

eCore

Network Architecture and Whitebox Evolution of Alibaba's Service-Oriented DCI Network

Dennis Cai (presenting on behalf of **Roy Jiang**), Alibaba Cloud

Manish Mukherjee, Cisco



October 13-16, 2025
San Jose, California



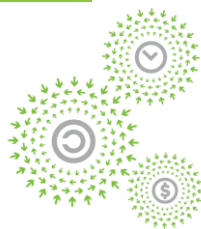


eCore

Network Architecture and Whitebox Evolution of Alibaba's Service-Oriented DCI Network

Dennis Cai (presenting on behalf of **Roy Jiang**) , Alibaba Cloud

Manish Mukherjee, Cisco



OPEN
PLATINUM®

Alibaba Cloud Global Network Footprint



29 Regions

91 AZs

3200+ Edge sites

Alibaba Global Network (AGN)

Current WAN

M2M Traffic

M2U Traffic

BGP

L3VPN

SR-TE

ISIS

Diffserv

v4/v6

MPLS

Router

Vendors



OCP
GLOBAL
SUMMIT

| 2025

Leading the Future of AI.

Cisco Confidential

eCore – Alibaba's Next Generation Predictable DCI Network

Architectural Goals

Simplicity

Stability

Service-Oriented Network

with Massive Scale

Design Principles

Simplicity
Single-chip & Single Stack

Composable Architecture
Easy to scale

Minimize Blast Radius

Service-Oriented

Operation Excellence

**Build and operate DCI
like DCN (fully automated)**

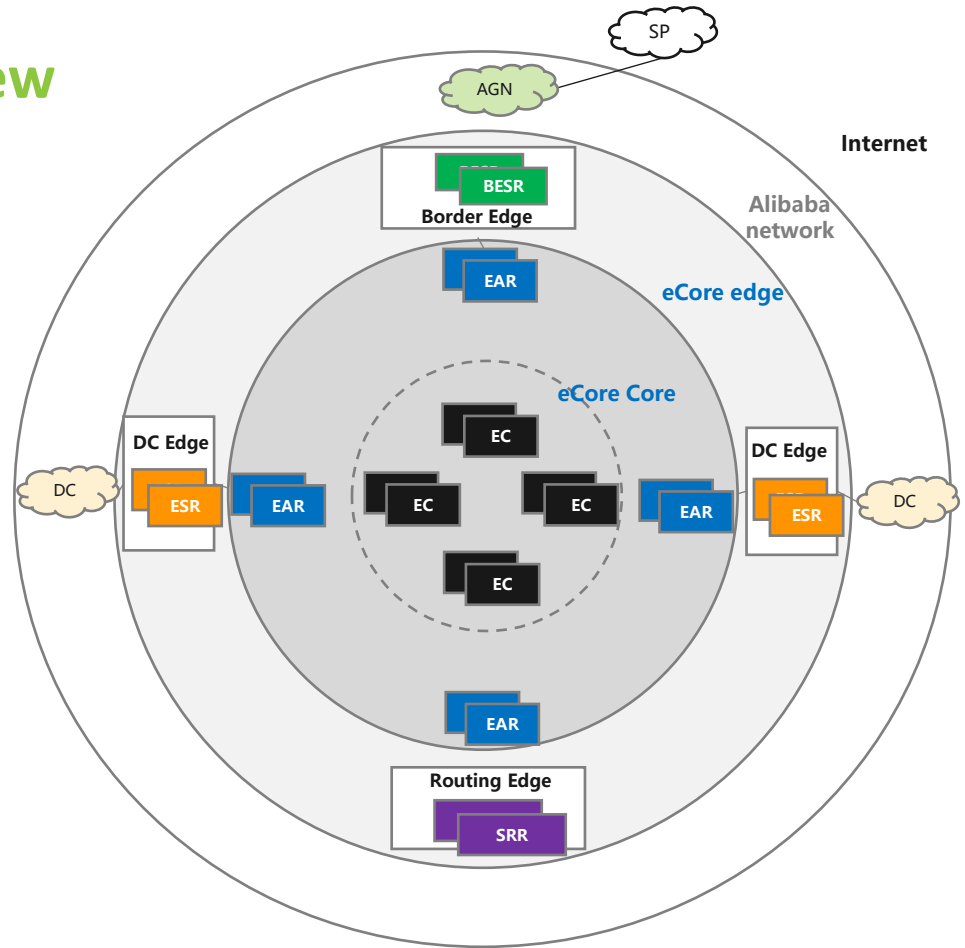
High Level Architecture Overview

Edge

- Role
 - SRv6 VPN PE router – overlay
- Categories
 - DC edge - ESR
 - Border edge - BESR
 - Routing edge - SRR (Service Route Reflector)

Core

- Role
 - SRv6 VPN P router - underlay
 - SRv6 Traffic Engineering
- Layers
 - National Core - EC
 - Regional Core - EAR (collocated with ESR)



Single-chip & Single-Stack

Single Chip

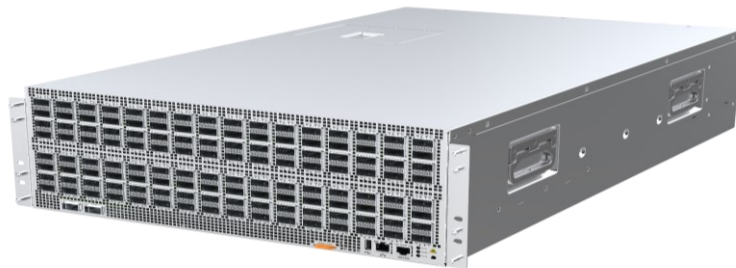
- Pizza-box
- **NO chassis**

Single Stack

- SRv6 only
- **NO MPLS**

In-house Developed Software

- AliNOS (based on SONiC)
- **Feature velocity + live patch**



SRv6 Enable

Deep Buffer

HW offload BFD

MACSec Enable

12.8T

Large LPM

In-house developed white-box router with Cisco Silicon One Q200

Dual-NOS/dual-chip: commercial pizza box from H3C/Ruijie with BRCM ASIC

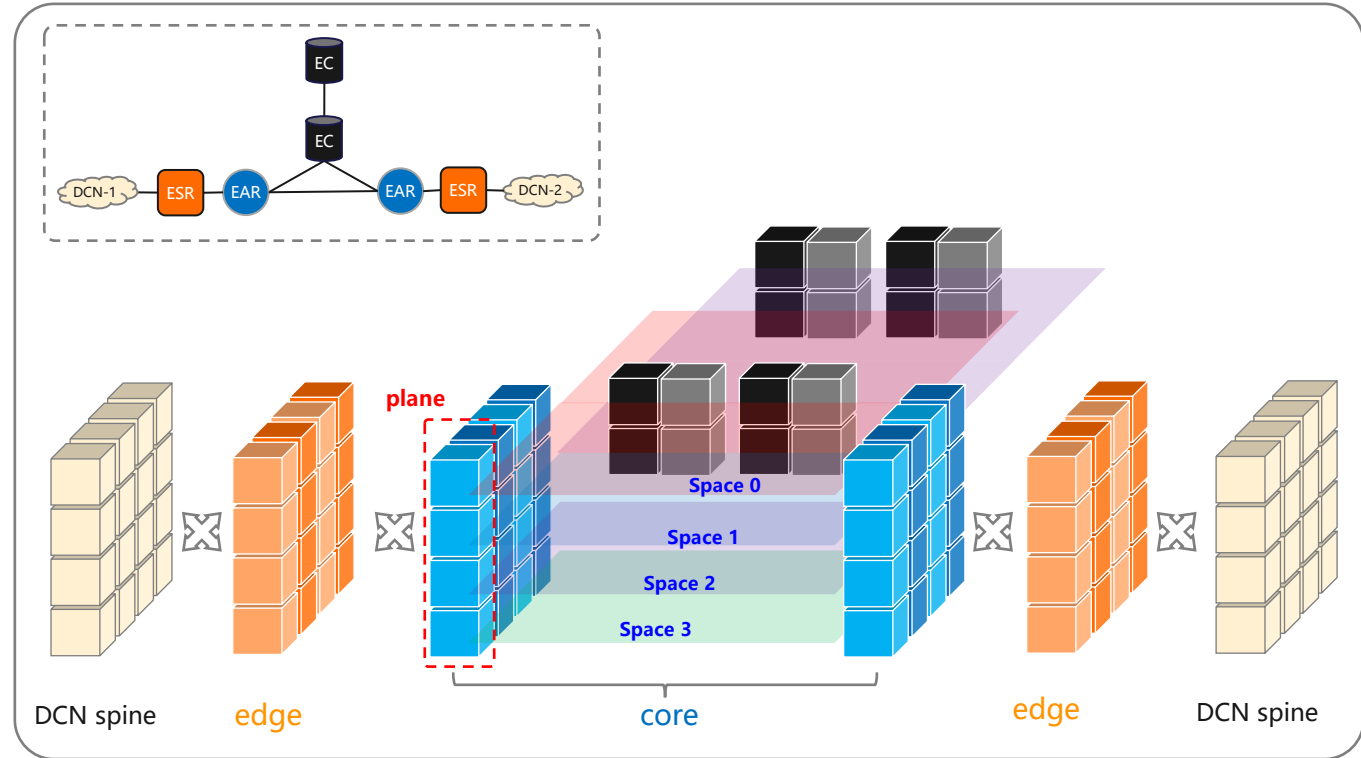
Composable Architecture

Standard Network Modules

- Node/Cluster/Group

Elastics and Resilience

- Multi-Plane (redundance)
- Multi-Space (capacity scaling)
- Multi-Routing Domain (failure isolation, small blast radius)



Minimize Blast Radius

IS-IS minimized

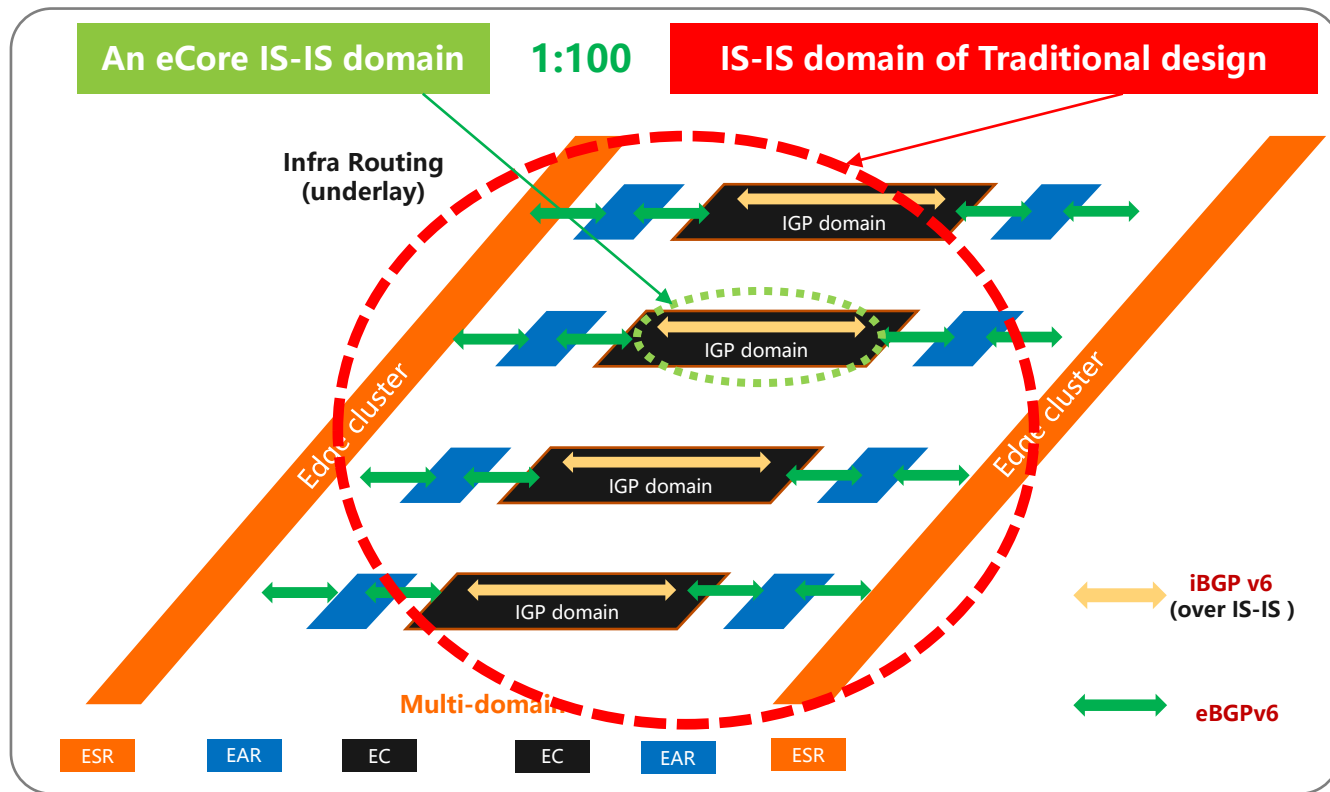
- ISIS domain = EC plane
- iBGP next-hop resolution only

BGP Underlay Routing

- SRv6 Inter-domain friendly
- **NO MPLS LSP**

Seamless Migration from Legacy MPLS Core

- SRv6 VPN over 6PE MPLS Core
- SR-MPLS TE



Service-Oriented: A Simple Way to Provide More Flexible TE Service

SRv6 VPN Overlay

- Service-SID (End.DT46)
- TE-Class + VPN

Very large SID space to represent flexible and granular TE services

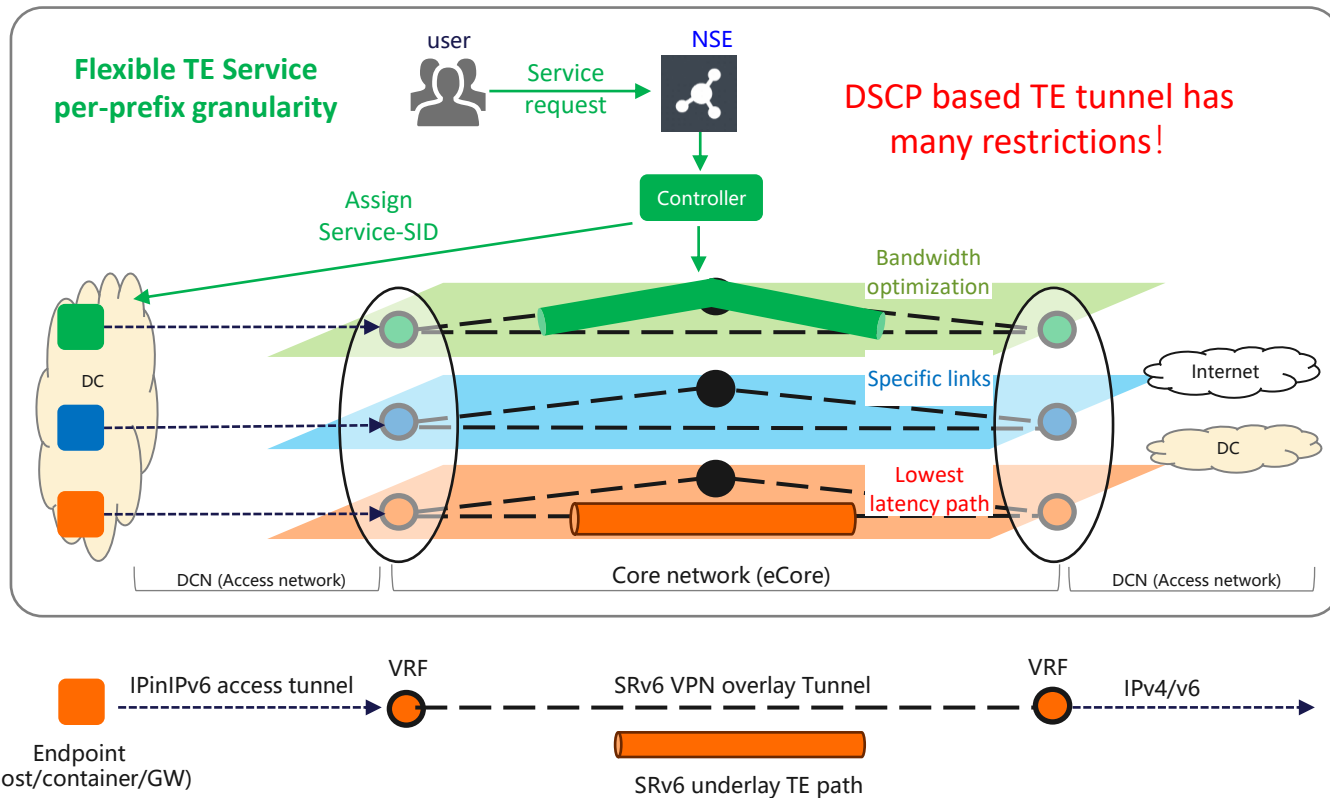
SRv6 TE Underlay

- TE SID (uN/uA)
- SRv6 policy + controller

SRv6 Tunnel Access

- IPinIPv6 terminated in VRF (End.DT46)

A simple way to steer traffic into SRv6 TE tunnel



Engineering Challenges and Solutions

Design

Pizza Box

BGP Distance Vector Routing

TE-Class Based VPN

Challenges

Node Explosion

Underlay Routing
Slow Convergence

Overlay Route
Replication

Solutions

Anycast SID

LAG + ECMP
Prefix Independent
Convergence(PIC)

SID Marking

Implementation

SONiC Phoenix Wing initiative
SRv6 VPN, SRv6 Policy, FIB loading time optimization, PIC, etc



Build and Operate DCI with DCN Paradigms

Deployment

Fully automated, like DCN

Service Monitoring

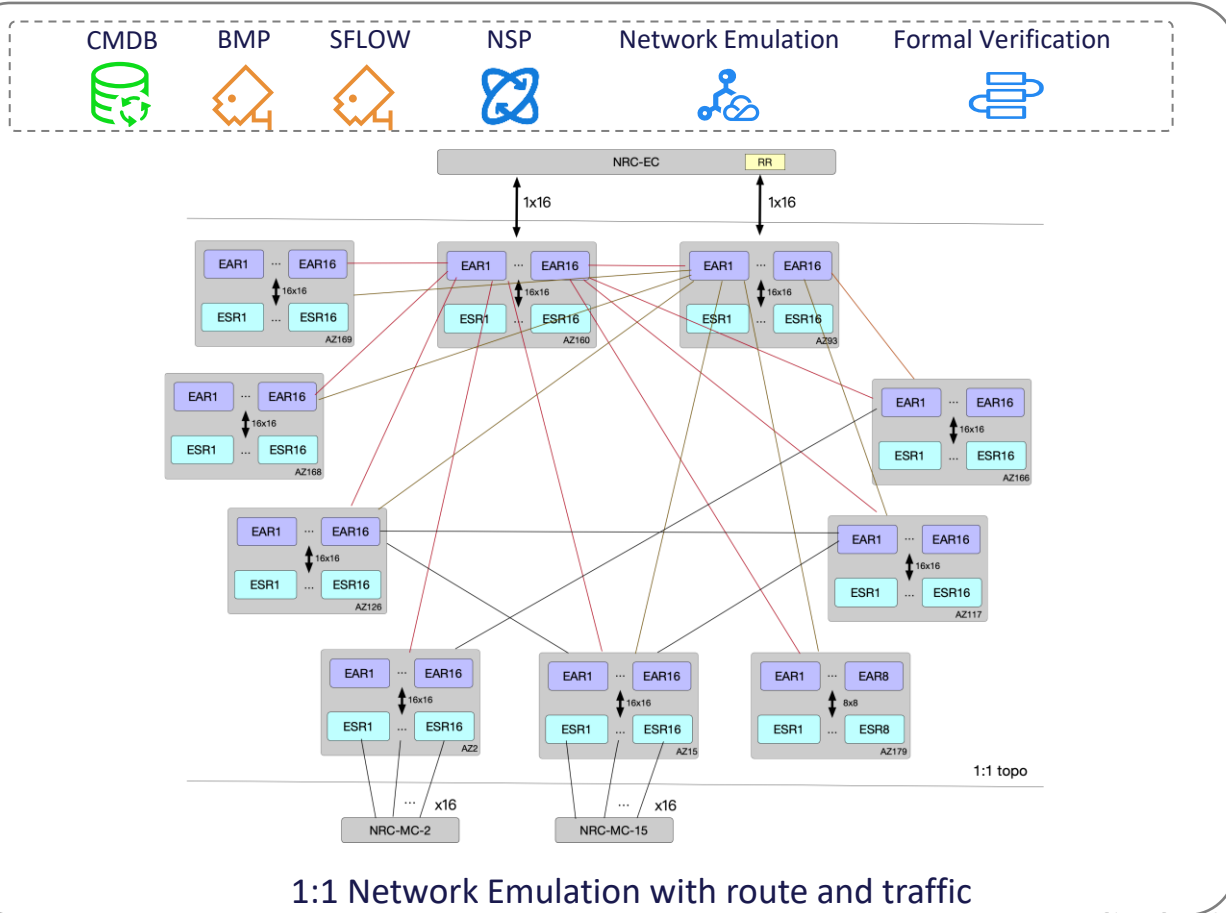
Per-service measurement

Software Upgrade

Live patch, fast upgrade

Verification

1:1 emulation with vSonic & formal verification for routing



Deployment Experience

**Automatic tools speed-up
massive rolling out**
(new region from months to
weeks)

Deployment: from big router to DCN-like whitebox system

- eCore deployment in first region – ~100 devices in 2 months with script tool
- Latest region deployment – 300 devices in 2 weeks with fully automated systems (modeling, auto-configuration, emulation, etc)

**Service monitoring is key to
detecting/locating failure**
(0 P1/P2 outage)

A silent packet drop case caused by a hardware defect

- Used to take a few hours to detect and locate
- Now detected in 1 min with service monitoring tool
- And located in 1 min with SFLOW based analysis

Live patch saves the day
(no more ISSU)

Network-wide bug fix

- Live patch on thousands of devices in just 1-2 weeks
- Traditional way of image upgrade needs a few months

eCore – the Results

Pain Points of AGN

- Device/protocol complexity
- Big failure domain and slow convergence
- High CAPEX and OPEX
- Supply-chain risk and feature velocity
- Rigid traffic engineering

eCore's Advantages

Pizza Box Single chip & single stack (SRv6)

10 x Fast convergence **1/100** ISIS domain size

50% CAPEX reduction **Operation Excellence** Fully automated

10 x Feature velocity (within weeks)

per-prefix Granularity TE

Acknowledgements

Special thanks to the [Cisco Segment Routing team](#) and [Silicon One team](#) for their outstanding support and collaboration in open communities—especially during the early stages of SRv6 Micro-SID and its implementation on our white-box platforms.



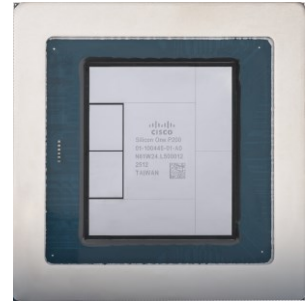
RFC 9800 - Compressed SRv6 Segment List Encoding
(June 2025)

<https://www.segment-routing.net>

Quantum Cat - 51.2T Alibaba | Cisco P200

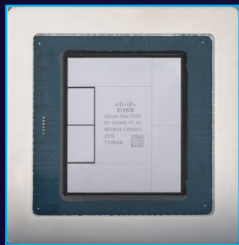


- Alibaba In-house developed 51.2T routing system
- 2xSKUs:
 - 128 * 400GE
 - 64 * 800GE
- Planned production by Q1CY26
- **Powered by Cisco Silicon One P200**



Silicon One P200 – Software Spotlights

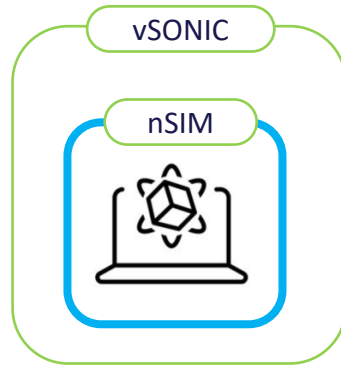
Digital Twin



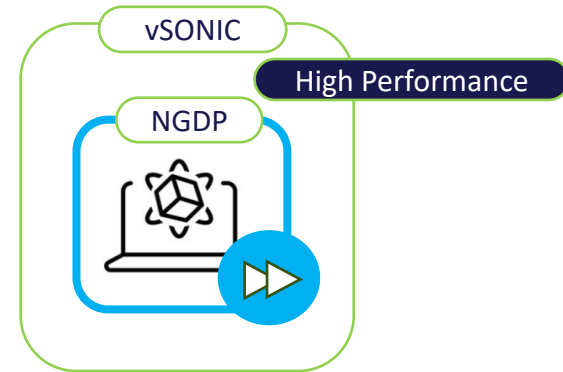
Programmability



P200 Digital Twin



Developer Friendly



Workflow Friendly

Integrate software months before Silicon arrives

Enable Network Emulation

Enable Faster Software Development, CI-CD

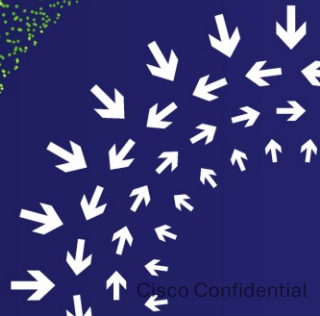
Silicon One P200 - Software Spotlights

Programmability / S-BFD Offload



OCP
GLOBAL
SUMMIT

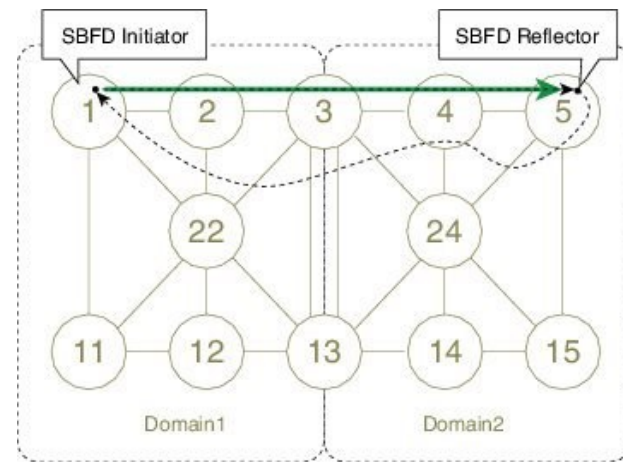
October 13-16, 2025
San Jose, California



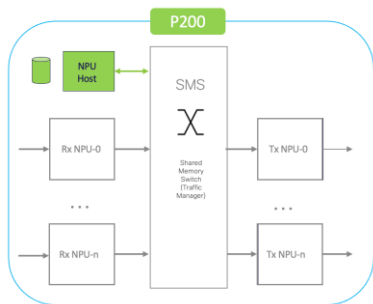
© Cisco Confidential

Seamless BFD (S-BFD) – RFC 7880

Feature	Initiator	Reflector
State Machine	Stateful - Maintains BFD state machine and detection mechanism	Stateless
Packet Handling	Generates and sends S-BFD control packets. SRv6 encap to monitor TE paths	Receive and reflect S-BFD control packets
Role	Monitor forwarding path and trigger a response	Swap SIP/DIP, discriminators



Programmability : Seamless BFD (S-BFD) Offload with P200



Session Scanner:
Periodic initiation



Periodic Send:
P4 generates S-BFD payload + headers
Routed by normal pipeline



Receive + Session Update:
P4 updates stateful session entry



Timeout Scanner:
Notify CPU if S-BFD session is detection "down"

Reflect:

- Validate own discriminator
- Swap header fields
- Send back to Initiator

- Thousands of stateful aggressive sessions
- SRv6 encap allows monitoring all paths

Unlimited Reflector session scale
(millions of PPS)

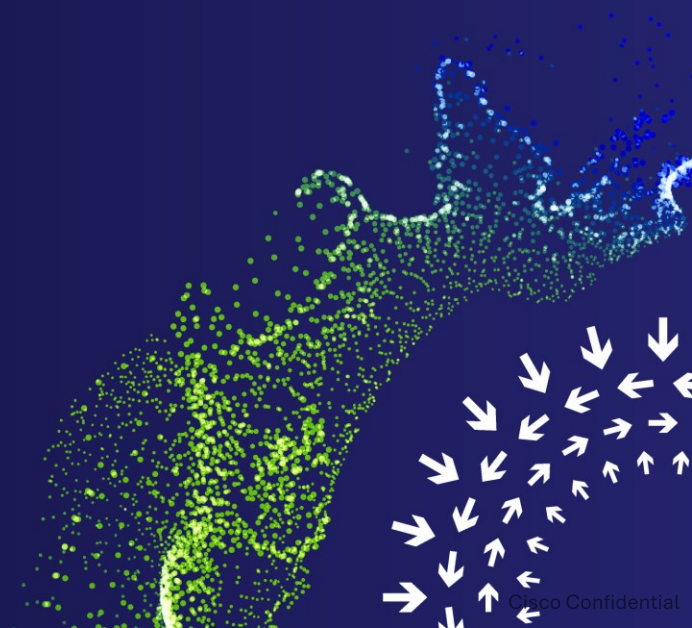
Silicon One P200 - Software Spotlights

SRv6 Update



OCP
GLOBAL
SUMMIT

October 13-16, 2025
San Jose, California



Enhancing SRv6 in SAI – Great Partnership with Alibaba

Category	Contributions	Status
Service Overlay and VPN	SRv6 VPN SID	Contributed
Traffic Engineering and QOS Policy	Class-based forwarding	Contributed
OAM and Fast Failure Detection	SRv6 Encap for BFD HW Offload	Contributed
	Seamless BFD (S-BFD) HW Offload - https://github.com/opencomputeproject/SAI/pull/2220	PR Opened
Performance Telemetry and Path Visibility	Path Tracing	Contributed
	SID Marking - https://github.com/opencomputeproject/SAI/pull/2222	PR Opened
	IPM - Integrated Performance Measurement	Upcoming

Silicon One P200 – Silicon, Systems, Software – Available Now!

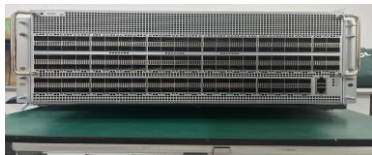
ALiNOS



IOS-XR, NX-OS

Common SAI

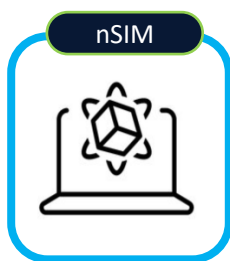
Common P4 and SDK



Alibaba Quantum Cat



Cisco 8223



- 512x112G Serdes
- **16GB** HBM deep-buffering
- **Software-selectable** Standalone or LC/DSF modes
- **10M+** IPv4/IPv6 routes
- **64K** ACL entries per forwarding engine (x16)
 - **1M+** ACL with prefix compression
- **Fully Programmable** Run-to-Complete NPU
- Rich DCI & Routing feature set
- **Integrated MACSec and TunnelSec**
- **SAI / SONiC** on Day-1

Call to Action

- SONiC Routing Working Group – [Phoenix Wing](https://lists.sonicfoundation.dev/g/sonic-wg-routing) enabling SRv6 in community SONiC:
<https://lists.sonicfoundation.dev/g/sonic-wg-routing>
- eCore's design & implementation will be open/opensourced
(Stay tuned for the upcoming release of the [eCore Deployment Report/Paper](#))
hopefully could help the industry to resolve the WAN issues existing for decades

Thank You!