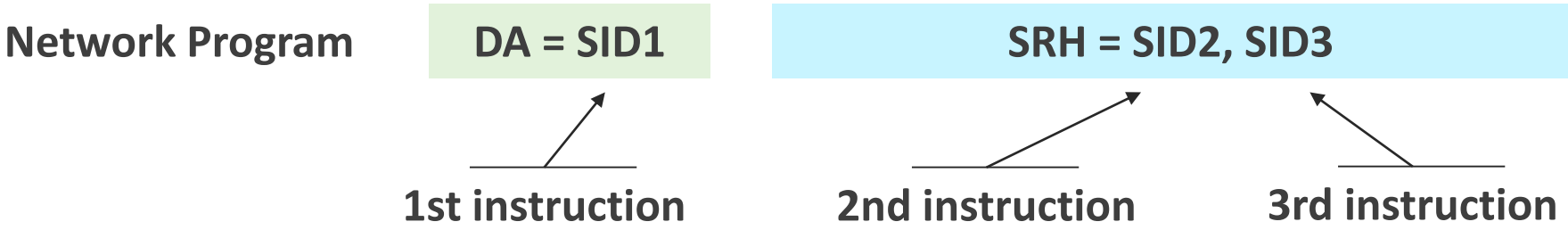
SRv6 unleashes IPv6 potential



SR for anything:
Network as a Computer

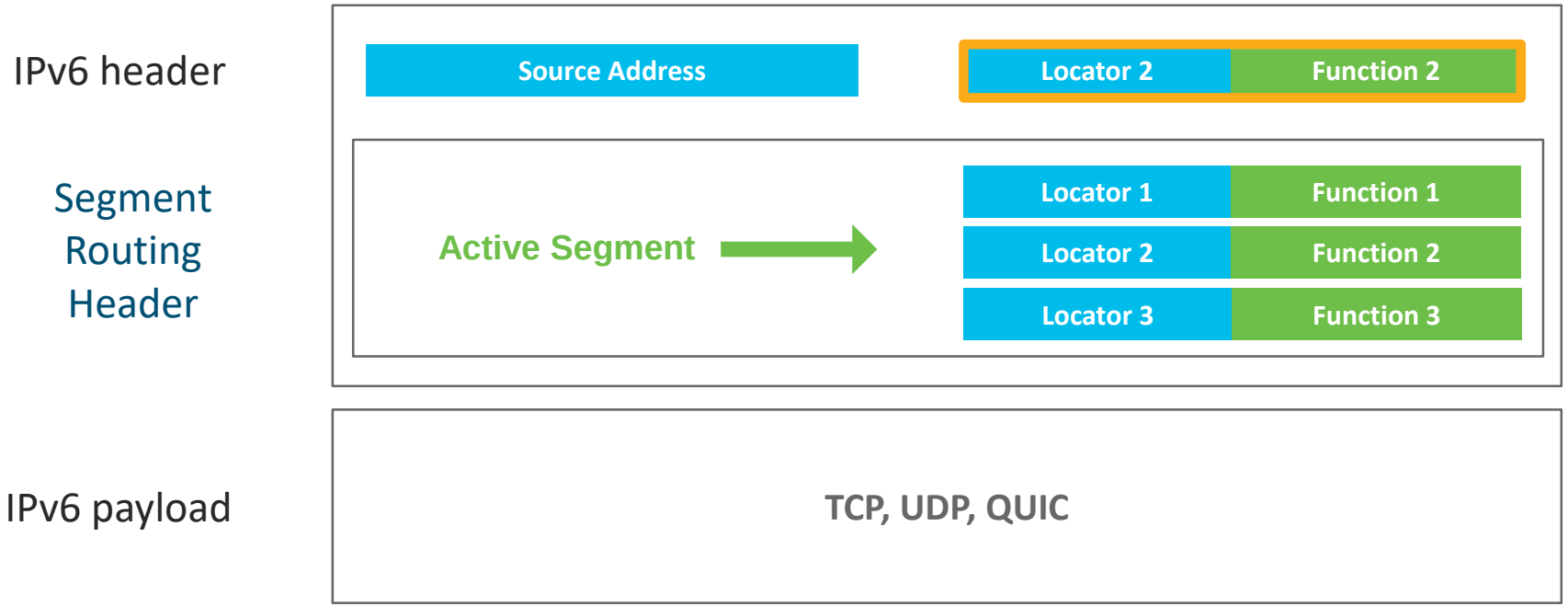


Network Programming

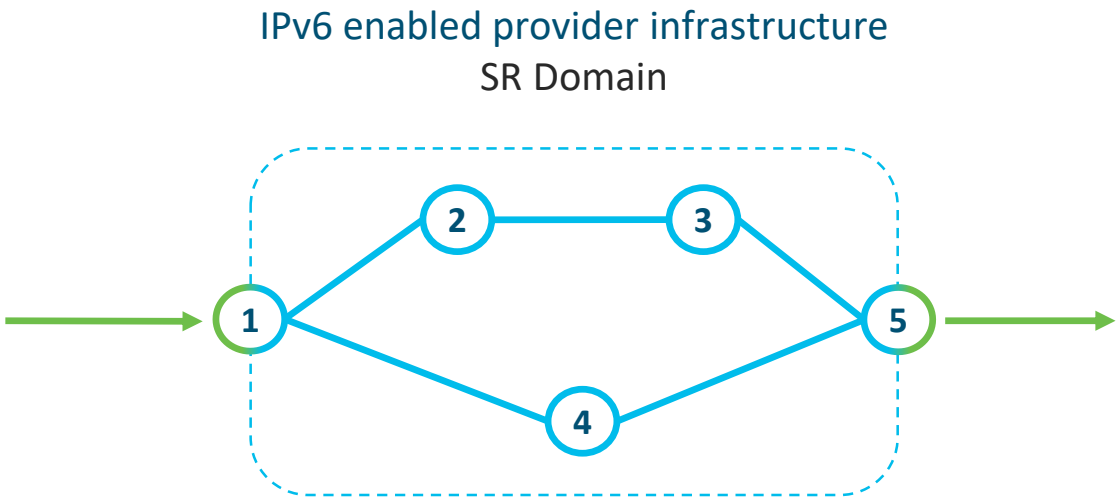


- A network program is a list of instructions (128-bit SRv6 SID)
- An instruction can be bound to any behavior
  - TE/FRR: END, END.X
  - VPN: END.DX, END.DT

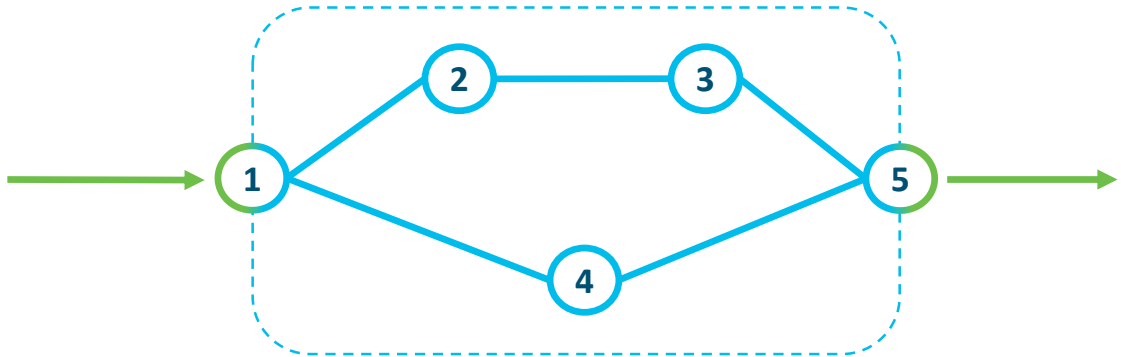
Network Program in the Packet Header



SRv6 Domain

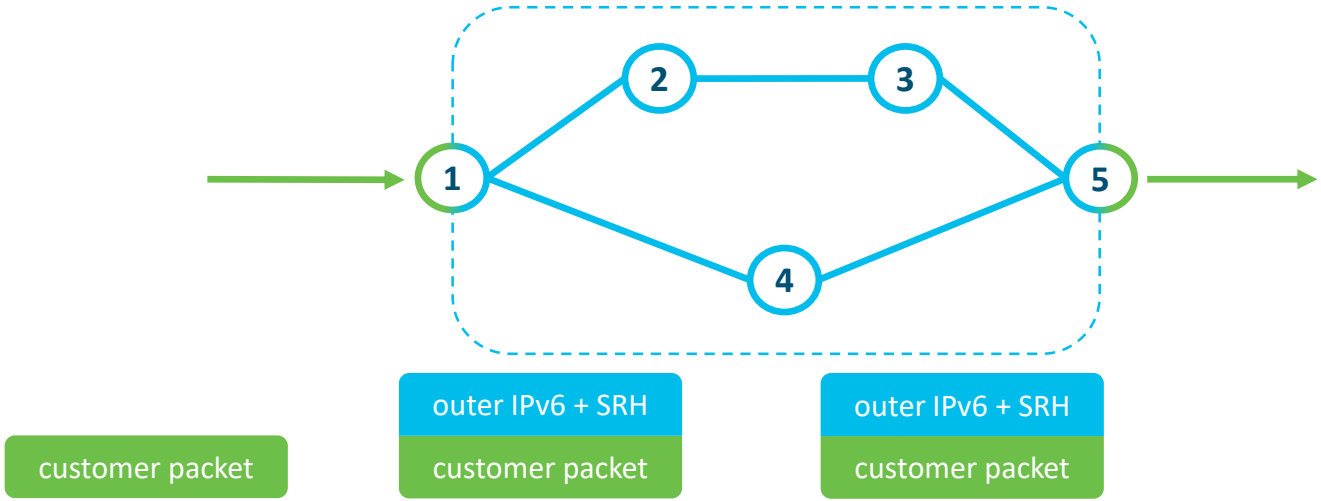


Encapsulation at the Domain ingress



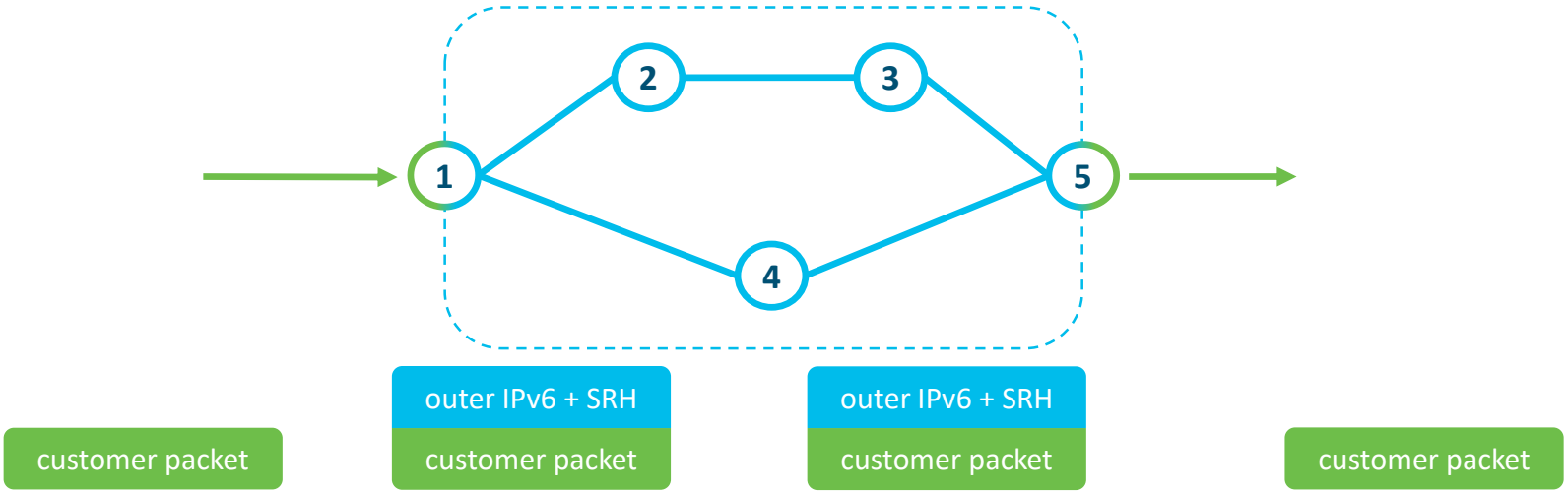
- IPv4, IPv6 or L2 frame is encapsulated within the SR Domain
- Outer IPv6 header includes an SRH with the list of segments

SRH of the outer IPv6 encapsulation



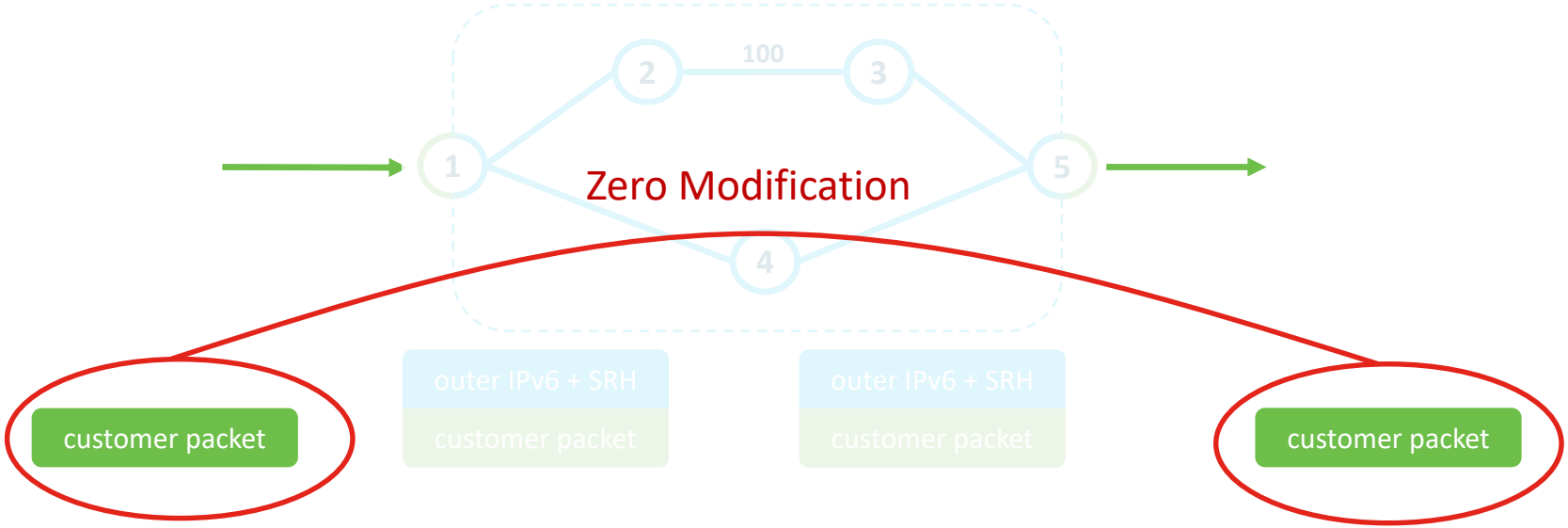
- Domain acts as a giant computer
- The network program in the outer SRH is executed

Decapsulation at Domain Egress



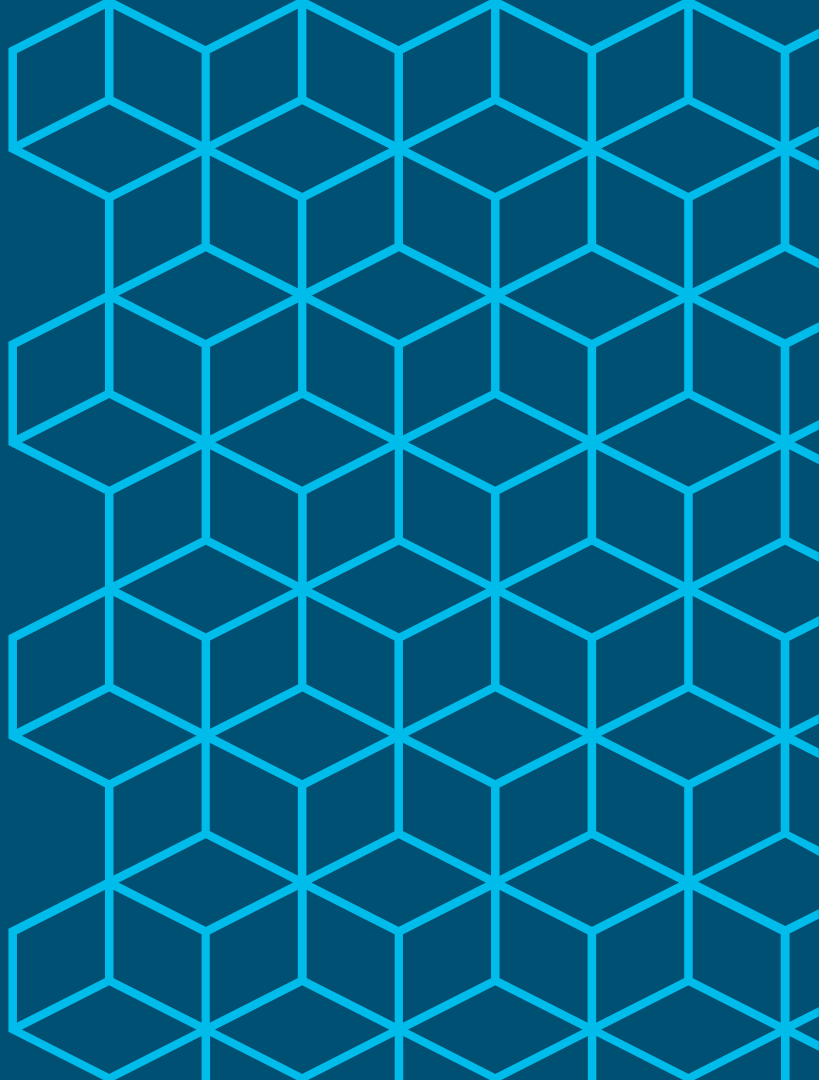
- Egress PE removes the outer IPv6 header as the packet leaves the SR domain

End-to-End Integrity



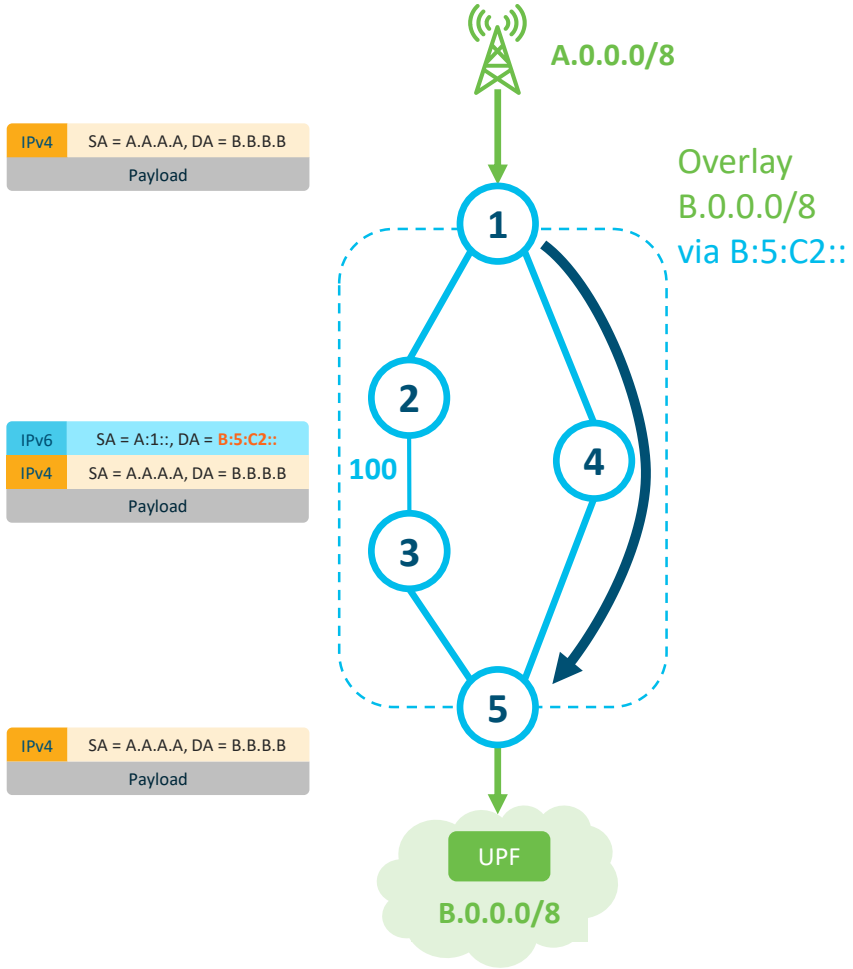
- End-to-end integrity principle is strictly guaranteed
  - Inner packet is unmodified
  - Same as SR-MPLS (MPLS stack is replaced by IPv6 outer header and SRH)

Deployment use-cases



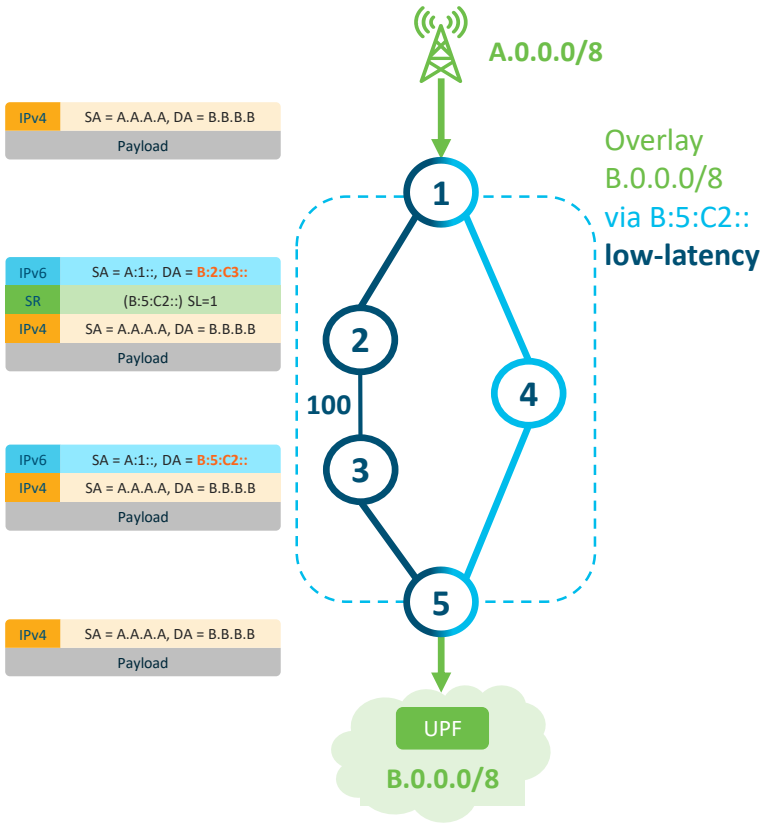
Overlay VPN

- Automated
  - No tunnel to configure
- Simple
  - Protocol elimination
- Efficient



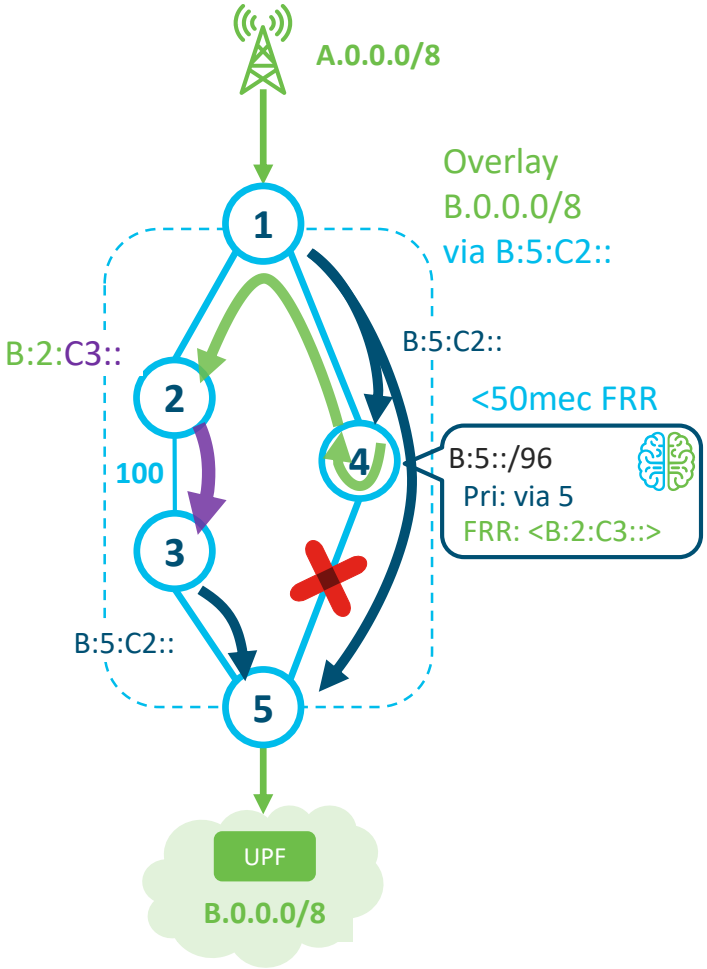
Traffic Engineering

- Automated SR TE policy triggered by SLA color of BGP route
  - No RSVP, no tunnel
- Automated Steering
  - No PBR steering complexity
- Inter-domain calculated by SR PCE
- SR native algorithms



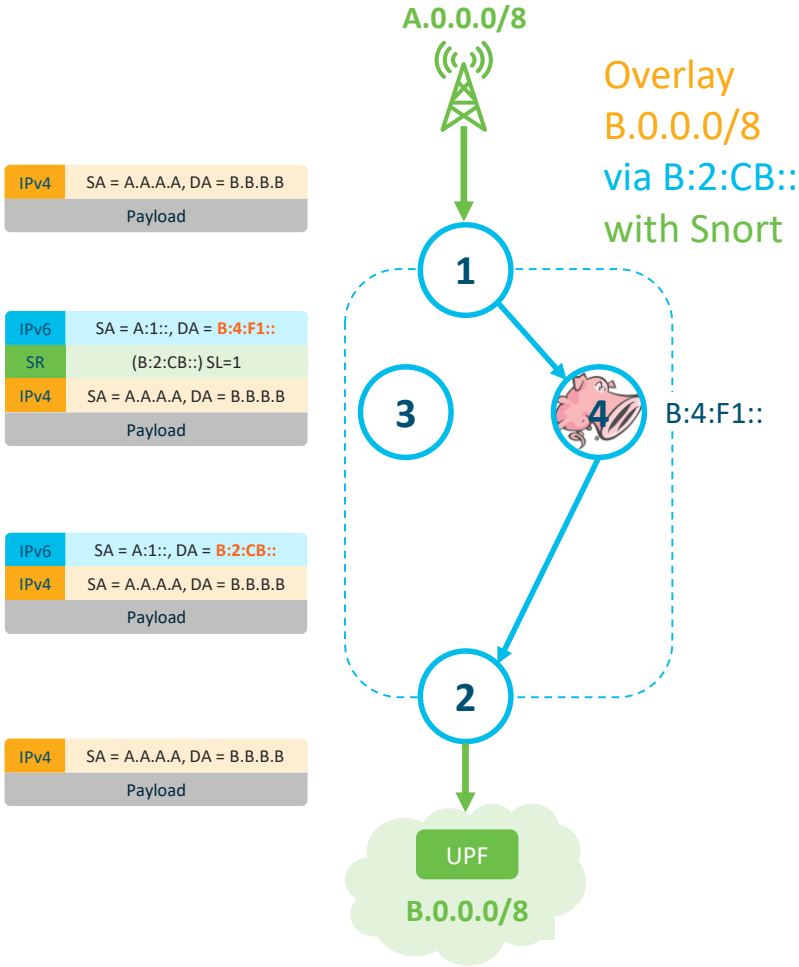
TILFA

- 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
  - avoid any intermediate flap via alternate path
- Incremental deployment
- Distributed and Automated Intelligence



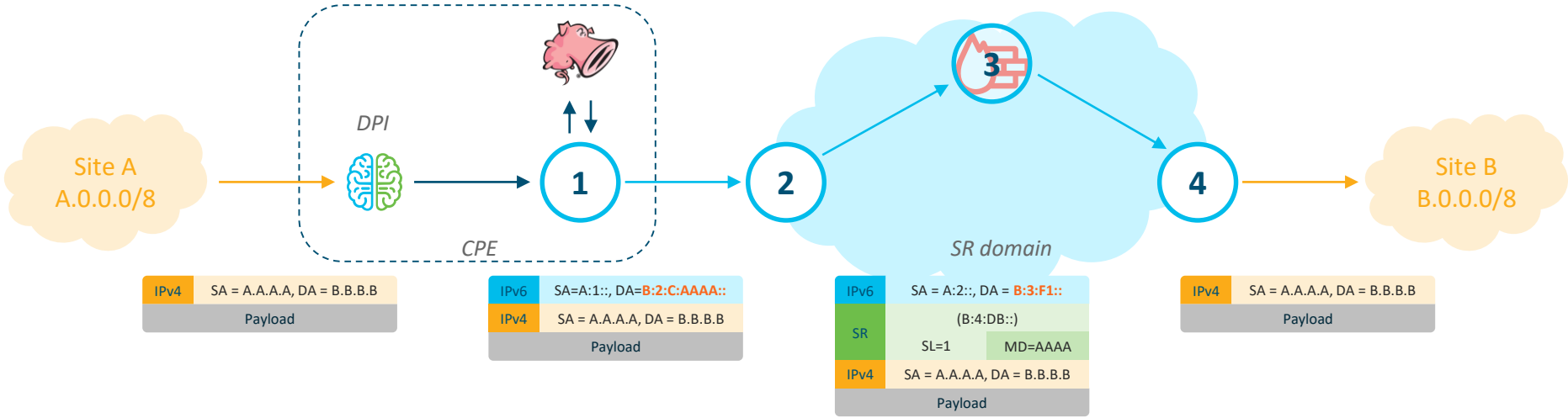
Integrated service chaining

- Instruction bound to a Network Function
  - Just another type of segment
  - Stateless in the fabric
  - Seamless integration with VPN and TE
- NF can leverage the SRH
  - Implement branching operation
  - Read / write metadata
- Open-source SR-aware NFs
  - Snort, iptables, nftables
  - Leverage native SRv6 support in Linux kernel

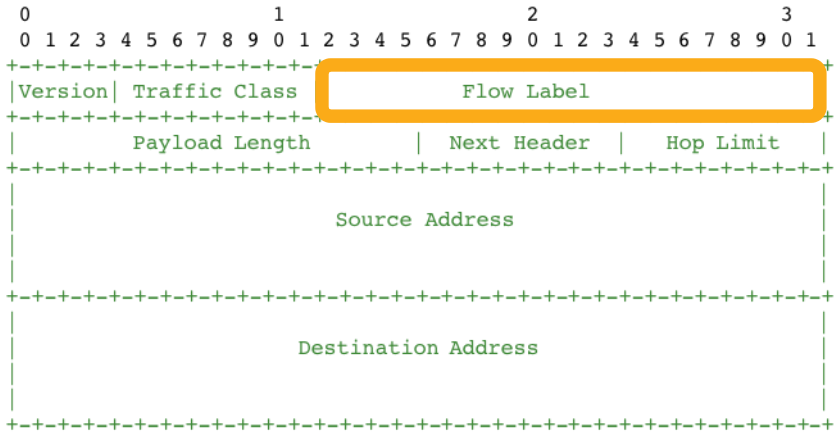


SRv6 service programming with metadata

- DPI inspects the traffic and indicates the classification result as a SID argument
- CPE may also perform some local service chaining (e.g. via Snort)
- Node 2 steers the traffic in an SR policy and copies the DA argument in an SRH TLV
- Firewall node 3 leverages this SRH metadata for smarter packet filtering



Load-balancing

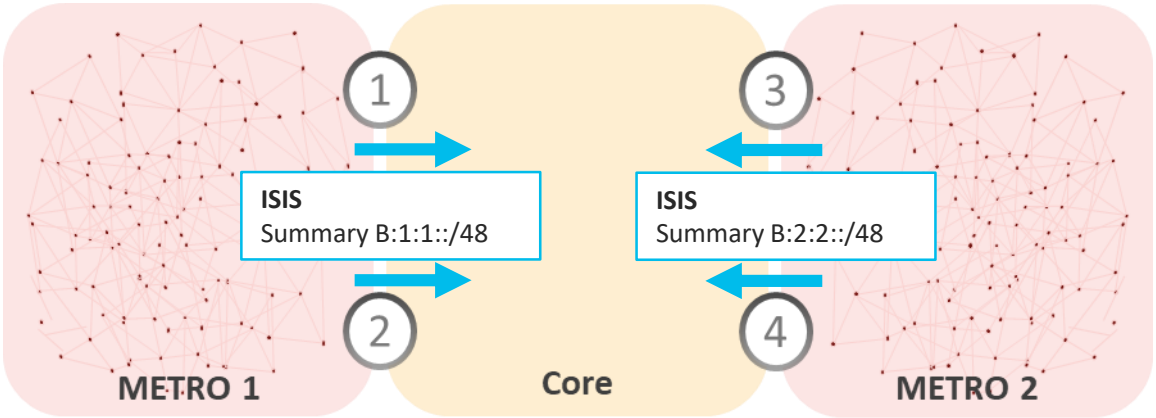


- 20-bit entropy
- No additional protocol
  - infamous mpls entropy label

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 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, iptables, nftables, snort, ExaBGP, Contiv-VPP

SoftBank Group

- First commercial SRv6 deployment

業新聞　2018年(平成30年)3月16日(金曜日)

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

ネットワークをシンプルにする

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.

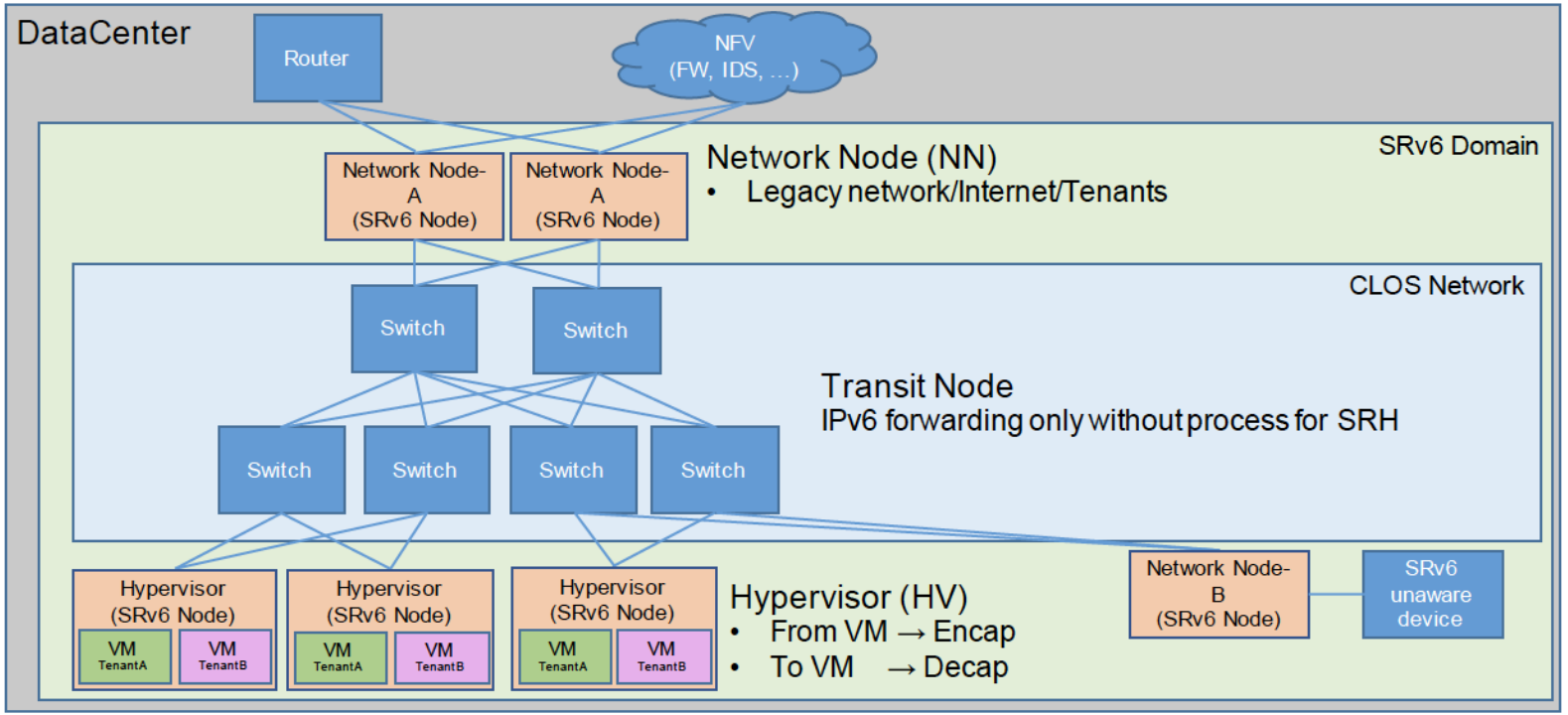
As of the end of 2019, the SRv6 network consists of:
- 1000 Cisco NCS 5500 routers
- 1800 Iliad's Nodeboxes
- The network services 4.5 million mobile subscribers (as of Q3 2019)
- The network is carrying 300 Gbps of commercial traffic at peak hours
- It is expected to grow to more than 4000 Nodeboxes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

Data Plane - Architecture



SRv6 Ecosystem

Network Equipment Manufacturers

CISCO

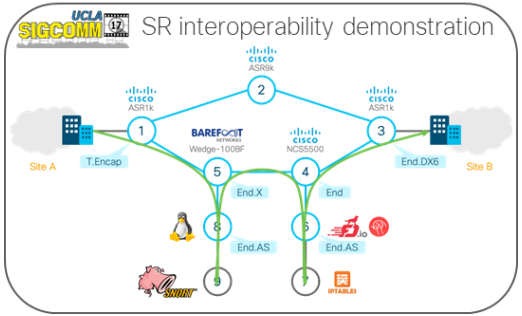
HUAWEI

Merchant Silicon

BAREFOOT NETWORKS

BROADCOM

MARVELL



Open-Source Applications

Pyroute2

SERA IPTABLES

SNORT

TCPDUMP

Contiv

WIRESHARK

exa

Open-Source Networking Stacks

fd.io

Smart NIC

intel

Mellanox TECHNOLOGIES

NFV Partners

TREND MICRO

ENEA Qosmos Division

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

Conclusion

Simplicity Always Prevails

~~LDP~~

~~RSVP-TE~~

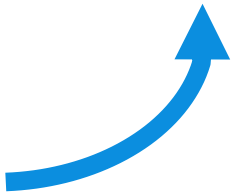
~~Inter-AS Option A/B/C~~

~~MPLS~~

~~UDP/VxLAN~~

~~NSH~~

Furthermore with more scale and functionality



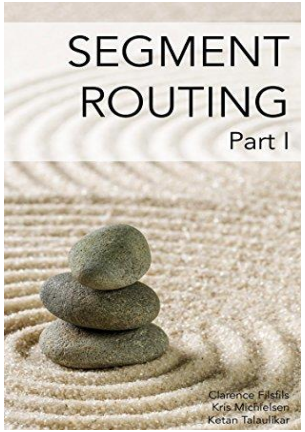
Segment Routing conclusion

- Strong industry support
- Fantastic deployment rate
- Bold architecture: network programming
- Numerous use-cases
  - FRR, TE, SDN, Overlay with SLA, NFV, Spray, SD-WAN, 5G & NS, ...

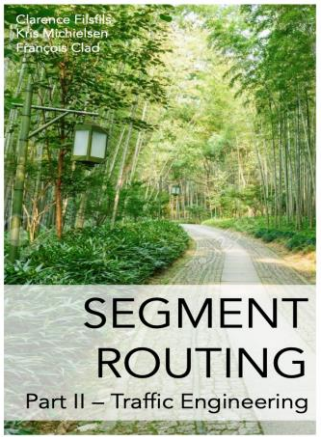
Partnering

- Track-record collaboration with operator
  - Focus on real operator needs
  - Seamless Deployment
  - Standardization
  - Multi-Vendor consensus
- Looking forward to working together

Stay up-to-date



amzn.com/B01I58LSUO



amazon.com/dp/B07N13RDM9

twitter.com/SegmentRouting

segment-routing.net

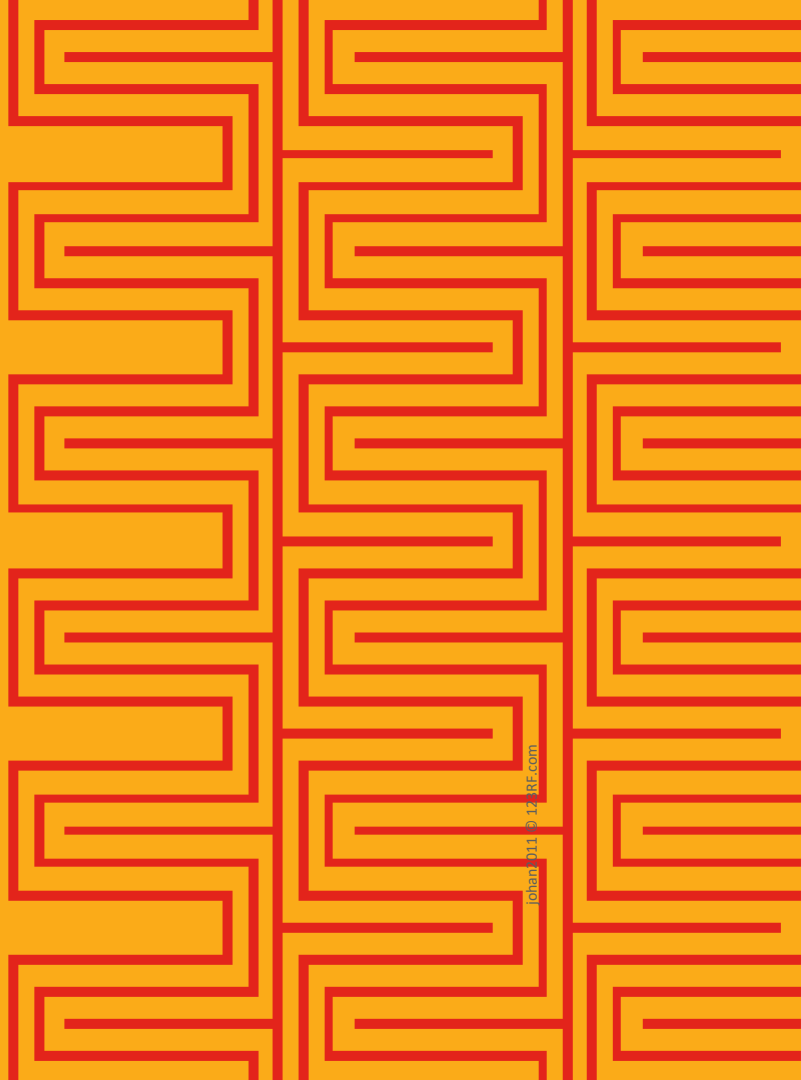
facebook.com/SegmentRouting/

linkedin.com/groups/8266623

Thank you!

ask-segment-routing@cisco.com

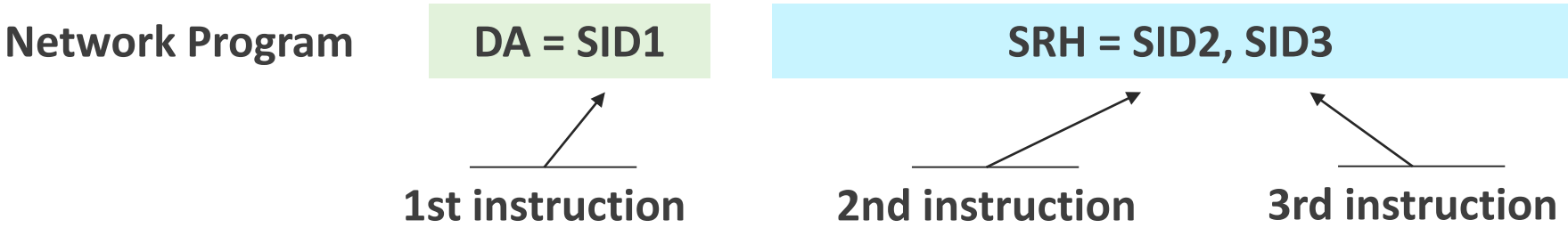
Micro-Program



A new set of (micro)-instructions

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

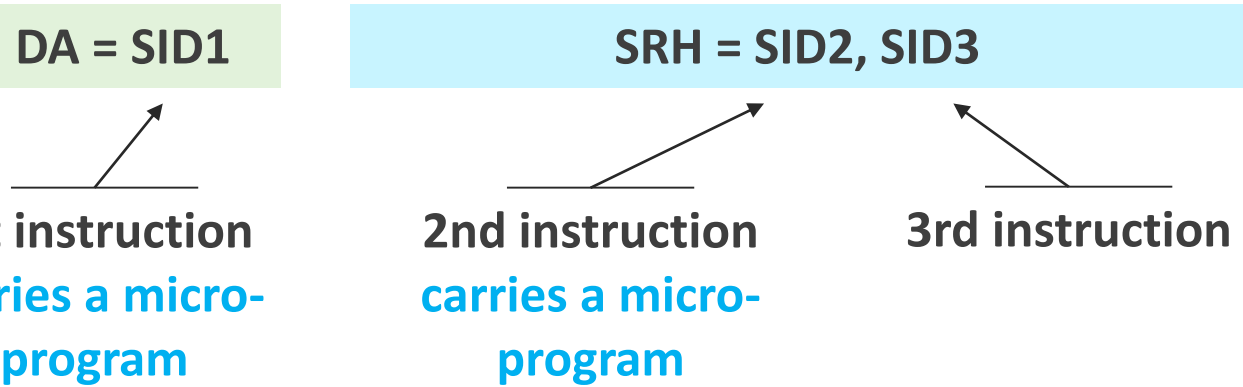
Network Program



- 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

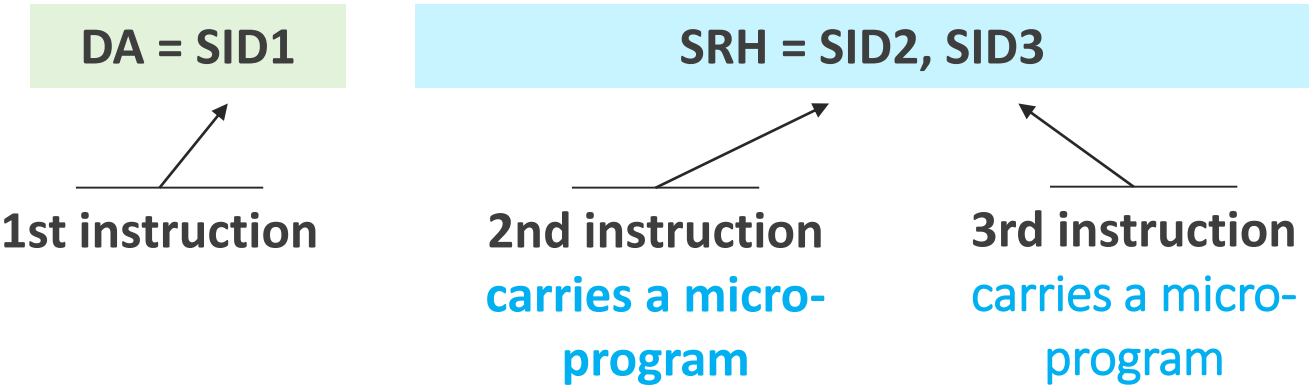
Any instruction could hold a micro-program

Network Program



Any instruction could hold a micro-program

Network Program



Micro-Program in an SRv6 SID

SRv6 SID = 128 bits = 8 groups of 4 nibbles

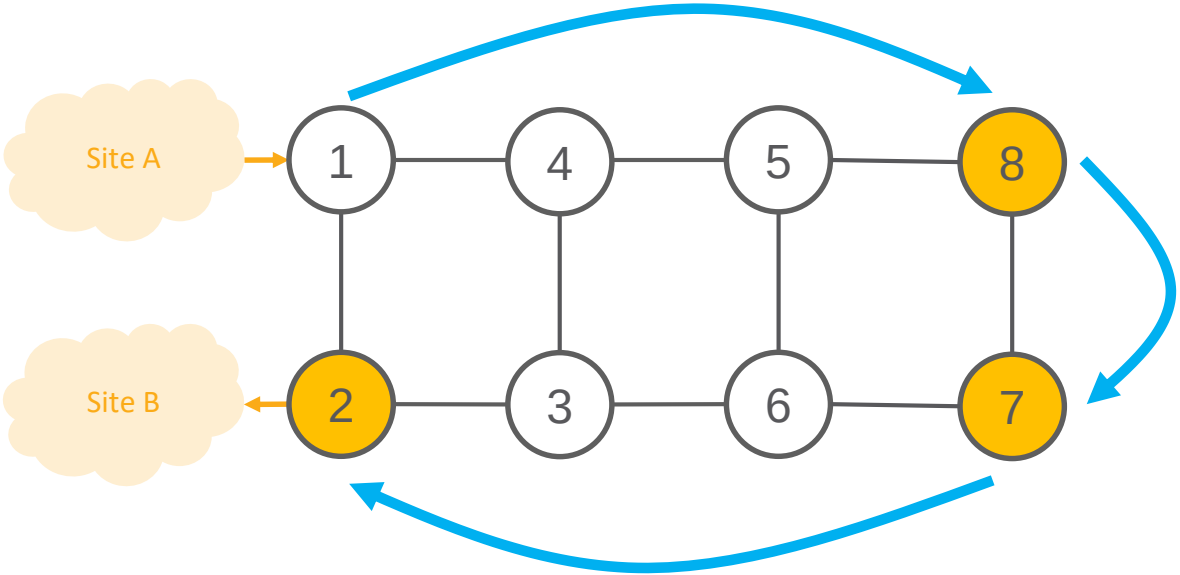
Assuming an allocation block in /32 (B:B::/32)

Assuming a micro-instruction ID in 4 nibbles

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

6 micro instructions per SRv6 Instruction

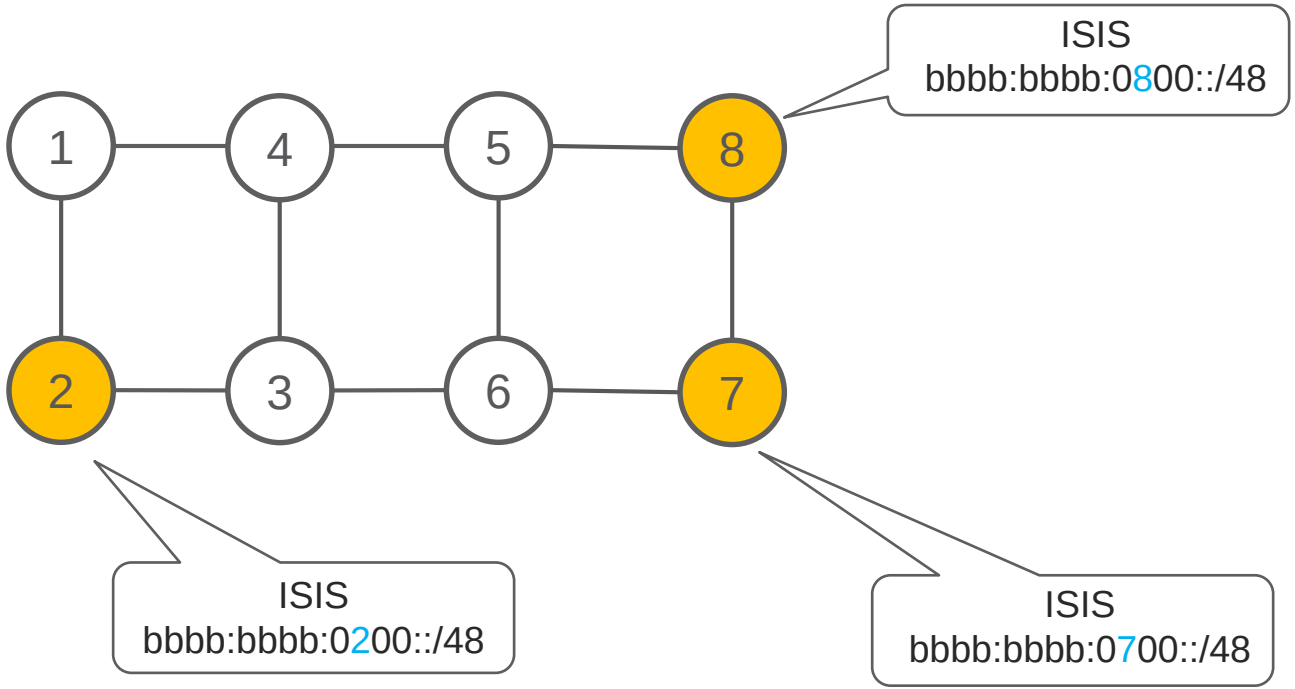
Illustration: go to 8 then 7 then 2 and decaps



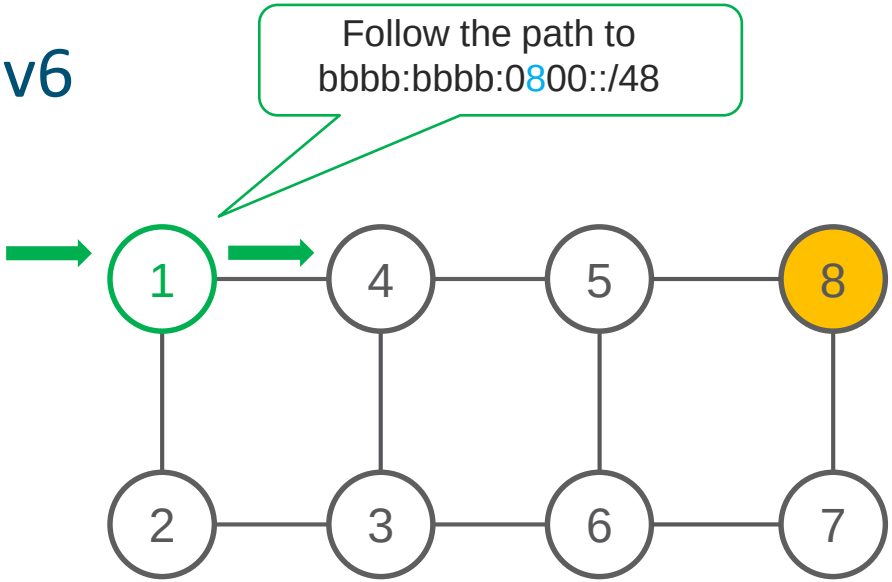
One single micro-program in the DA is enough

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

Basic IP Routing: no new extension

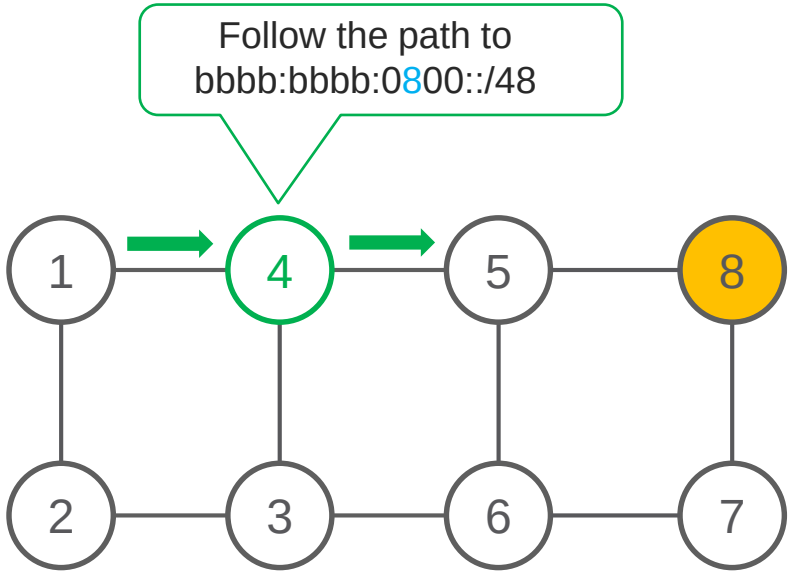


@1: basic IPv6



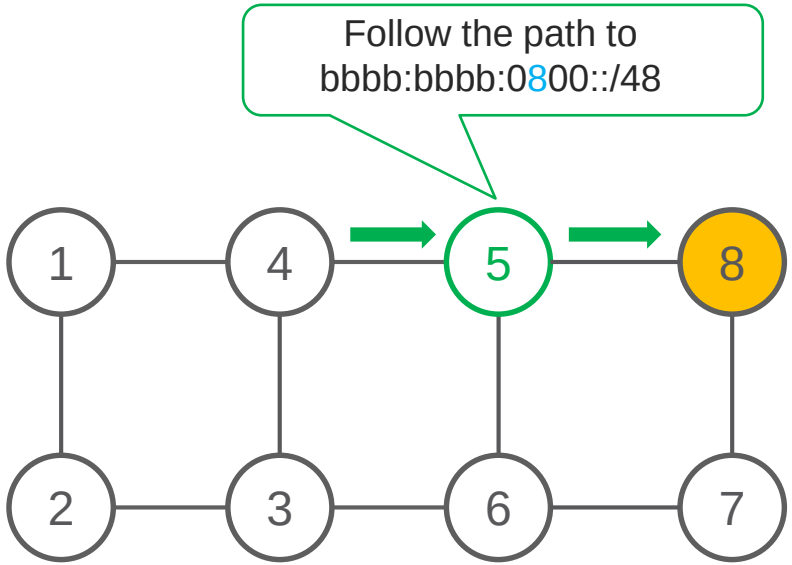
DA = bbbb:bbbb:0800:0700:0200:0000:0000:0000

@4: basic IPv6



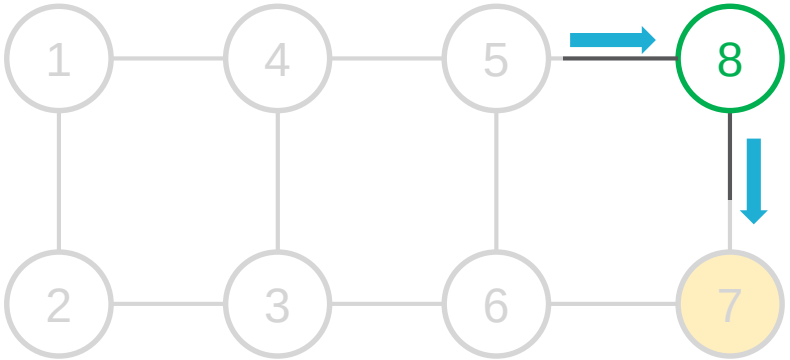
DA = bbbb:bbbb:0800:0700:0200:0000:0000:0000

@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

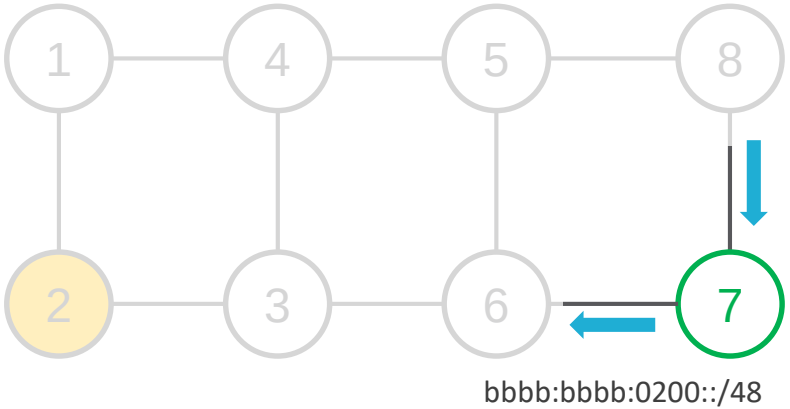
FIB Longest-Match bbbb:bbbb:0800::/48 → SRv6 Instruction:

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:

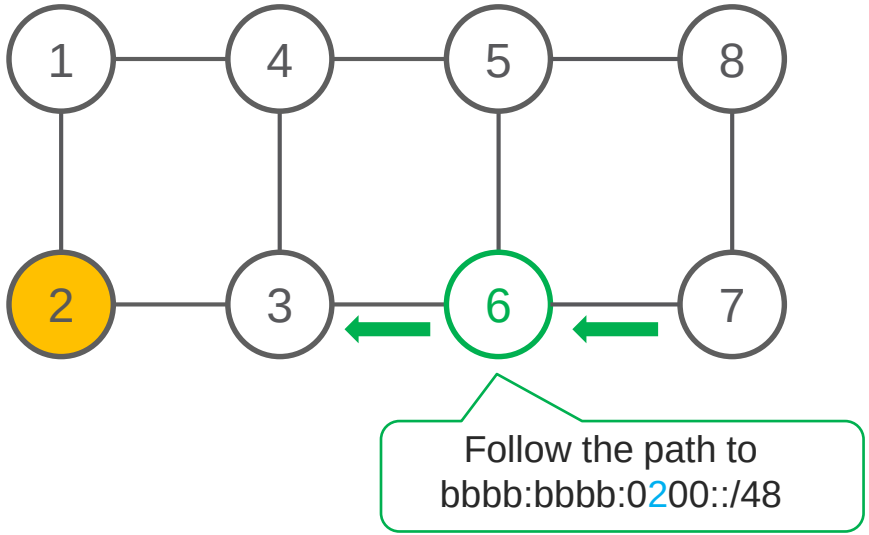
FIB Longest-Match bbbb:bbbb:0700::/48 → SRv6 Instruction:

Shift micro-Program by one micro-Instruction

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

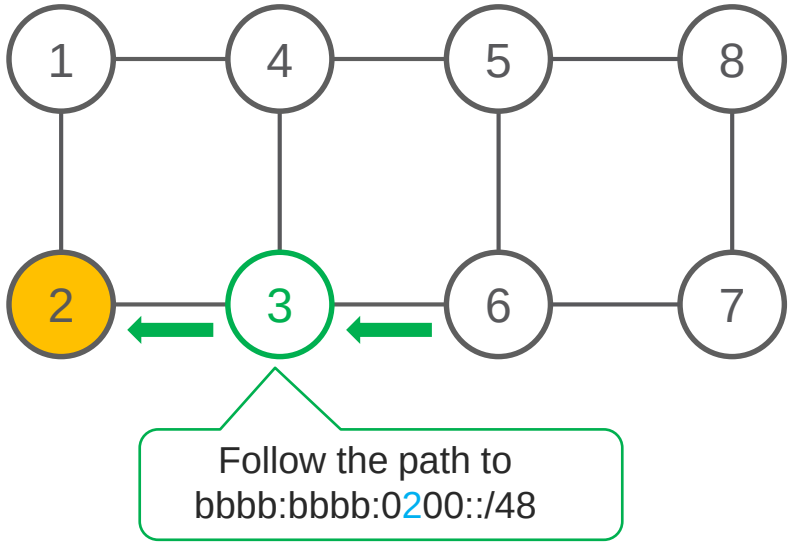
Lookup the updated DA and forward

@6: basic IPv6



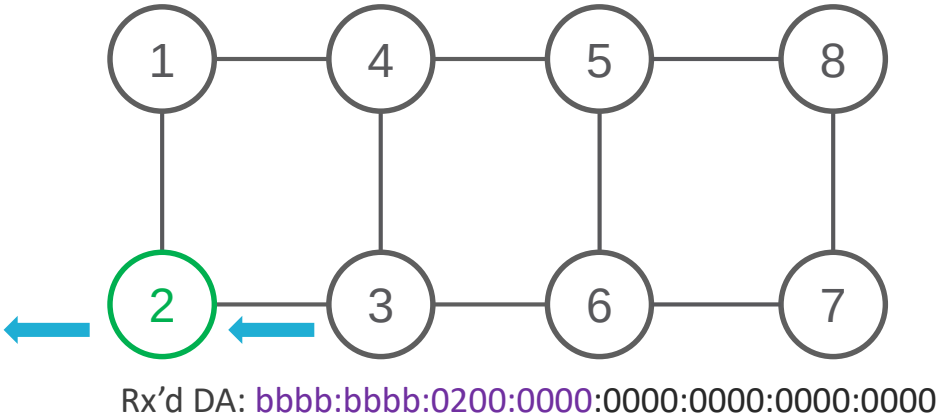
DA = bbbb:bbbb:0200:0000:0000:0000:0000:0000

@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

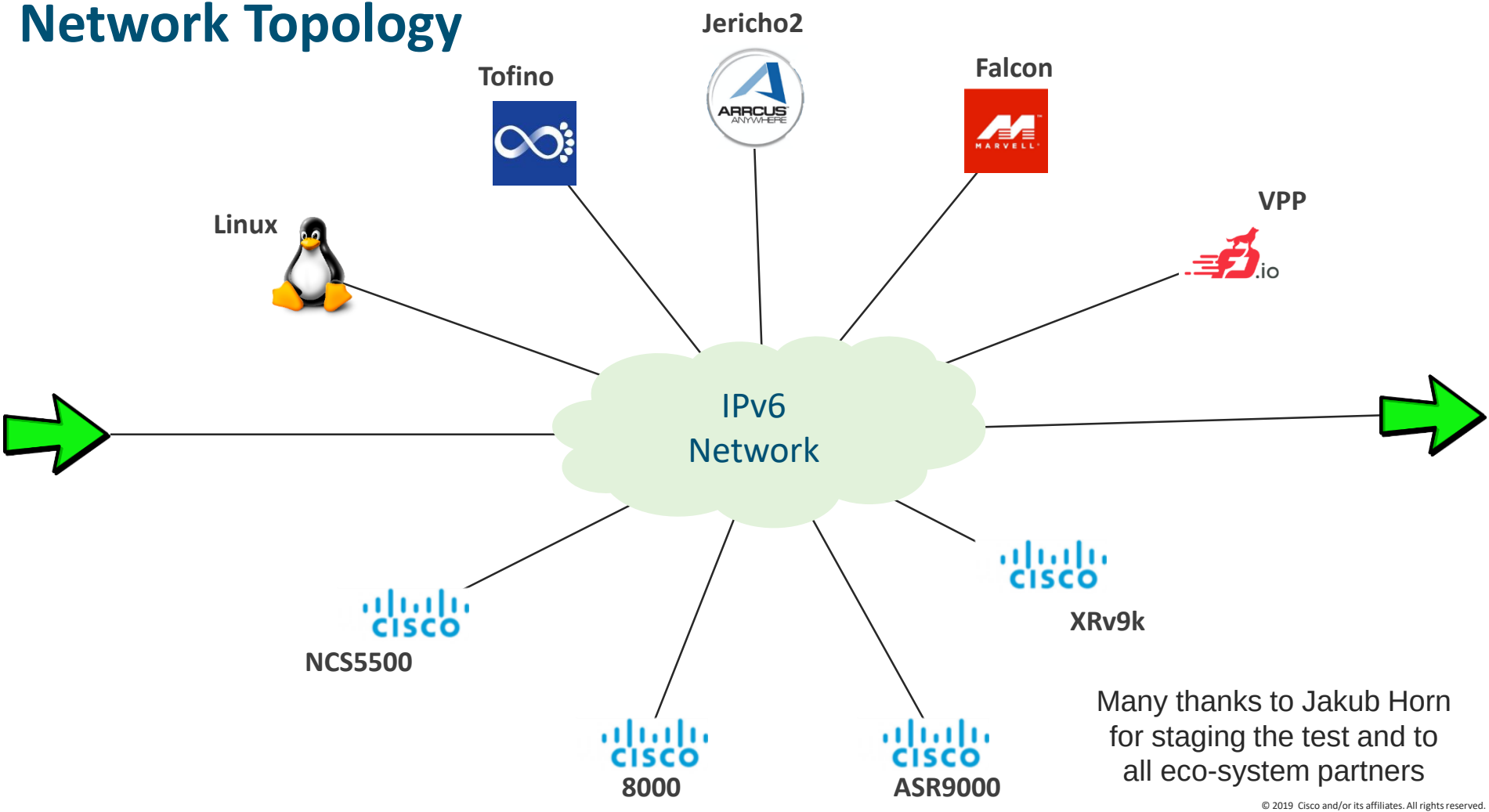
Alibaba

- 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