

Segment Routing v6

Deterministic AI Backend

Networking and Unified DC

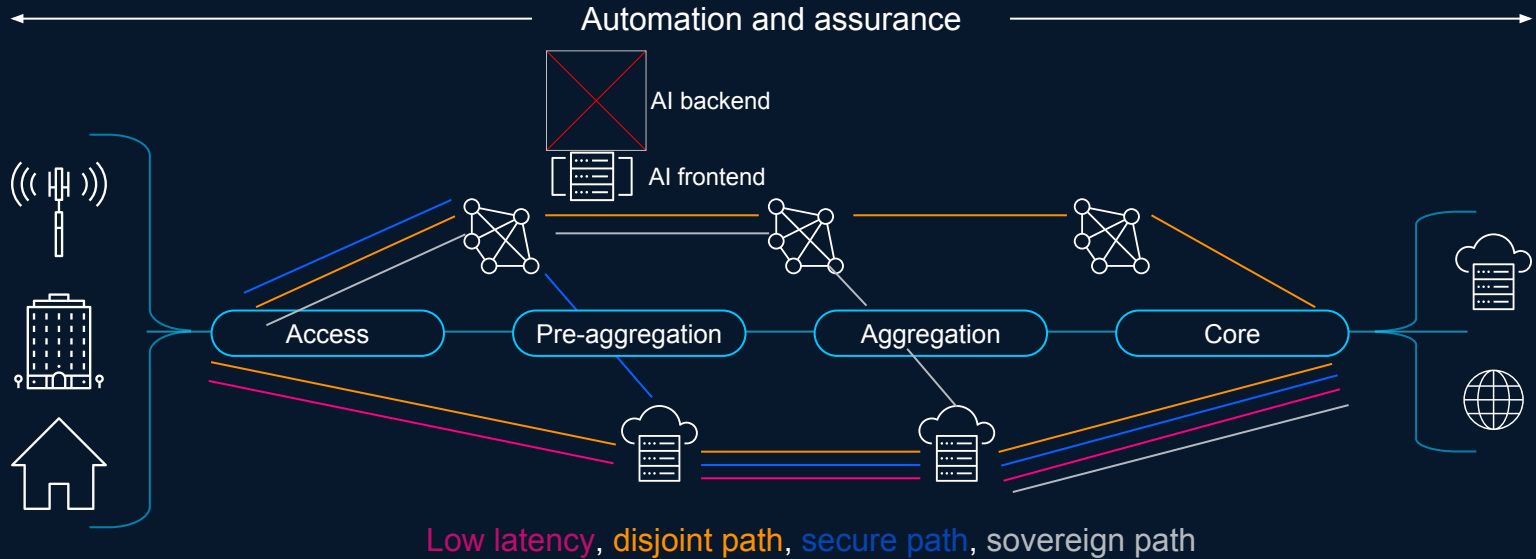
Frontend Design



Alberto Donzelli
Principal Solution Engineer, Cisco

Segment Routing IPv6 (SRv6)

A multi-domain programmable fabric for service creation

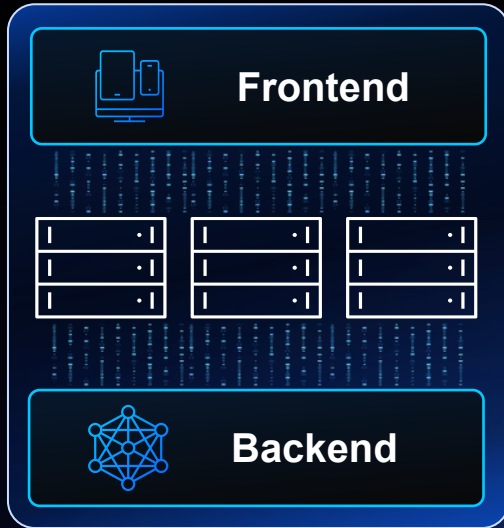


SRv6-Powered Data Center Interconnect

DCI growth for inference and training is massive



System design choices are changing over time



Data Center 1

Modular system for 1.6T
Dis-aggregate with 51.2T deep-buffered switches

High capacity ZRP and ZR+ optics

SRv6 for AI DCI roles

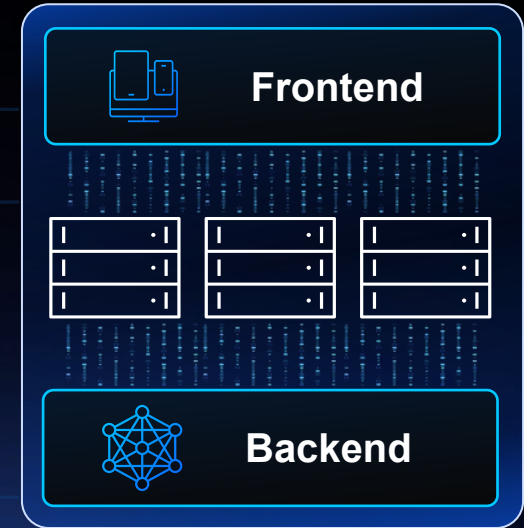
Utilizes Micro-SID (uSID) to reduce header overhead,
improving performance for AI and
high-throughput connectivity

Enhanced Traffic Engineering

Improved scalability

Reference Architecture

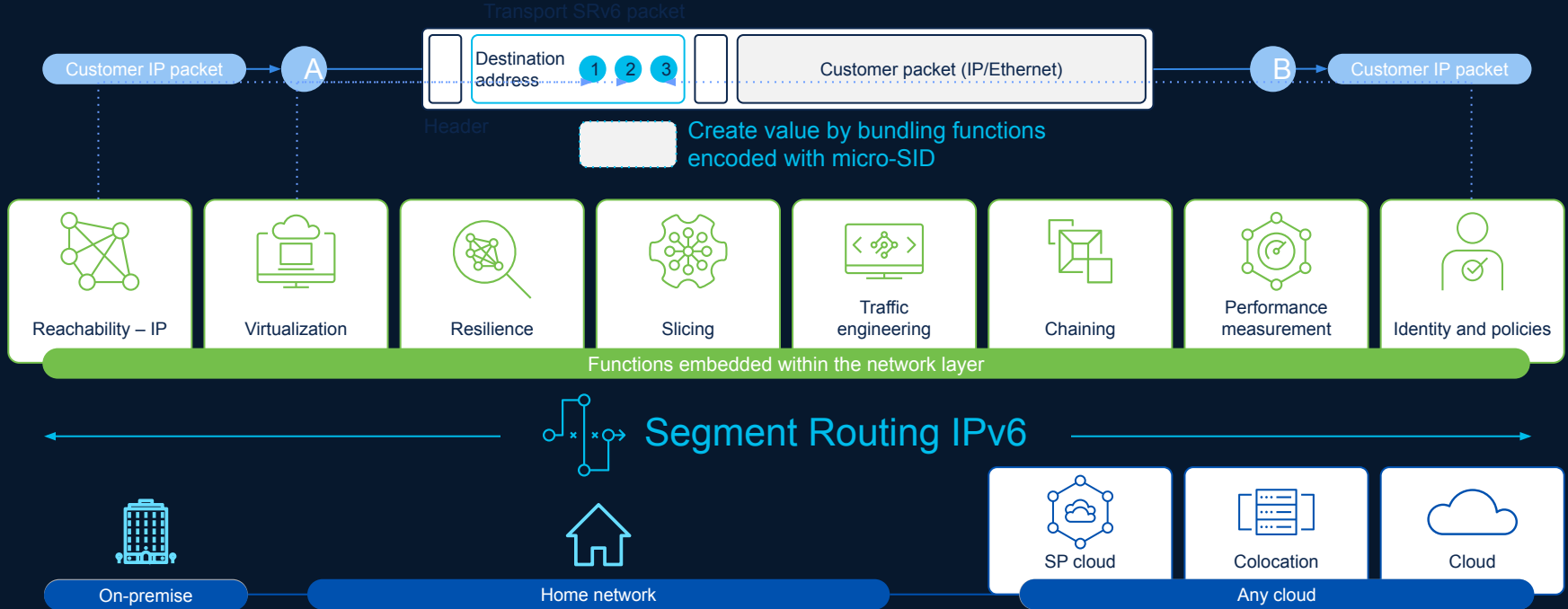
Expanding Cisco CRA to cover DC connectivity



Data Center 2

SR IPv6 - a service programming platform

Deploy diverse functions where needed



Fully Standardized with strong Cisco commitment and leadership
Editor of 96% IETF RFCs | Co-author of 100% IETF RFCs



SRv6... at record-speed



America
S



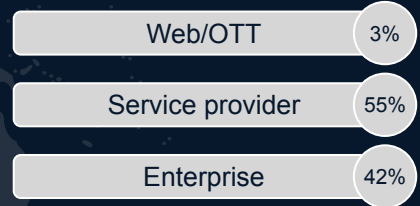
EME
A



APJ
C



- Deployed
- Active testing
- Deployment planned



SRv6 uSID - Rich Ecosystem

Open-Source Networking Stacks

Network Equipment Manufacturers



Merchant Silicon



Access: A100 Edge: K100
Core: Q200, P100
ToR/Spine: G100, G200

DNX – Jericho
XGS – Tomahawk



Smart NIC / DPU



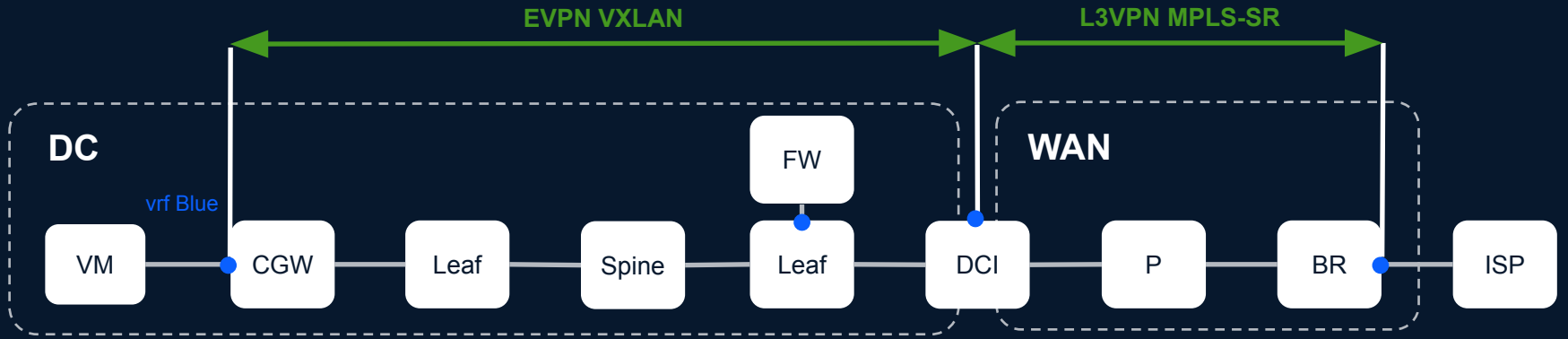
Open-Source Applications



Partners



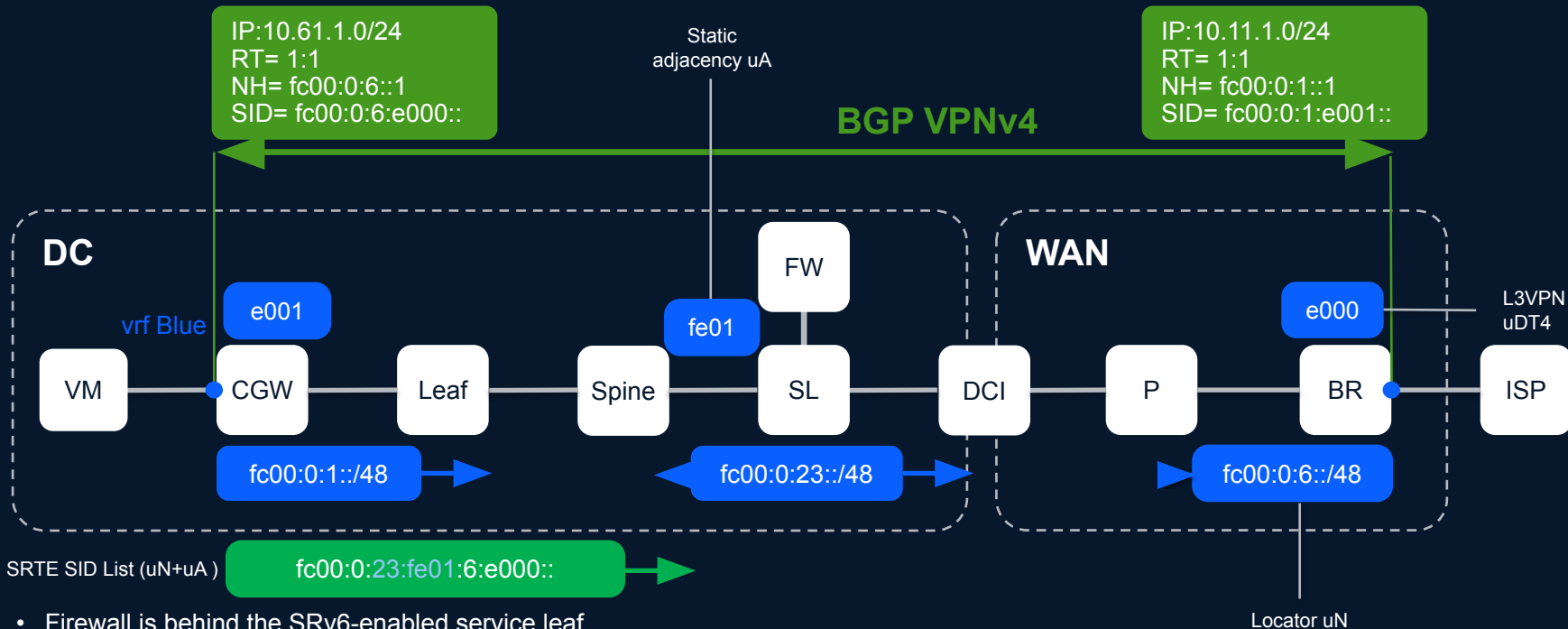
Current architecture of the DC frontend network



- IPv6 only infrastructure in the data center.
- Multivendor data center and WAN networks approach.
- Cloud gateway and firewalls are NFVs running on hosts. Nebius develops them.
- VXLAN based overlay between the cloud gateway and DCI.
- DCI does “stitching” between EVPN VXLAN and L3VPN MPLS-SR.



Services chaining with SRv6 uSID (e.g., firewall)



- Firewall is behind the SRv6-enabled service leaf.
- SRv6 TE-policy at CGW and BR to set the explicit path through the firewall (SL uN+uA).
- Firewall is a plain IPv6 forwarder, inspects the inner packet, does not change the outer IPv6 header.
- Cloud Gateway CGW is VPP with GoBGP control plane.



SRv6 — Powering the World's Most Powerful AI Data Center

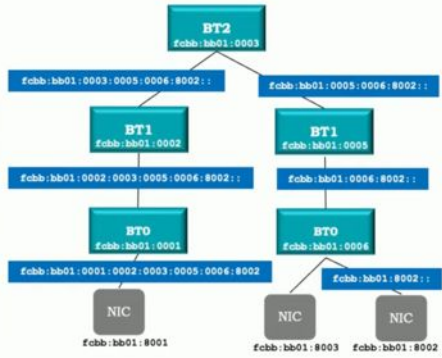



Sept 2025: Microsoft announced Fairwater DC in Wisconsin. The world's Most Powerful AI Data Center.

House hundreds of thousands of GPUs, operating in seamless clusters.

Microsoft Blog [\[Link\]](#)

SRv6



Why SRv6 in AI Backend?

- Training traffic is very bursty.
- Traditional ECMP doesn't work
 - Few elephant flows with little entropy
 - If flows collide on link, tail latency increases, JCT increases

How SRv6 works in the AI Backend?

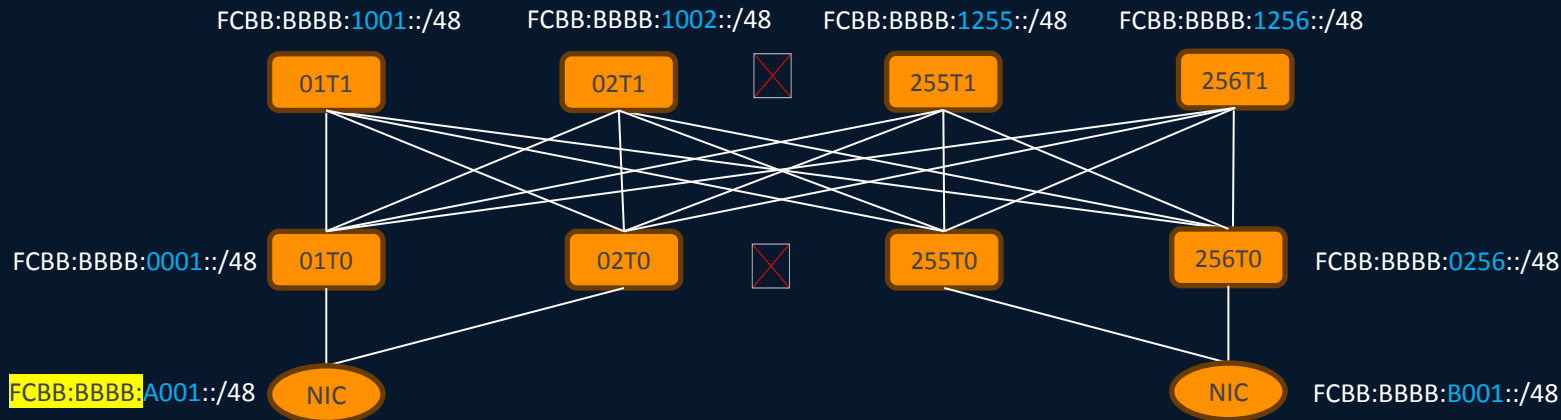
- The uSID stack is communicated to the NIC endpoints as well as to the switches
- The source NIC populated the uSID stack for the complete network path
 - The packet can contain upto 6 uSIDs (without SRH)
 - SRH header is optional
- The switches are SRv6 transit devices processing the uSID stack
 - Drop packet if uSID stack doesn't match
 - Strictly source routing only
 - Target NIC decapsulates the packet and processes it
- Both uN and uA (for multi-path) can be supported

Oct 2025: Microsoft explained the key enabling technology for the Fairwater DC **SRv6!!**

SRv6 used for deterministic placement of AI flows in the fabric **achieve optimal load-balancing**

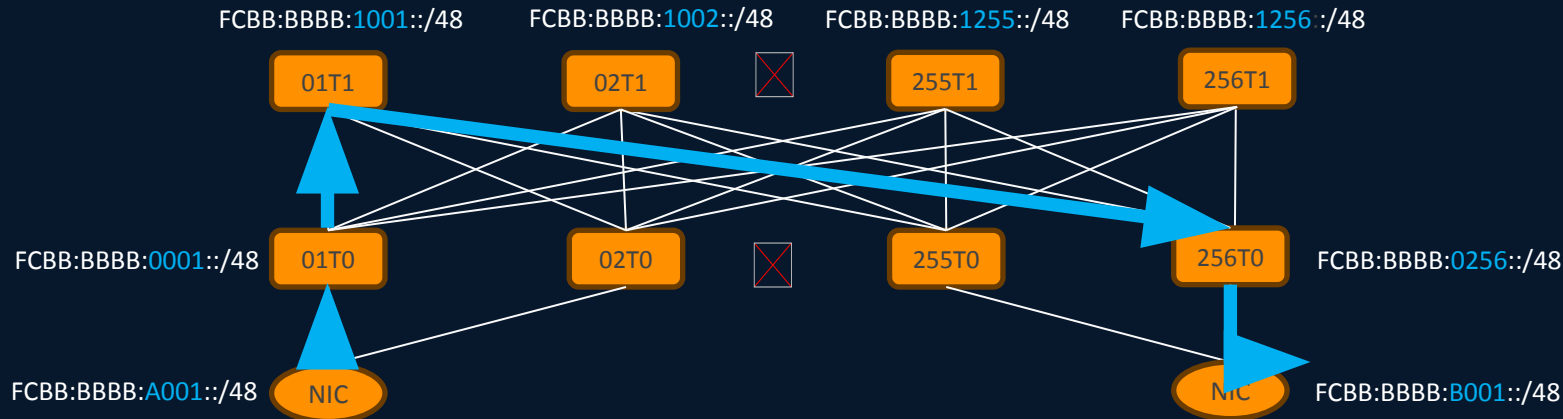
Guohan Lu, Microsoft. OCP Global Summit [\[Link\]](#)

SRv6 uSID – Deterministic Path Placement



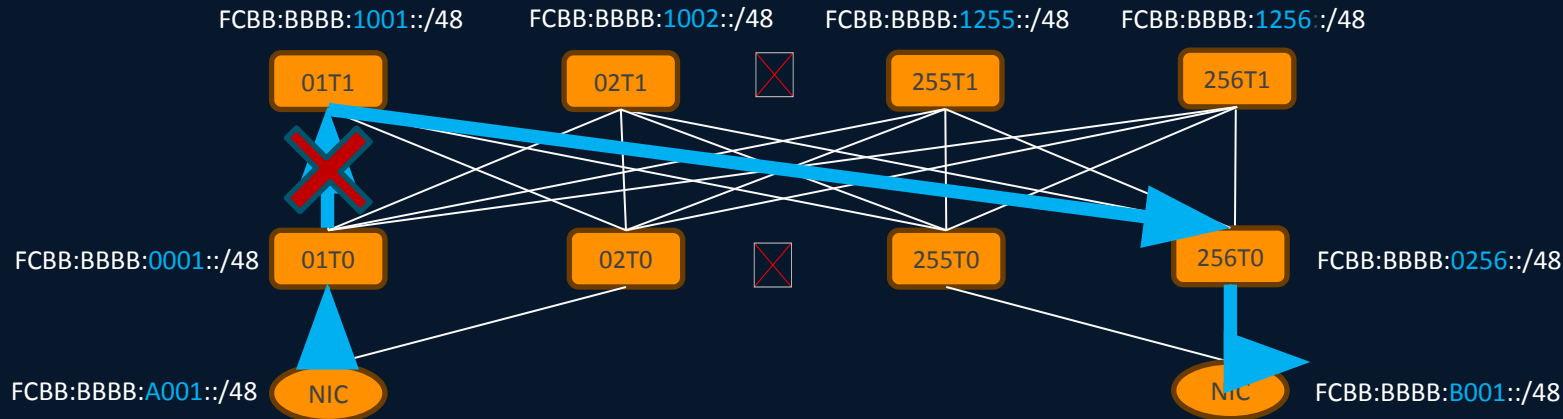
- SRv6 Block - FCBB:BBBB::/32
- SRv6 Locator - FCBB:BBBB:NNNN::/48
 - 01T0 => FCBB:BBBB 0001::/48

SRv6 uSID – Deterministic Path Placement



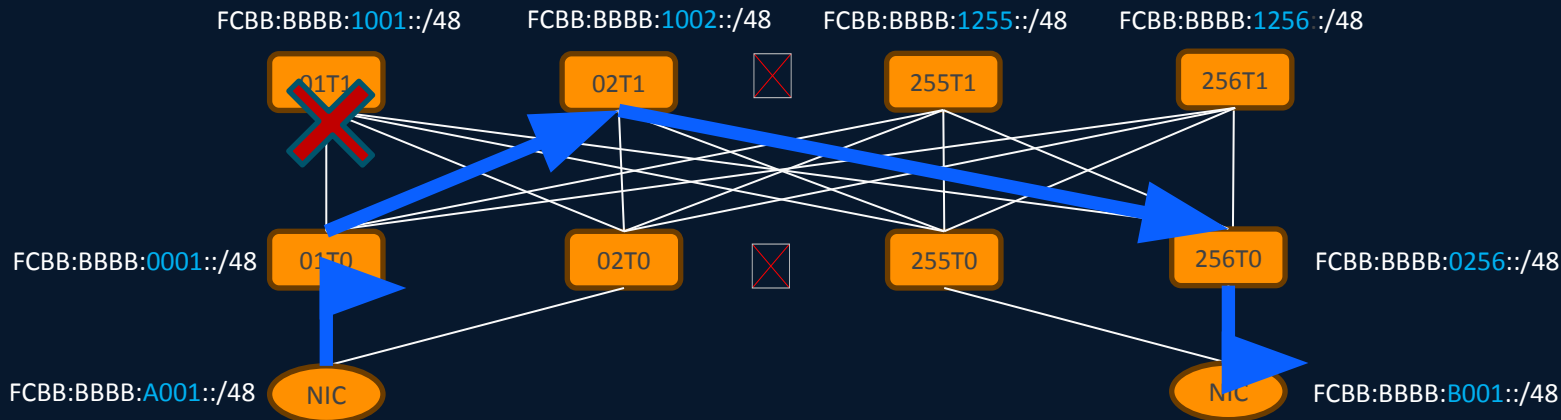
- NIC A to NIC B
 - Via FCBB:BBBB 0001 1001 0256 B001
 - Go to 01T0 then 01T1 then 256T0 then NIC B

SRv6 uSID – Deterministic Path Placement



When congestion happens, congestion notification is sent (e.g packet trimming)
The source retransmit, changing path

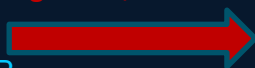
SRv6 uSID – Deterministic Path Placement



- NIC A to NIC B

- via FCBB:BBBB:0001 1001 0256 B001
- Go to 01T0 then 01T1 then 256T0 then NIC B

Congestion/Failure "01T1"



- NIC A to NIC B

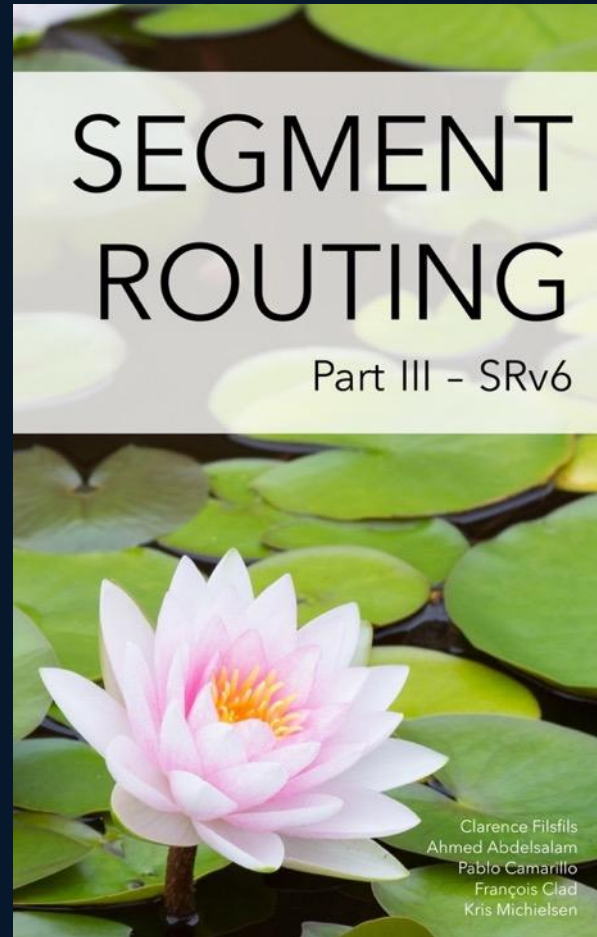
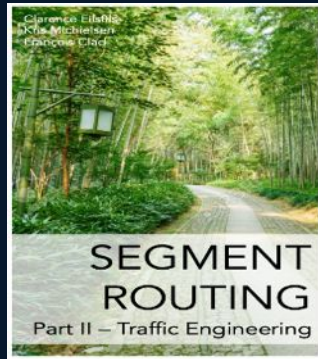
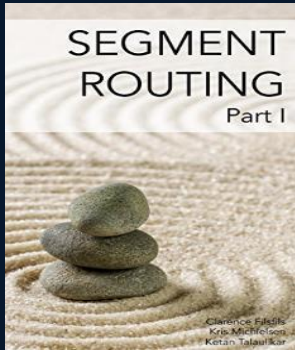
- via FCBB:BBBB 0001 1002 0256 B001::
- Go to 01T0 then 01T1 then 256T0 then NIC B



Conclusions

- SRv6 makes IP more scalable, richer with backward compatibility
- SRv6 provides a Unified IP Solution
 - SP Architecture from access to Internet through DCI and WAN
 - AI Fabric: Backend, Front End, DCI and WAN
- Cisco SR team on top of innovation, standardization and deployments
- A lot of new SRv6 use-cases are coming!





segment-routing.net



