SRv6 uSID in SONiC

SONiC Summit, Vienna, 2024

Ahmed Abdelsalam

Engineering Technical Leader, Cisco Systems

IP is back and better than ever.

Build anything

Simplified, scalable, and versatile networks that are self-sufficient

Self-sufficiency is standard

End-to-end policy

- From Host to Internet through DC, Access, Metro, Core, Cloud
- No protocol conversion or gateways at domain boundaries

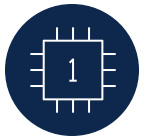
Any service, without any shim

- VPN, Slicing, Traffic Engineering, Green Routing, FRR, NFV

Better scale, reliability, cost, and seamless deployment in Brownfield

Essential embedded assurance

Active probing between Fabric Edges **along all ECMP paths**



High-capacity probe generation and ingestion powered by Silicon One (14MPPS)

Continuous **routing monitoring**

Advanced analytics and intelligent service optimization driven by AI

Measure everything

Embedded SLA monitoring and IPM within the network is essential

IP is back and better than ever.

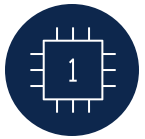
Build anything

Simplified, scalable, and versatile networks that are self-sufficient

SRv6 uSID

Essential embedded assurance

Active probing between Fabric Edges **along all ECMP paths**



High-capacity probe generation and ingestion powered by Silicon One (14MPPS)

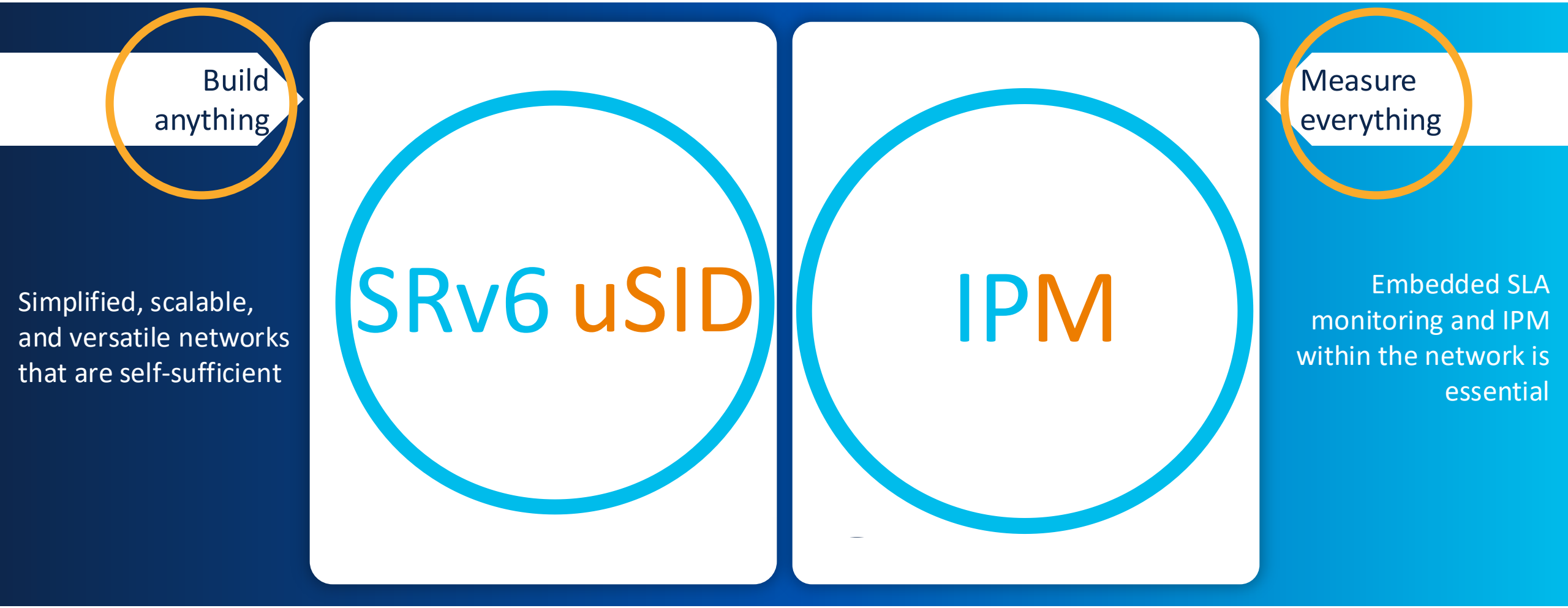
Continuous **routing monitoring**

Advanced analytics and intelligent service optimization driven by AI

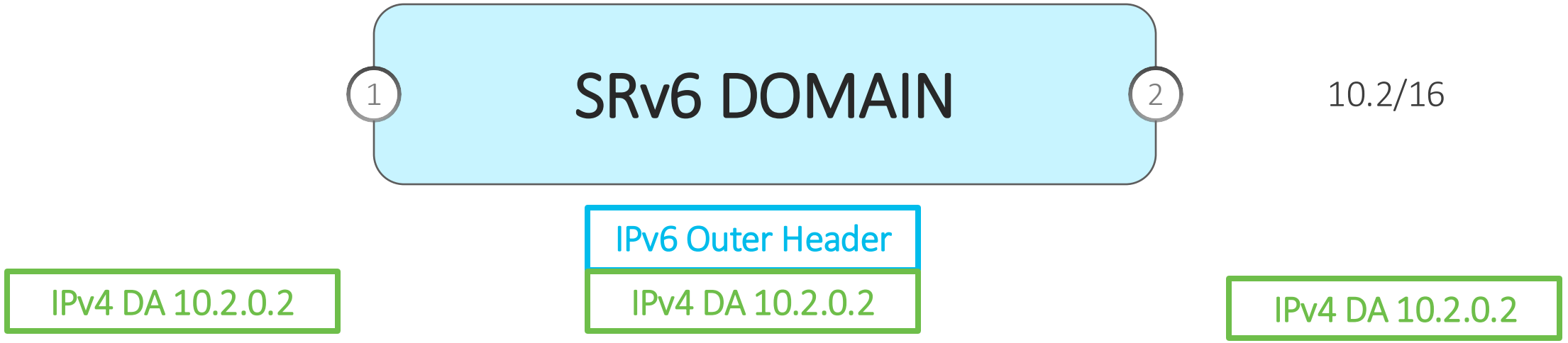
Measure everything

Embedded SLA monitoring and IPM within the network is essential

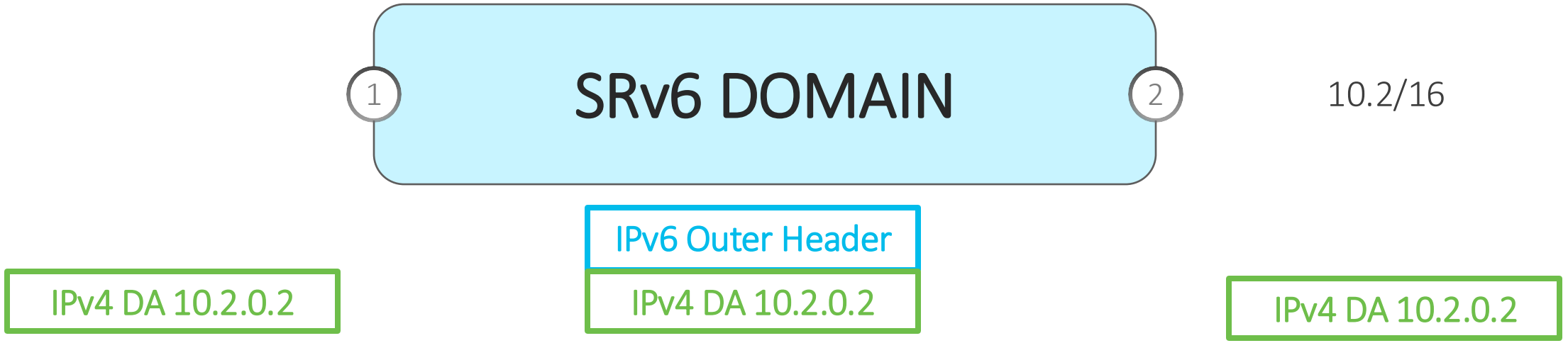
IP is back and better than ever.



Transparent Service



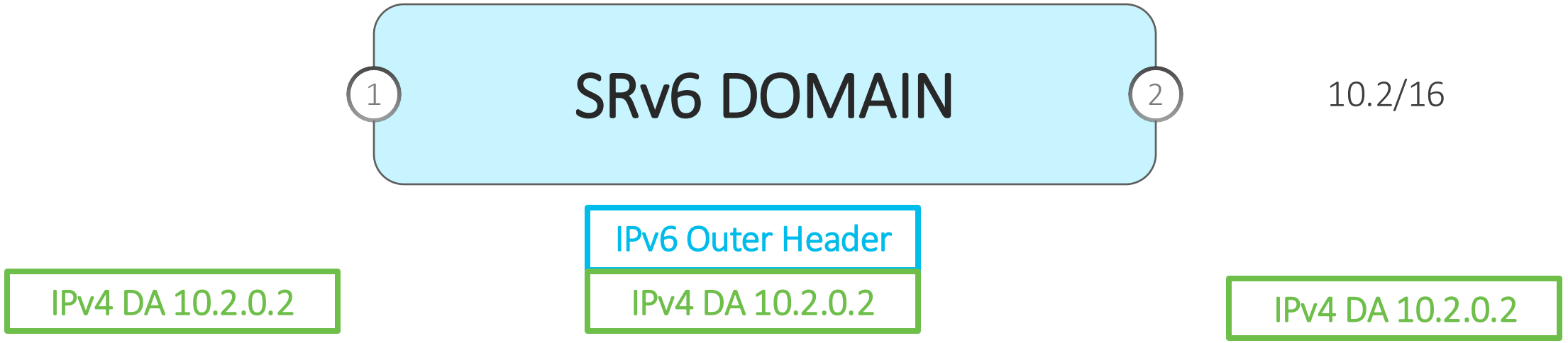
Transparent Service



DA = FCBB:BBBB:0000:0000:0000:0000:0000:0000

IPv6 Block uSID1 uSID2 uSID3 uSID4 uSID5 uSID6

Transparent Service



DA = FCBB:BBBB:0000:0000:0000:0000:0000:0000

IPv6 Block uSID1 uSID2 uSID3 uSID4 uSID5 uSID6

uSID can bound to any instruction (VPN, TE, NFV)

SRv6 uSID

- Build Anything
  - Any combination of underlay, overlay, service chaining, security…
  - VPN, Slicing, Traffic Engineering, Green Routing, FRR, NFV
- Any Domain
  - Access, Metro, Core, DC, Host, Cloud
  - End-to-End Stateless Policy
  - No protocol conversion or gateways at domain boundaries
- Seamless Deployment in Brownfield
- Built day-1 for Automation
- Standardized, Rich Eco-system, Rich Open Source (SONiC)

Rich SRv6 uSID Ecosystem

Network Equipment Manufacturers

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

Merchant Silicon

Broadcom, Marvell, Barefoot Networks (Intel)

Open-Source Applications

Pyroute2, IPTABLES SERA, GoBGP, exa, BGP, eBPF, cilium, Wireshark, FRRouting, Snort, GoBMP

SRv6 uSID

Open-Source Networking Stacks

SONiC, Linux, fd.io, P4, Open Compute Project, OvS Open vSwitch, VyOS

Smart NIC / DPU

Intel, NVIDIA

Partners

kaloom, Trend Micro, ENEA Qosmos Division, NoviFlow, ixia A Keysight Business, spirent Promise. Assured., Red Hat, f5, 6WIND, Keysight Technologies

SRv6 is Proposed Standard

Architecture

- SR Architecture – RFC 8402
- SRTE Policy Architecture – RFC 9256

Data Plane

- SRv6 Network Programming – RFC 8986
- IPv6 SR header – RFC 8754

Control Plane

- SRv6 BGP Services – RFC 9252
- SRv6 ISIS – RFC 9352
- SR Flex-Algo – RFC 9350

Operation & Management

- SRv6 OAM – RFC 9259
- Performance Management – RFC 5357

Strong Commitment and Leadership

Editor of	96% IETF RFCs
Co-author of	100% IETF RFCs

Simplicity Always Prevails

Over 80000 uSID routers deployed

Outperform MPLS - Daniel Voyer (Bell Canada)

- Native Optimum Slicing
  - SLID is encoded in Flow Label
- HW Linerate Push: 3 times better
  - J2 uSID linerate push: 30 uSIDs >> 10 MPLS Labels
- HW Counter and FIB consumption: 4 times better
  - uSID requires 4 times less counters and FIB entries than MPLS
- Routing scale: 20 times better
  - uSID supports summarization. MPLS requires host routes.
- Lookup efficiency: 2 to 3 times better
  - uSID can process 2 to 3 SIDs in a single lookup (LPM nature)
- Load-balancing: optimum and deterministic
  - uSID provides HW friendly entropy (fixed offset, shallow)

Bell SRv6 uSID Deployment
Paris 2022

Dan Voyer – Technical Fellow – Bell Canada

Outperforms VxLAN – Gyan Mishra (Verizon)

- Seamless Host support for Network Programming
  - 6 uSID's in outer DA: RFC2460 IPinIP with opaque DA
- TE in the DC
  - elephant flows exist, asymmetric fabrics exist, TE is needed
- TE in the Metro/Core from the host
  - An SRv6 uSID DC allows for the application to control the network program in the metro/core without complex DPI and protocol conversion at the DC boundary,
- uSID DC provides lower MTU overhead (~5%)
  - Lower MTU overhead means lower DC cost
- Vendor, Merchant and SONIC/SAI maturity
  - uSID support across DC vendor (Cisco), Merchant (Cisco, Broadcom, Marvell), Sonic/Sai (Alibaba deployment)

SRv6 uSID DC Use-Case
Paris 2023

Gyan Mishra - Associate Fellow – Verizon

SRv6 uSID enabled SONiC

- 2021: SRv6 uSID in SONIC collaboration started
  - Very rich ecosystem (Alibaba, Cisco, Broadcom, Intel, others)
  - Several major milestones
- 2022: Alibaba Routing Platform
  - SONiC based
  - SRv6 uSID enabled
- 2023: Alibaba SRv6 uSID deployment
  - Alibaba and Cisco Blog
- 2024:
  - TE Policy
  - BFD offload
  - Infra enhancements

Features		SONiC Status
Underlay		
	SRv6 uSID	Available: SONiC 202211 [Link] [Link]
	TE Policy	Target: SONiC 202411
	BFD Offload	Target: SONiC 202411
Overlay		
	BGP VPN	Available: SONiC 202305 [Link]

Eddie Ruan - Senior Staff Engineer – Alibaba

Deployment Outcomes

98% Reduction in ISIS LSP in Core	73% Reduction in total ISIS routes in Core	96% Reduction in number of BGP sessions on Core routers 55 to 2	SRv6 transforms Rakuten IP Transport network to world class converged 5G SDN-ready network
90% Reduction in total routes to data centers 6500+ → 500	86% Reduction in config touch-points 13-15 -> 2	90% Improvement in convergence time 500ms to 50ms	

Rakuten

Akash Agrawak - Technical Director – Rakuten Mobile

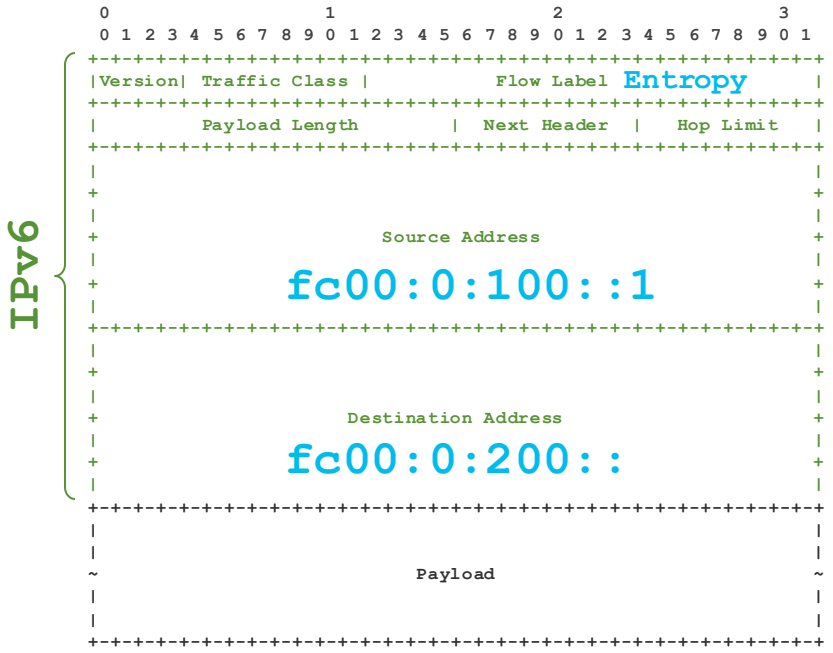
Why SRv6 uSID?

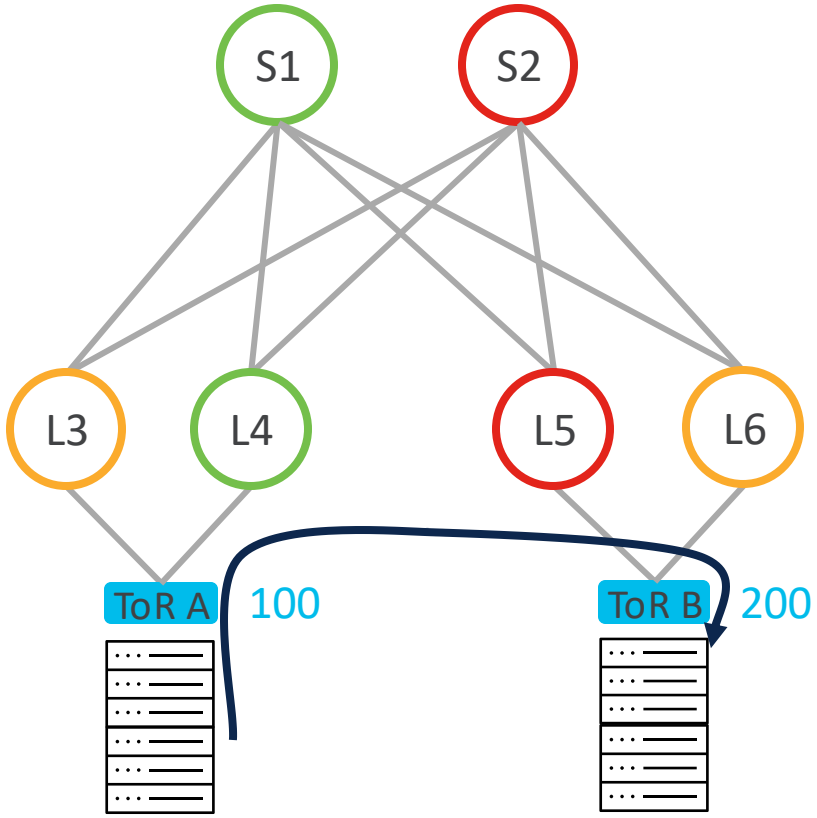
End-to-End: Outperforming existing solution

- The only solution that is end-to-end across segments
  - uSID outperforms MPLS along any comparison analysis
  - uSID outperforms VxLAN along any comparison analysis
  - uSID outperforms NSH along any comparison analysis
- Simple IPv6 encapsulation
  - 6 uSIDs in the IPv6 DA
- Benefits
  - VxLAN/MPLS gateways are eliminated
  - NSH is eliminated
  - TE/FRR/VPN/Service Chaining is naturally integrated
  - Fulfills the DC Traffic Engineering requirements
  - Integrated Measurements
    - Deterministic Per-Path Measurements

Simple IPv6 encapsulation

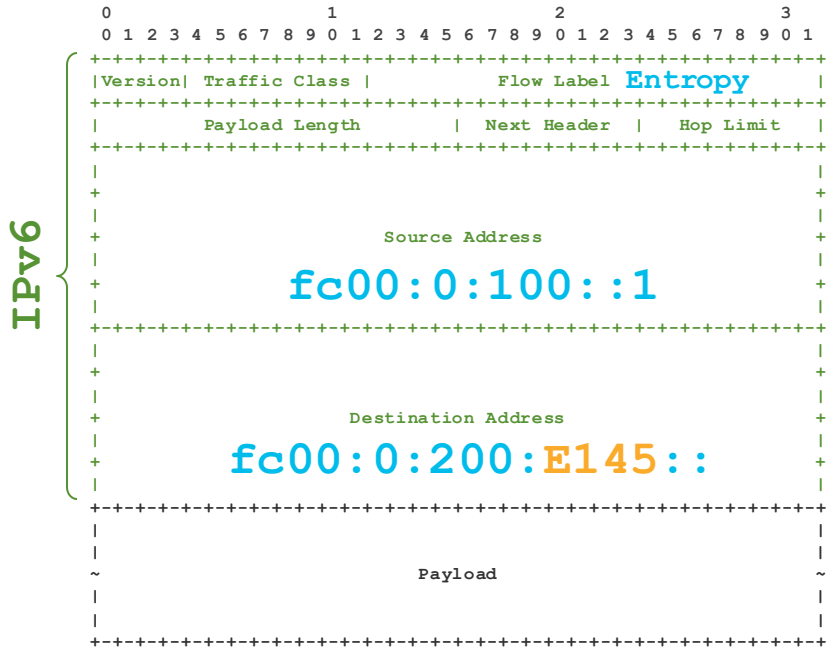
- Best effort overlay from 100 to 200 only requires simple IPv6 encap

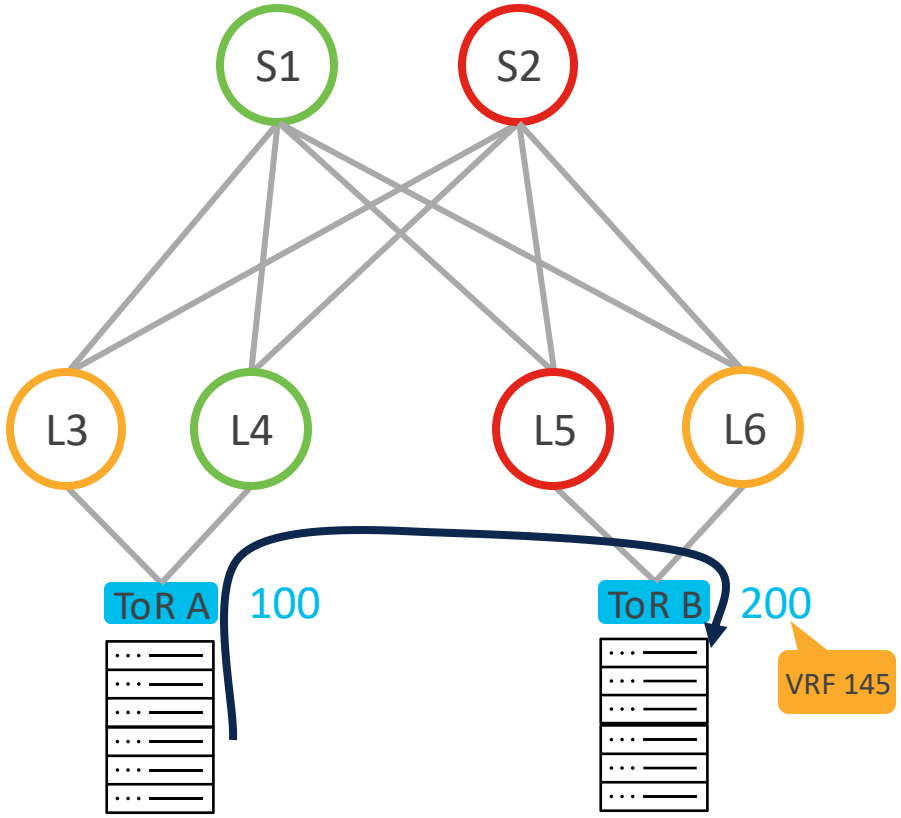




Simple IPv6 encapsulation

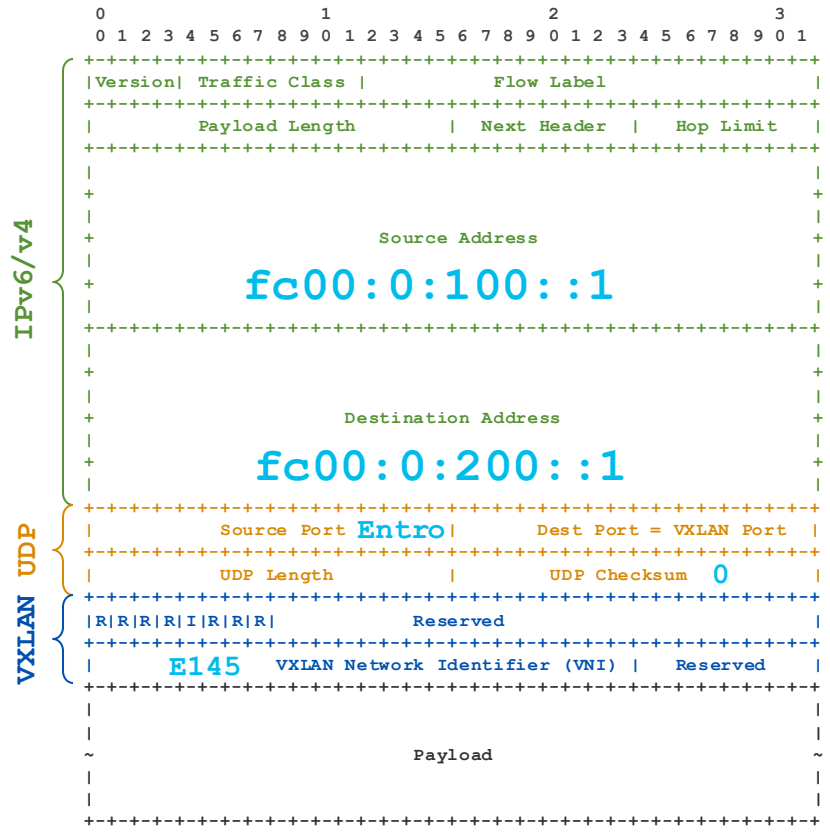
- Best effort overlay from 100 to 200 only requires simple IPv6 encap
- If we need to create a VPN service, we only add a single uSID

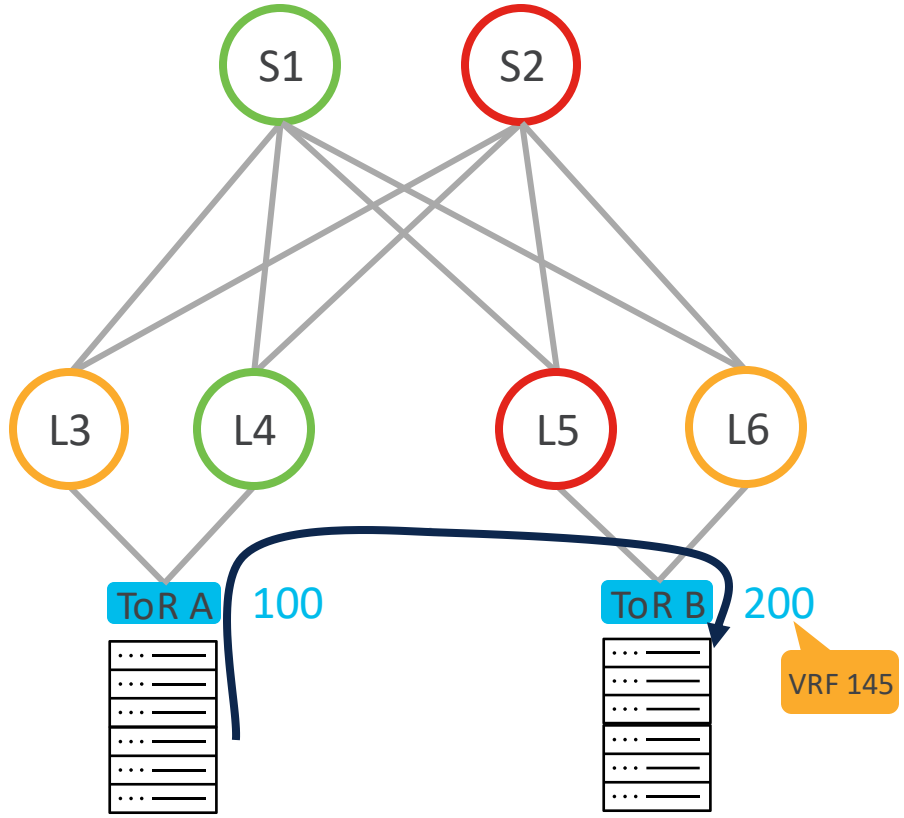




Non-Simple Encapsulation

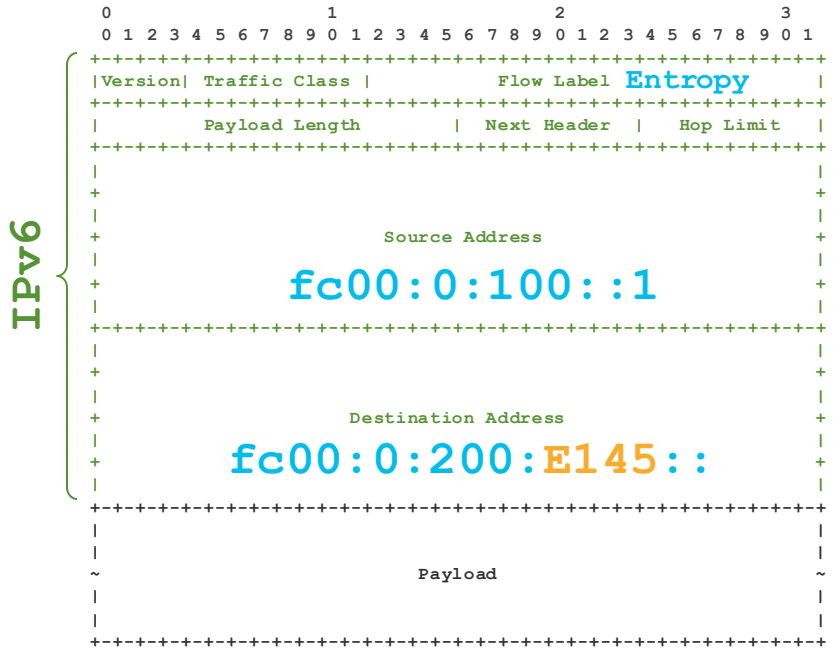
- To compare it with VXLAN….

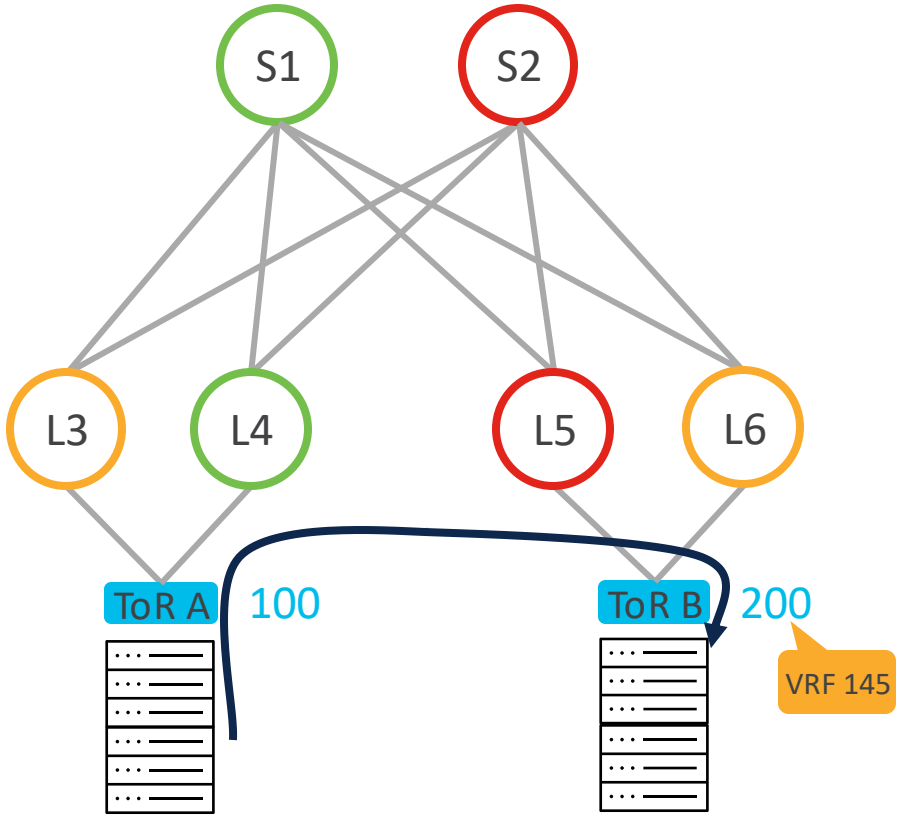




Simple IPv6 encapsulation

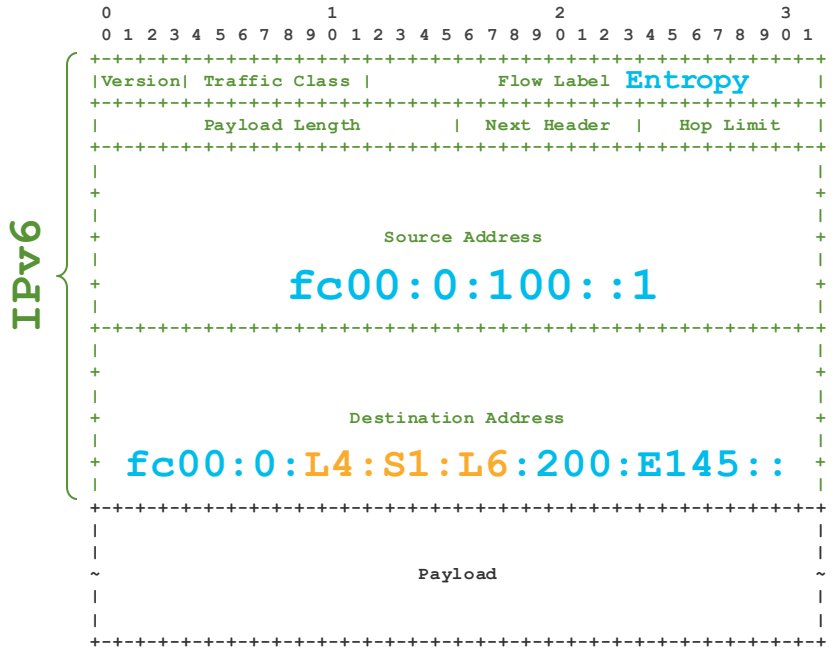
- Best effort overlay from 100 to 200 only requires simple IPv6 encap
- If we need to create a VPN service, we only add a single uSID

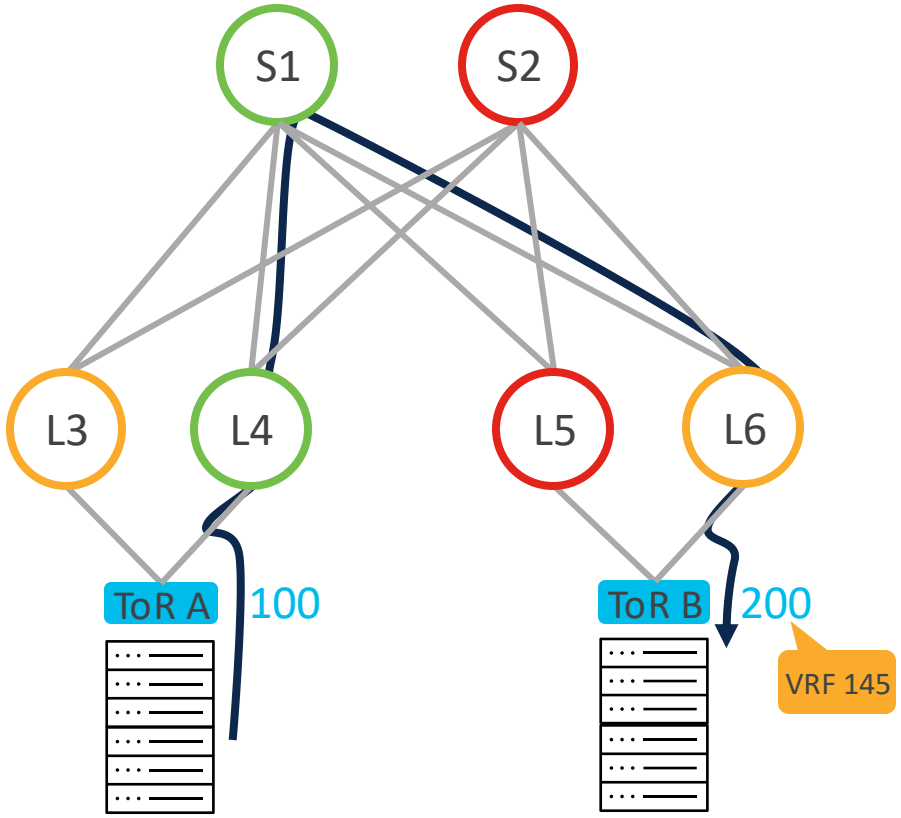




Simple IPv6 encapsulation

- Best effort overlay from 100 to 200 only requires simple IPv6 encap
- If we need to create a VPN service, we only add a single uSID
- If we want to enforce a Traffic Engineered path <4, 1, 6> we add three uSIDs:





- Apple-to-apple comparison:
  - SRv6 uSID provides more functionality AND less overhead than VXLAN

SRv6 uSID in SONiC

The Journey

- 2021: SRv6 uSID in SONiC collaboration started
  - Very rich ecosystem (Alibaba, Cisco, Broadcom, Intel, others)
  - SRv6 uSID support in SAI (1.19) and SONiC (202111)
- 2022: Alibaba Routing Platform
  - SONiC based and SRv6 uSID enabled
  - SRv6 uSID VPN support in SONiC (202211)
- 2023: Alibaba SRv6 uSID deployment
  - Across all of China
  - SRv6 uSID VPN enhancement in SAI (1.12) and SONiC (202305)
  - SONiC Routing WG (Link) and Alibaba & Cisco Blog (Link)
- 2024: SONiC PhoenixWing Initiative
  - Endorsed by Alibaba, Cisco, Microsoft, Broadcom, Linux Foundation (Link)
  - Phase1: SRv6 uSID readiness in SONiC 202411
  - Phase2: IP Measurements (IPM)

SAI: SRv6 uSID Support

- SAI Header
  - https://github.com/opencomputeproject/SAI/blob/master/inc/saisrv6.h
  - SRv6 Headend/Encapsulation behaviors
  - SRv6 Endpoint behaviors
    - SRv6 Underlay
    - SRv6 Overlay/VPN
    - Endpoint Behavior flavors
  - SRv6 Traffic Engineering behaviors
- SAI Adapter
  - Cisco Silicon One has full implementation of SRv6 uSID SAI header
  - Deployed at Alibaba

SAI Header

- List of supported behaviors
  - Headend: H.Insert, H.Insert.Red, H.Encaps, H.Encaps.Red
  - Endpoint: End, End.X, End.T, End.DX4, End.DX6, End.DT4, End.DT6, End.DT46, uN, uA, End.B6.Encaps, End.B6. Encaps.Red, End.B6.Insert, End.B6.Insert.Red
  - Flavors: PSP, USP, USD
  - Traffic Engineering: SRv6 SID List & stats (Packets & Bytes)
- RFC8986: SRv6 Network Programming
  - https://www.rfc-editor.org/rfc/rfc8986.html
- High Level design & documentation
  - https://github.com/opencomputeproject/SAI/blob/master/doc/SAI-IPv6-Segment-Routing-Update.md
  - https://github.com/opencomputeproject/SAI/blob/master/doc/SAI-IPv6-Segment-Routing-VPN.md

SONiC: SRv6 uSID Support

- Supported SRv6 features
  - SRv6 Headend/Encapsulation
  - SRv6 Endpoint behaviors
    - SRv6 Underlay
    - SRv6 Overlay/VPN
    - Endpoint Behavior flavors
  - SRv6 Traffic Engineering behaviors
- Across all SONiC component
  - ConfigDB, AppDB, OrchAgent, ..
  - FPM Module & Fpmsyncd (for the integration with FRR)

SONiC Components: SRv6 uSID Support

- HLD & DB Schema:
  - https://github.com/sonic-net/sonic-swss-common/blob/master/common/schema.h
  - https://github.com/sonic-net/SONiC/blob/master/doc/srv6/
- ConfigDB
  - SRV6_SID_LIST_TABLE, SRV6_MY_SID_TABLE, SRV6_POLICY_TABLE, SRV6_STEER_MAP
- AppDB
  - SRV6_SID_LIST_TABLE, SRV6_MY_SID_TABLE, ROUTE_TABLE
- Orchagent
  - New orchagent called “srv6orch” to program SRv6 in the SDK via SRv6 SAI API
- FPM Module & Fpmsyncd
  - Extended to support SRv6 and integration with FRR

FRR: SRv6 uSID support

- BGP: SRv6 uSID L3VPN (RFC 9252)
  - VPNv4: IPv4 address family
  - VPNv6: IPv6 address family
  - VPNv46: IPv4 and IPv6 using a single SRv6 SID
- ISIS: SRv6 uSID underlay (RFC 9352)
- Zebra:
  - SRv6 Locator support
  - SRv6 SID Manager
    - Responsible for SID allocation/management.
    - Ensure full interoperability between FRR and other vendors including Cisco IOS XR and many others

Alibaba deployment & Alibaba Routing Platform

SONiC Alibaba Group

USE CASE

SONiC Alibaba Data Center Network Use Case

Overview

Alibaba Cloud's worldwide infrastructure spans 28 regions, providing a combined total of 86 available zones and an extensive network featuring over 3,200 CDN nodes.

The cornerstone of this cloud infrastructure is built on Alibaba predictable network infrastructure. The overarching strategy behind the development of the Alibaba network revolves around the utilization of a unified protocol stack, IPv6/SRv6, alongside single-chip white boxes which the control plane is based on SONiC and the data plane is programmable. To establish a predictable network, it is important to strike a harmonious balance between optimizing both endpoints and the network infrastructure. Moreover, it's crucial to foster a collaborative design that bridges applications and the network, resulting in an endpoint-network synergy system.

https://sonicfoundation.dev/wp-content/uploads/sites/21/2024/01/sonic-usecase-alibaba-data-center-network_011824a.pdf

Authors:

- *Pablo Camarillo, Technical Lead at Cisco*
- *Clarence Filsfils, Cisco Fellow and Cisco*
- *Ahmed Abdelsalam, Technical Lead at Cisco*
- *Dennis Cai, Head of Network Infrastructure at Alibaba Cloud*
- *Eddie Ruan, Director of Network System Software at Alibaba Cloud*
- *Yuanchao Su, Director of Network Architecture & Planning at Alibaba Cloud*

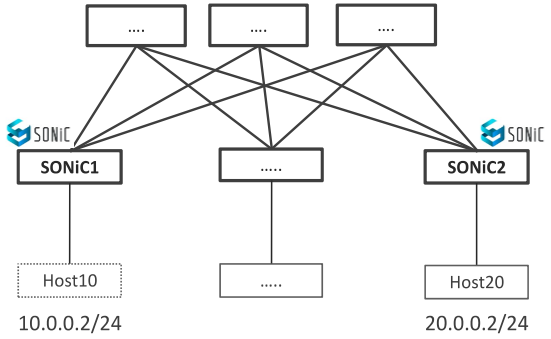
https://sonicfoundation.dev/driving-innovation-alibaba-and-cisco-co-dev-srv6-sonic-router/

Demo

- SRv6 uSID in SONiC with Cisco Silicon One
- L3VPN between SONiC devices
  - https://www.segment-routing.net/demos/20230602-demo-srv6-usid-in-sonic-with-cisco-silicon-one/

Demo: SRv6 uSID L3VPN in SONiC with Cisco Silicon One

- SRv6 uSID DC Fabric
- SONiC + Cisco Silicon One Q200.
- L3VPN between SONiC1 and SONiC2 to connect host10 and Host20



SONiC Routing WG

SONiC Working Group on Routing sonic-wg-routing@lists.sonicfoundation.dev

Working Group on Routing Area -- Routing Performance/Scale etc.

Group Information

278 Members
220 Topics , Last Post: Sep 12
Started on 04/02/23

- Contributors from both FRR and SONiC community (Alibaba, MSFT, Cisco, NVidia, Dell,..)
- The goal is to improve white box supports in routing.
- Initial focus areas
  - SRv6 uSID
  - Communication among FRR, kernel and SONiC
  - BGP loading time
  - Others

SONiC PhoenixWing Initiative

- Launched at the 4th Networking Open-Source Tech Conference in Beijing (May 25, 2024).
  - Endorsed by Alibaba, Cisco, Microsoft, Broadcom, Linux Foundation
  - Phase1: SRv6 uSID readiness in SONiC 202411
  - Phase2: IP Measurements (IPM)
- SONiC Foundation Blog
  - https://sonicfoundation.dev/phoenixwing-initiative-kick-off-in-beijing/
- Endorsement videos
  - https://www.segment-routing.net/news/2024-05-25-phoenixwing-initiative-kick-off-in-beijing/

Conclusion

- SRv6 uSID
  - Delivers any service (VPN, TE, NFV)
  - Across any domain (Access, Metro, Core, DC, Host, Cloud)
  - No shim Layer. No protocol conversion or gateways at domain boundaries
  - Outperforms VXLAN/MPLS/NSH
  - Fully standardized and very rich ecosystem
- Fully supported in SAI/SONiC/FRR
  - Deployed at Alibaba
  - Very rich SONiC ecosystem
  - SONiC Routing WG
  - PhoenixWing


CISCO