SRv6 NEXT-C-SID – “IPv6 uSID”

Clarence Filsfils

Flexibility

- uSID Block can be Private or Public
- uSID Block can be independent from the IP address block
- Multiple uSID Blocks can be allocated: one per Flex-Algo
- uSID Block length can be 16, 32 or 48-bit
- uSID length can be 8, 16, 32… in fact any length
  - And within a program uSID of different lengths can be combined

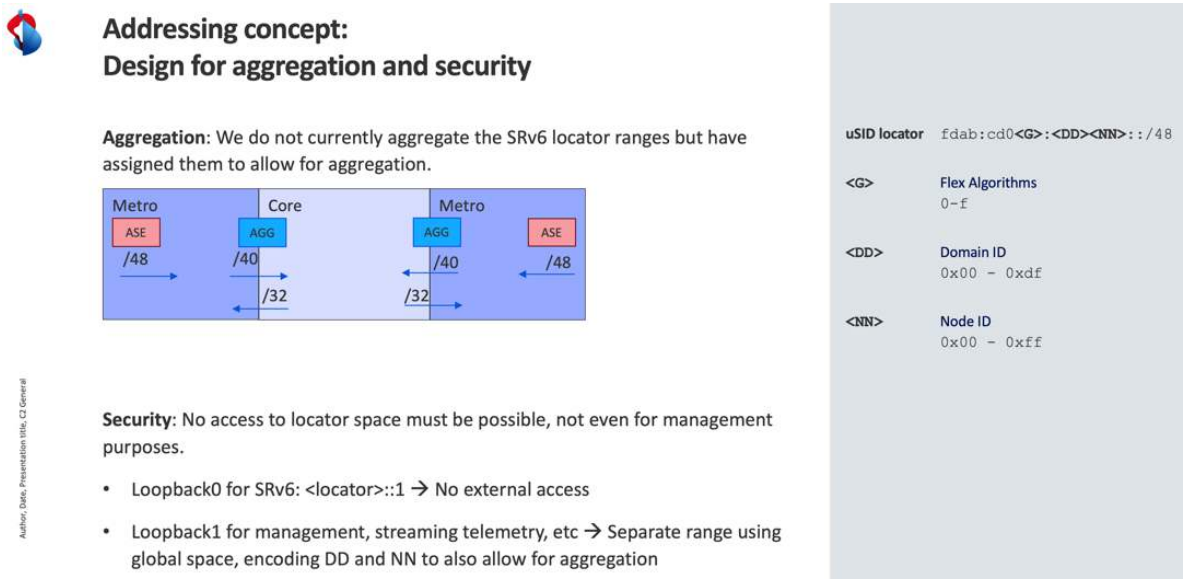
A Typical Deployment

“IPv6 uSID infra is configured in Rakuten network on 14k+ devices, 70% services running on uSID successfully”

Akash Agrawal, Rakuten

Swisscom: Introducing SRv6 to an existing network

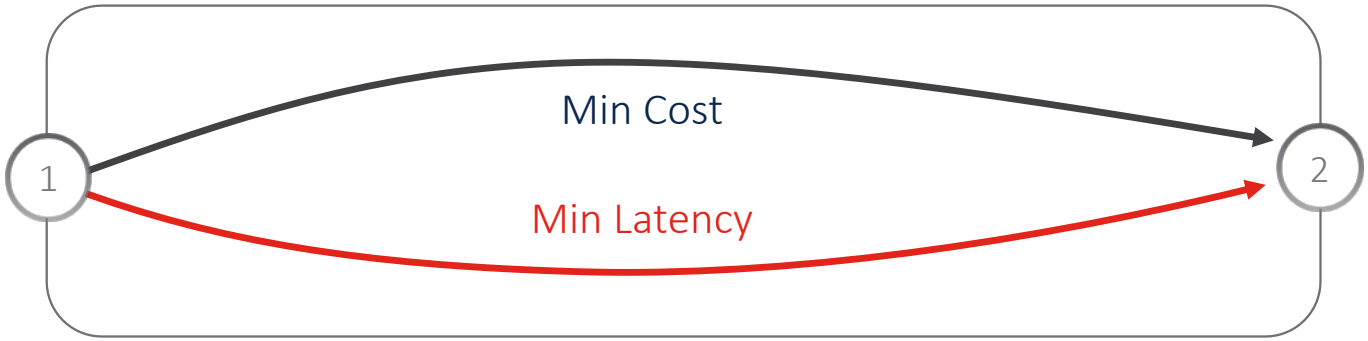
- uSID deployed using unique Local Addresses (ULA)
  - Non routable
- Locator prefix: **fdab:cd0<G>:<DD><NN>::/48**
  - **fdab:cd0/24** block for uSID allocation
  - **<G>**: Flex Algorithms (0-f)
  - **<DD>**: Domain ID (0x00 - 0xdf)
  - **<NN>**: Node ID (0x00 - 0xff)
- Summarization:
  - Leak single /40 prefix to the Core
- Migration approaches:
  - Approach 1: Back-to-back connections
  - Approach 2: Dual-connected PE



https://www.segment-routing.net/conferences/2023-02-10-SRv6-deployment-swisscom/

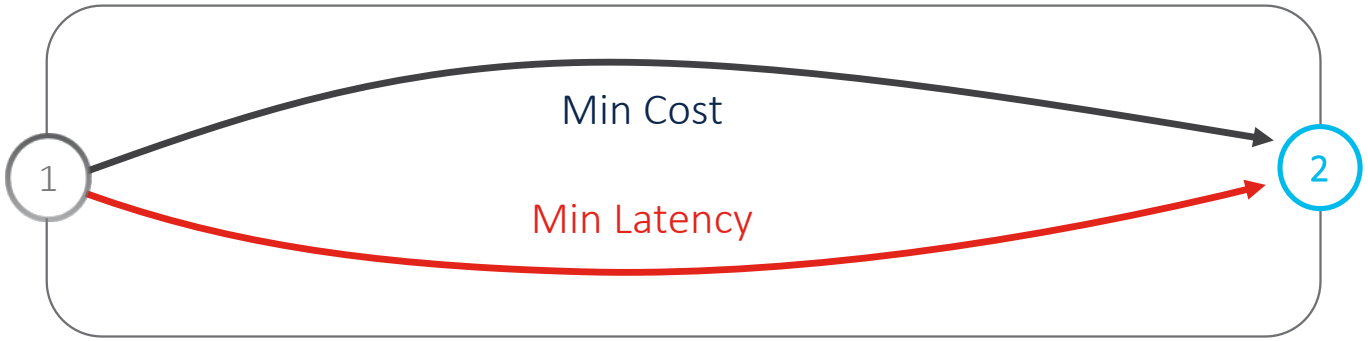
IGP with 2 Algorithms

- IGP cost ⇔ Min Cost
- Flex-Algo 128 ⇔ Min Latency



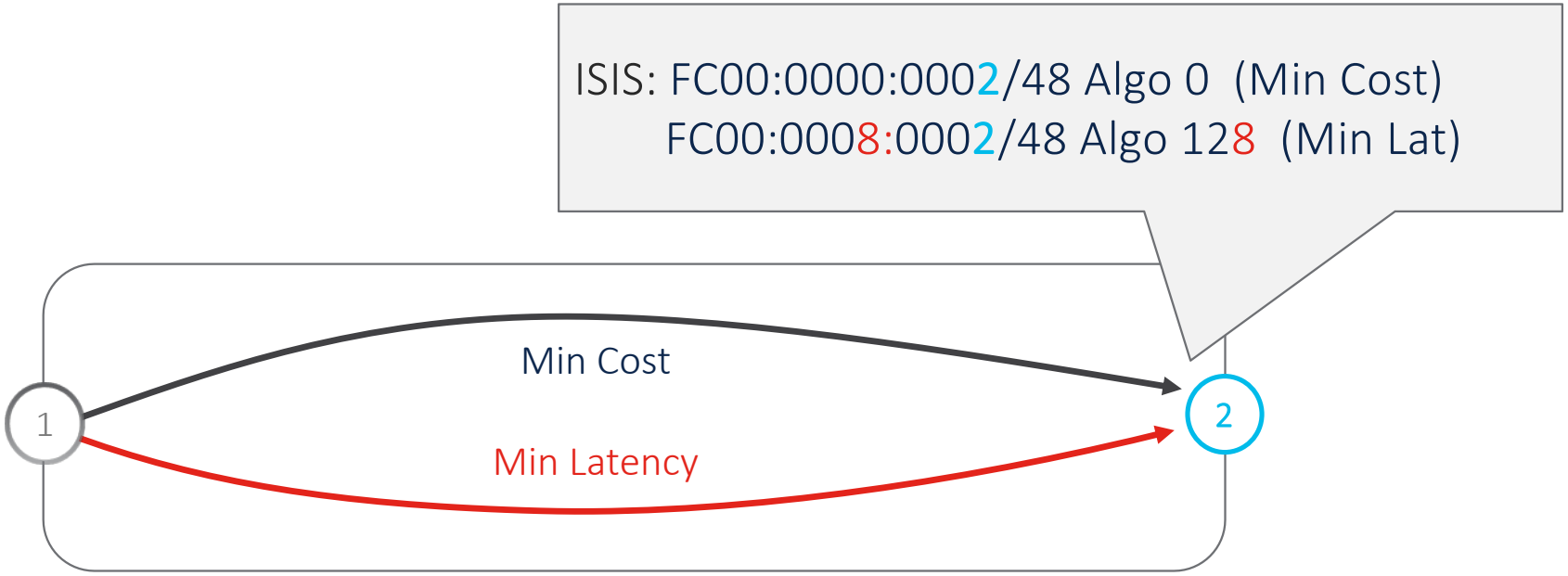
32-bit Private Block, one per Algo

- Min Cost ⇔ FC00:0001::/32
- Min Latency ⇔ FC00:0008::/32



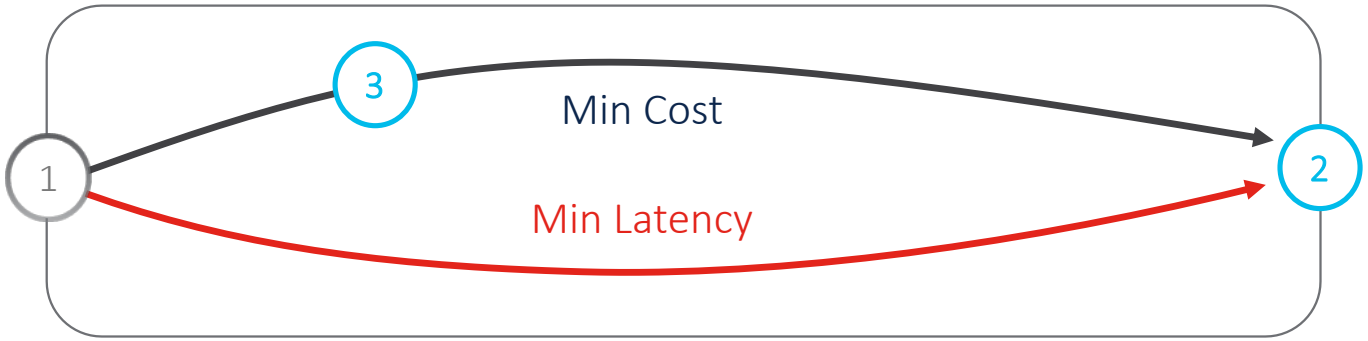
16-bit uSID

- Represented as 4 nibbles
- Globally Significant: if first nibble is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D}
- E.g. Node 2 is 0x0002



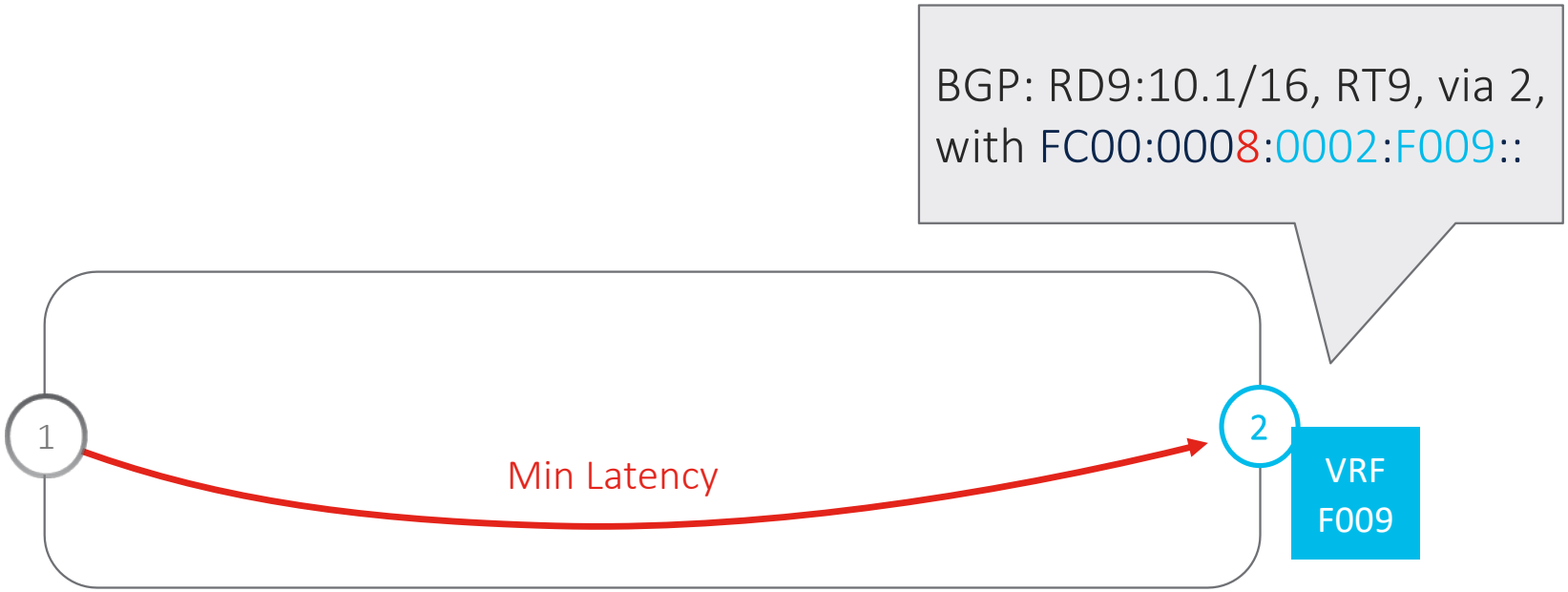
Seamless Deployment

- Node 3 forwards as per legacy longest match FC00:0000:0002/48



16-bit uSID – Locally Significant

- If first nibble is {E, F}
- @ Node 2, 0xF009 is local bound to VRF 9



Hardware Efficiency

- Node 2 processes two uSID's with one single longest match
- FC00:0008:0002:F009::/64 ➔ It is me and this is for VRF 9

Intuitive & Rich

- FC00:0000:0002:F009:0000:0000:0000:0000
  - Min Cost Path to node 2 and then lookup in VRF9
- FC00:0008:0002:F009:0000:0000:0000:0000
  - Min Lat Path to node 2 and then lookup in VRF9
- FC00:0000:0003:0002:F009:0000:0000:0000
  - Min Cost Path to node 3 then 2 and then lookup in VRF9
- FC00:0000:0003:F005:0002:F009:0000:0000
  - Min Cost Path to node 3, VNF 5, min cost path to 2 and then lookup in VRF9
- A program reads left to right and has 6 uSID's in the DA
- uSID 0000 means "end of program"

Ultra Scale

- Global uSID
  - 4 billions while only consuming 0.2% of ULA
  - More is possible
- Local uSID
  - 4 billions leveraging wide 32-bit uSID’s for ultra scale edge services
- Routing Summarization
- Best compression efficiency

Shifting to the Next uSID

- Node 3 receives DA FC00:0000:0003:0002:F001:0000:0000:0000
- Node 3 longest matches FC00:0000:0003/48 => “It is me”
- Node 3 shifts the programs by 1 uSID and the DA becomes FC00:0000:0002:F001:0000:0000:0000:0000
- Node 3 looks up the updated DA and matches FC00:0000:0002/48
- Node 3 forwards the packet to node 2
- Linerate across the whole portfolio

If more than 6 uSID's are required

Outer DA: FC00:0000:0001:0002:0003:0004:0005:0006

uSID1 uSID2 uSID3 uSID4 uSID5 uSID6

SRH: FC00:0000:0007:0008:0009:0010:0011:0012

uSID7 uSID8 uSID9 uSID10 uSID11 uSID12

- 12 uSID's with SRH holding a single SID
- 18 with SRH holding 2 SID's
- SRH rarely needed as 6 uSID's in DA are enough most of the time

Benefits

Grand Architecture with HW-Efficiency

- Revolutionary Network Programming Model (Turing Complete)
  - The IP Destination Address (DA) holds up to 14 instructions
    - 2-byte block, 1-byte uSID's
  - SRH extension header holds additional instructions (rarely needed)
- Any behavior can be bound to the instruction
  - Shortest path according to cost, latency with exclusion of unsecured links
  - TDM-alike behavior (one instruction per hop/interface)
  - TE, FRR, NFV, Cryptography…
- Linerate across our entire portfolio

Novel Architecture with Brownfield

- Classic Longest-Match at Legacy IP node
- The network program is opaque to legacy node
- Alibaba, Swisscom, Bell… are all brownfield deployments

Unified Core Metro Access DC Cloud IP solution
Outperforms per-domain custom shim (MPLS, VxLAN)

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

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

Rich SRv6 uSID Ecosystem

Network Equipment Manufacturers

CISCO, HUAWEI, NOKIA, ciena, JUNIPER NETWORKS, ZTE, ARRCUS NETWORK DIFFERENT, ARISTA, H3C

Merchant Silicon

BROADCOM, MARVELL, BAREFOOT NETWORKS (intel)

SR v6
Segment Routing

Open-Source Networking Stacks

.io, P4, OPEN Compute Project, SONiC, FRRouting

Smart NIC

intel, Mellanox TECHNOLOGIES

Open-Source Applications

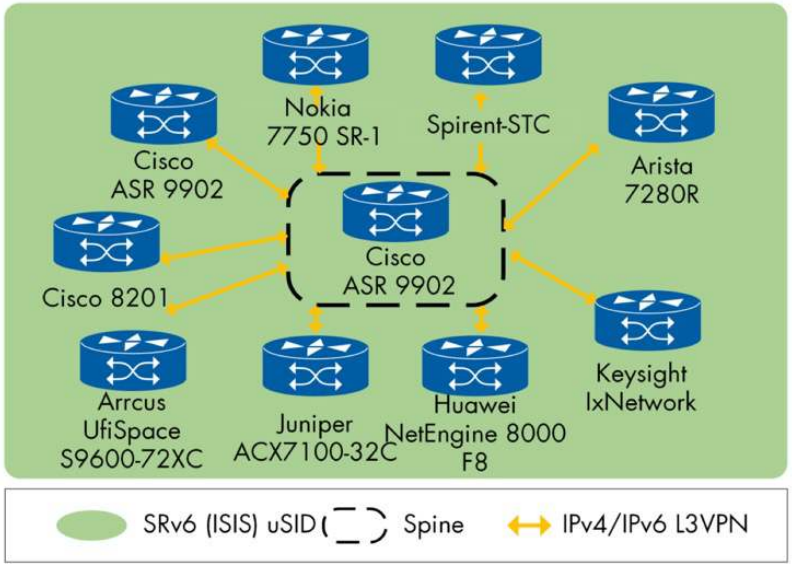
Pyroute2, IPTABLES SERA, GoBGP, exa, BGP, eBPF, cilium, TCPDUMP, WIRESHARK, SNORT, GoBGP

Partners

kaloom, TREND MICRO, ENEA Qosmos Division, KEYSIGHT TECHNOLOGIES, NoviFlow, ixia A Keysight Business, spirent Promise. Assured.

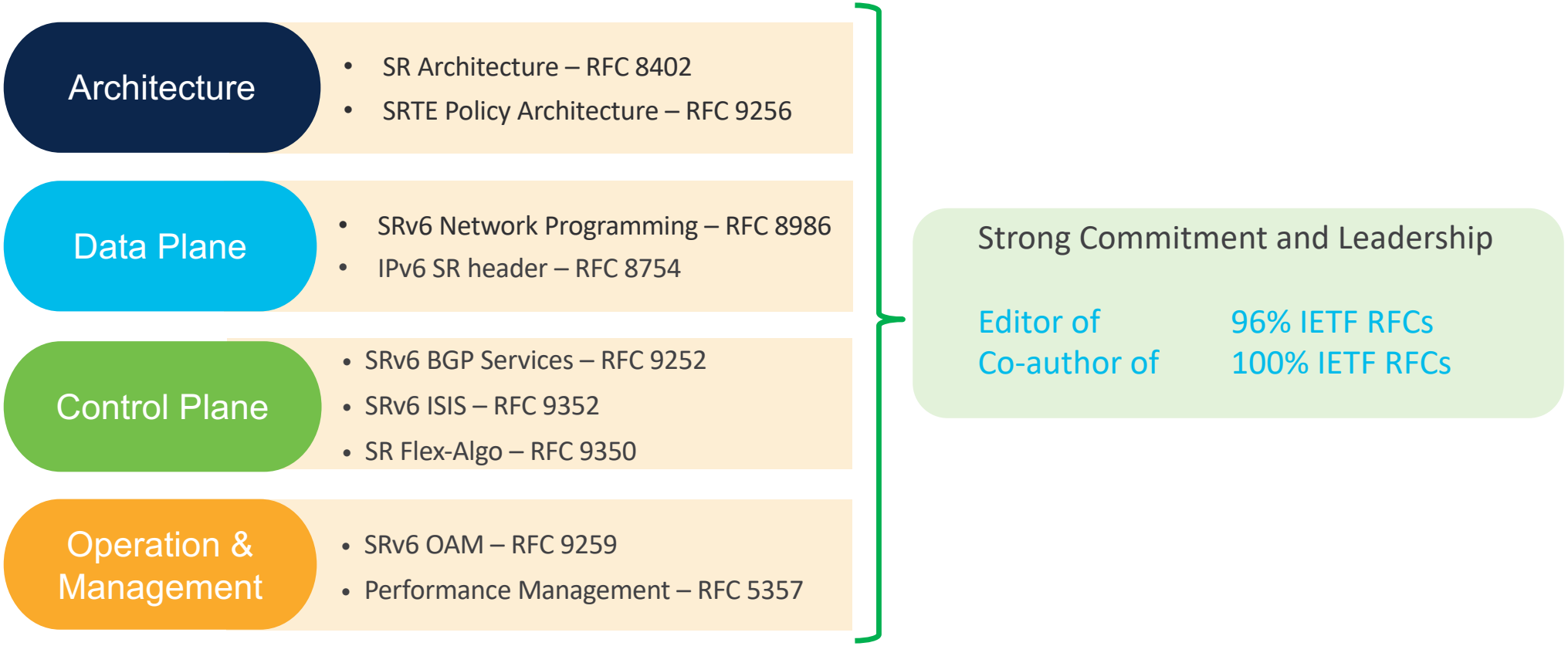
EANTC 2023 testing

- Successful multi-vendor interoperability test for SRv6 **uSID** over 11 implementations.
- Full BGP-based overlay services over SRv6 **uSID** with TI-FLA, UPA, and SR-TE.
- Multiple Silicon families from BRCM, Cisco, Huawei, Juniper, Nokia



https://eantc.de/fileadmin/eantc/downloads/events/2023/EANTC-InteropTest2023-TestReport.pdf

SRv6 is Proposed Standard



Value

Any Service over IP without any shims

- TDM
- Disjointness
- BW
- Latency
- Secured Routing

Unified Solution

- No DPI at VxLAN/MPLS boundaries
- No GW

Native Host and Cloud

- uSID is in IP
- MPLS is neither in the IP socket nor in the cloud

Better Reliability

- 0 Net Outage in 4 years of commercial service & 50k+ deployed routers

Seamless Brownfield Deployment

- Alibaba
- Bell
- Swisscom

Value

Any Service over IP without any shims

- TDM
- Disjointness
- BW
- Latency
- Secured Routing

Unified Solution

- No DPI at VxLAN/MPLS boundaries

Native Host and Cloud

- uSID is in IP
- MPLS is neither in the IP socket nor in the cloud

Better Reliability

- 0 Net Outage in 4 years of commercial service & 50k+ deployed routers

Cheaper

- Lower Overhead
- Smaller instructions (8 or 16 bits)
- Elimination of shim's
- HW Linerate

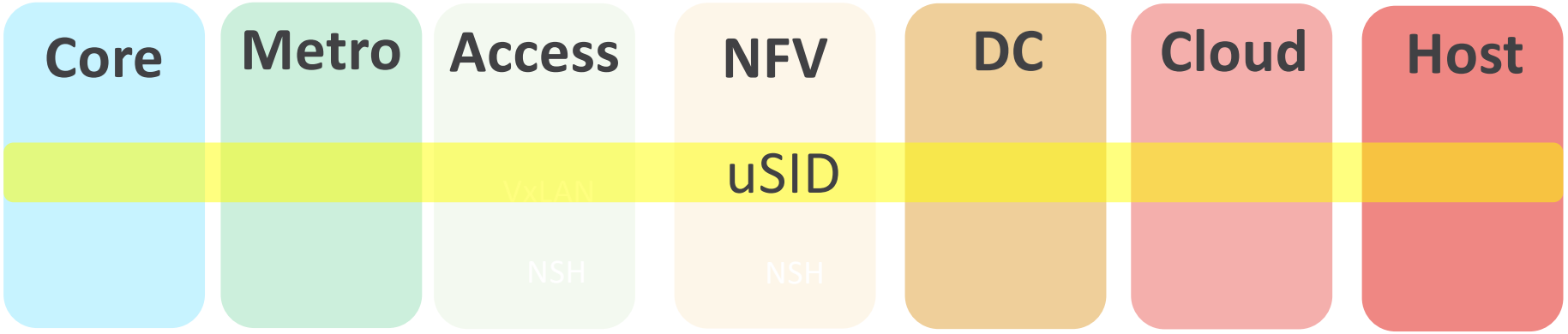
Seamless Brownfield Deployment

- Alibaba
- Bell
- Swisscom

IPv6 uSID

- Most of the nodes just perform classic IPv6 as defined 25 years ago
- Some of the nodes enable and use the uSID network programming by simply using the available space in the outer DA
- SRH is very rarely used but available for ultra-scale use-case

Operator Endorsement across Unified Solution



Dan Voyer
Bell Canada
Paris 2022

Dan Bernier
Bell & NoviFlow
Paris 2022

Gyan Mishra
Verizon
Paris 2023

Dan Bernier
Bell Canada
Paris 2023

Simplicity Always Prevails