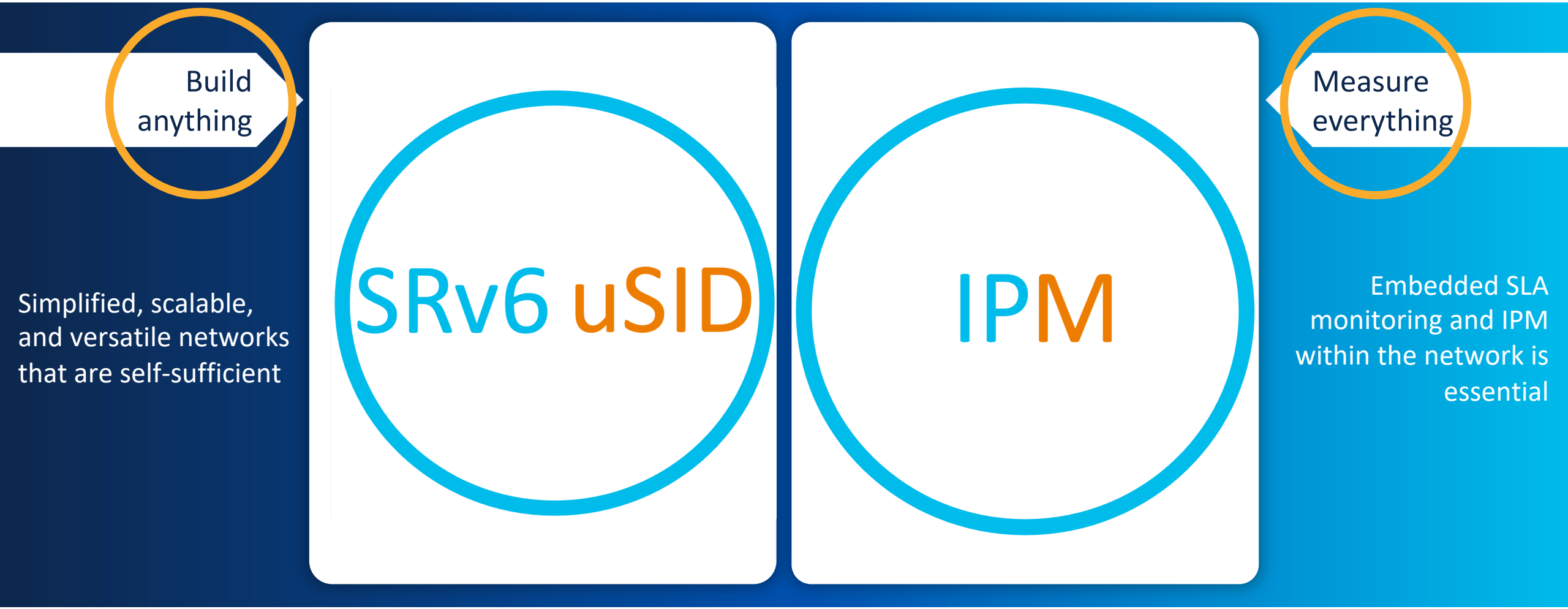
SRv6 uSID Build Anything

IPM Measure Everything

Clarence Filsfils

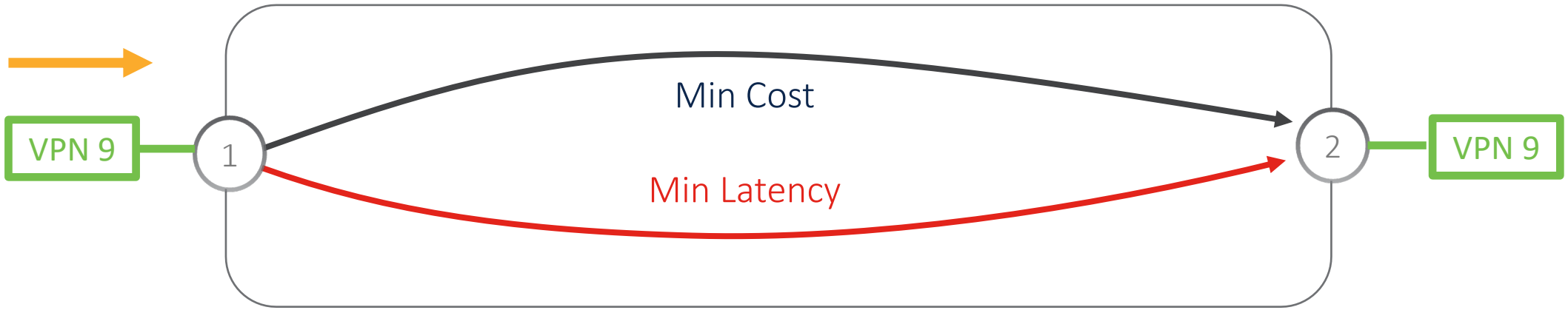
IP is back and better than ever.



A Typical Deployment

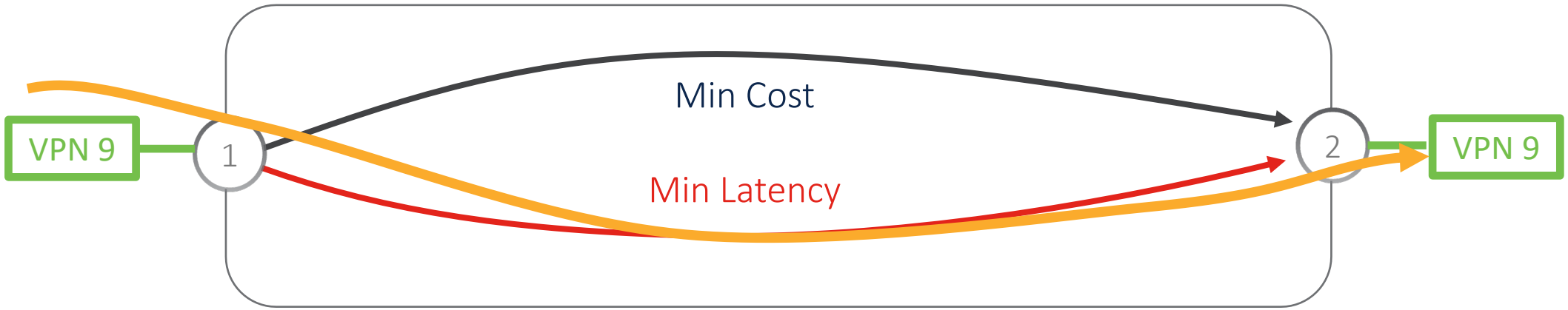
VPN with Underlay TE (Min Latency)

- The VPN “inner” packet gets in the VRF of 1
- Inner packet can be IPv4, IPv6, Ethernet, TDM…



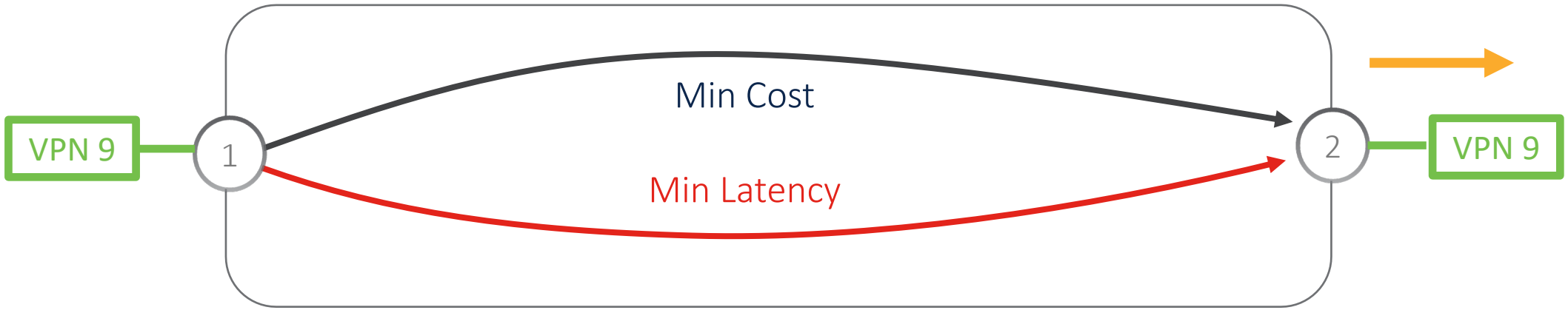
VPN with Underlay TE (Min Latency)

- An outer IPv6 header is added with Destination Address (DA)
- FD00:0008:0002:F009:0000:0000:0000:0000
- The outer DA holds the end-to-end stateless network program



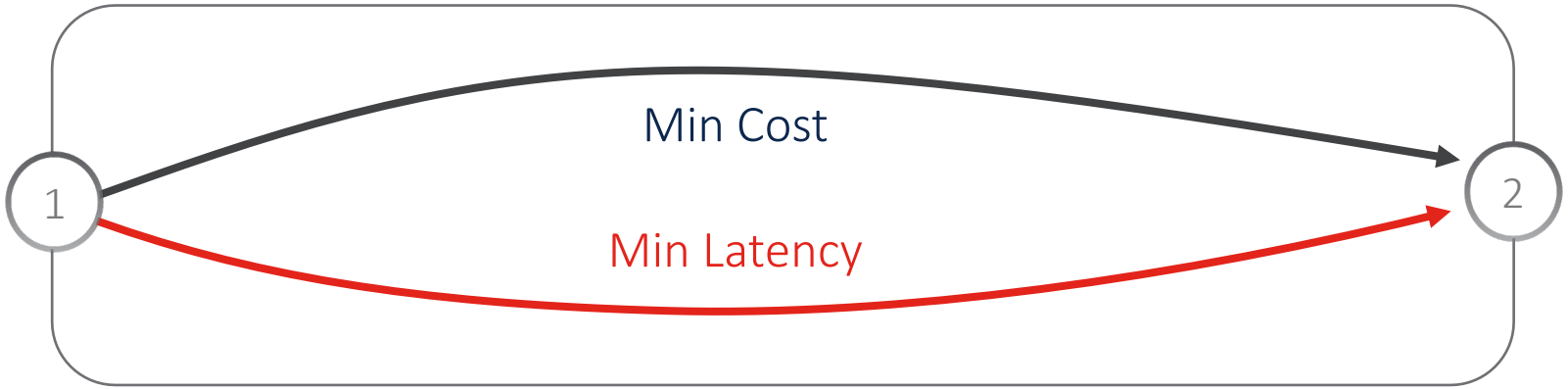
VPN with Underlay TE (Min Latency)

- The outer header is popped
- The inner packet is sent to VPN 9 site
- The desired end-to-end behavior has been delivered without any per-flow state



IGP with 2 Algorithms

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

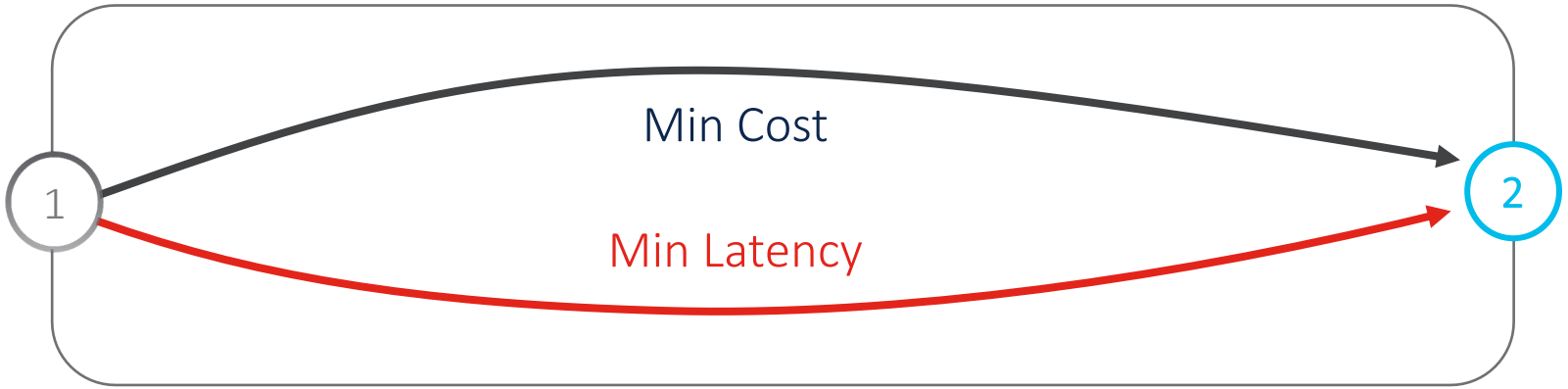


Allocate uSID’s from the Private Space

- All known deployments allocate from Private Address space (FD::/8)
- Public space is also possible

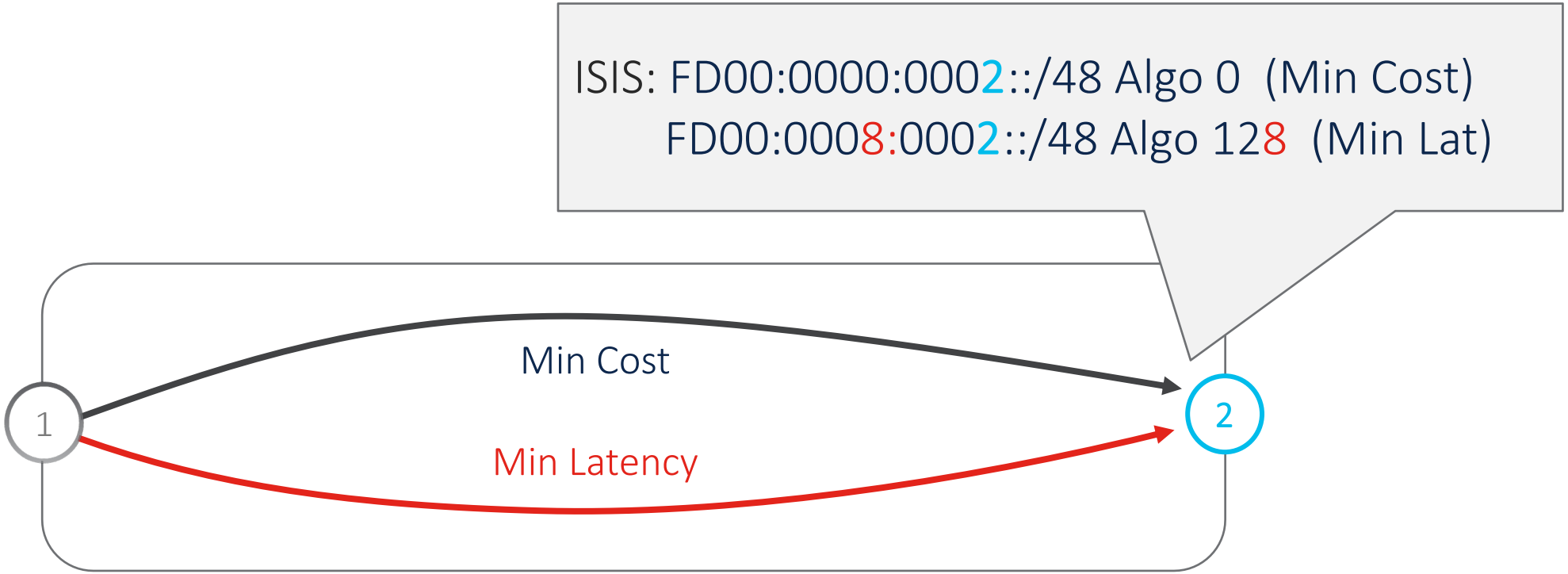
32-bit Private Block, one per Algo

- Min Cost ⇔ FD00:0000::/32 (0 because base algo)
- Min Latency ⇔ FD00:0008::/32 (Flex Algo 128)



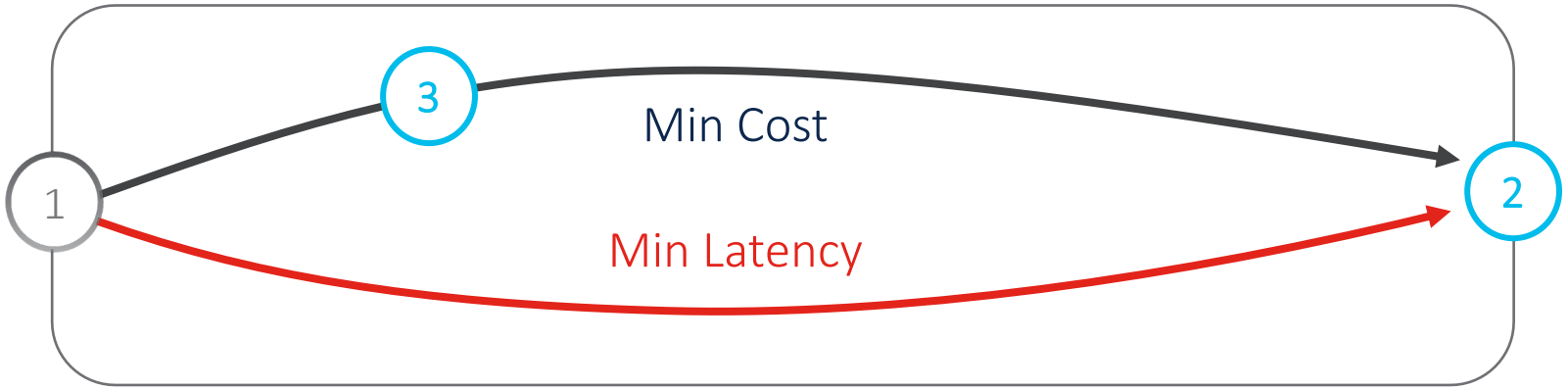
16-bit uSID

- 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



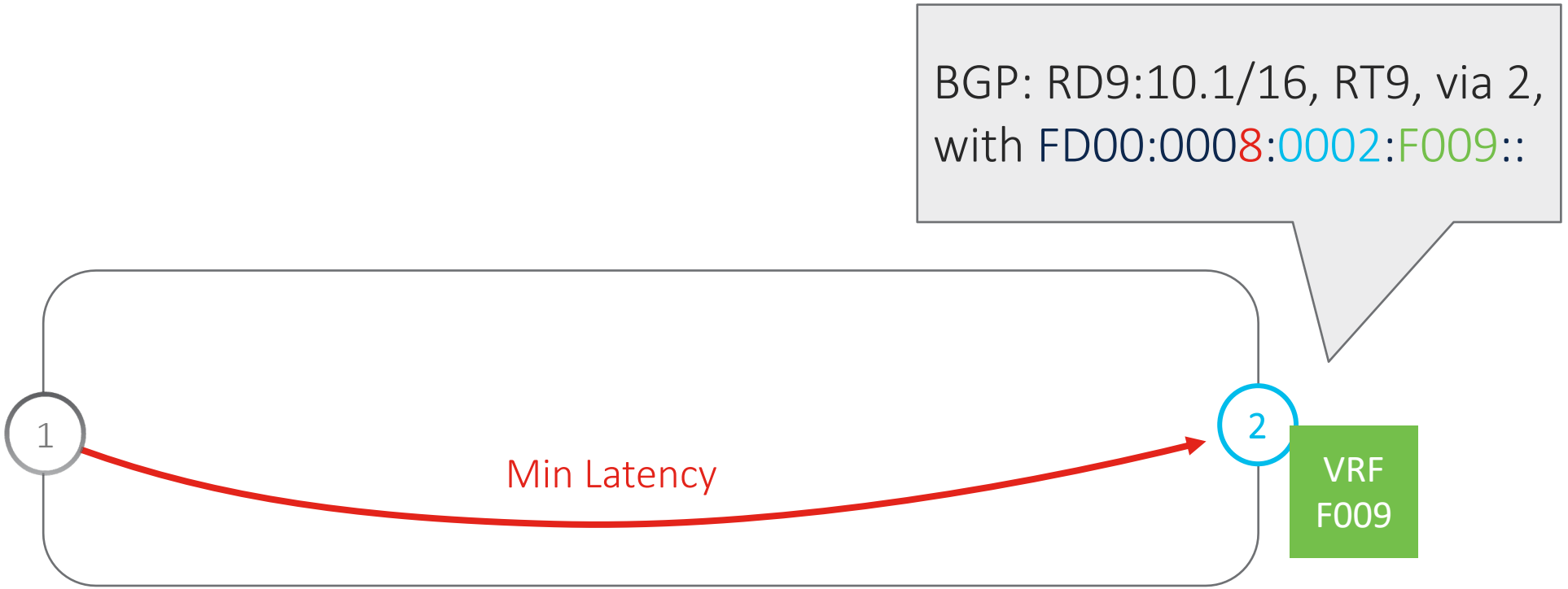
Seamless Deployment

- Node 3 forwards as per legacy longest match FD00:0000:0002::/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
- FD00:0008:0002:F009::**/64** ➔ It is me and this is for VRF 9

Intuitive & Rich

- FD00:0000:0002:F009:0000:0000:0000:0000
  - Min Cost Path to node 2 and then lookup in VRF9
- FD00:0008:0002:F009:0000:0000:0000:0000
  - Min Lat Path to node 2 and then lookup in VRF9
- FD00:0000:0003:0002:F009:0000:0000:0000
  - Min Cost Path to node 3 then 2 and then lookup in VRF9
- FD00:0008:0003:F005:0002:F009:0000:0000
  - Min Lat Path to node 3, VNF 5, Min Lat 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

If more than 6 uSID’s are required

Outer DA: FD00:0000:0001:0002:0003:0004:0005:0006

uSID1 uSID2 uSID3 uSID4 uSID5 uSID6

SRH: FD00:0000:0007:0008:0009:0010:0011:0012

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

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
- Standardized, Rich Eco-system, Rich Open Source (SONiC)

Outperform 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, Ericsson, ZTE, Arrcus, Arista, H3C

Merchant Silicon

Broadcom, Marvell, Barefoot Networks (Intel)

SRv6 uSID

Open-Source Applications

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

Open-Source Networking Stacks

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

Smart NIC / DPU

Intel, NVIDIA

Partners

kaloom, Trend Micro, ENEA Qosmos Division, NoviFlow, ixia A Keysight Business, Spirent Promise. Assured., KEYSIGHT TECHNOLOGIES, 6WIND, f5

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

Over 80000 uSID routers deployed

Inter-DC/Metro Traffic Engineering across all of China
Eddie Ruan

14k+ devices, 70% services on uSID
Akash Agrawal

Telefonica Vivo - uSID Deployment

LACNIC Blog > IPv6 > Implementation of SRv6 uSID in Telefónica VIVO's Infrastructure

Implementation of SRv6 uSID in Telefónica VIVO's Infrastructure

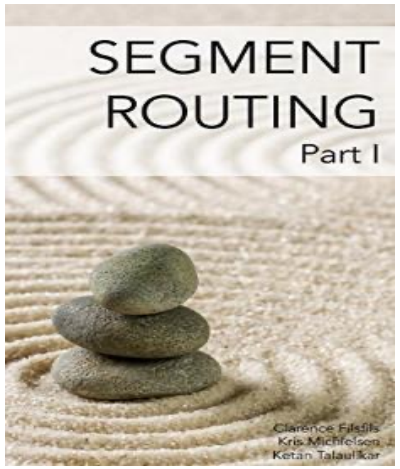
24/01/2024

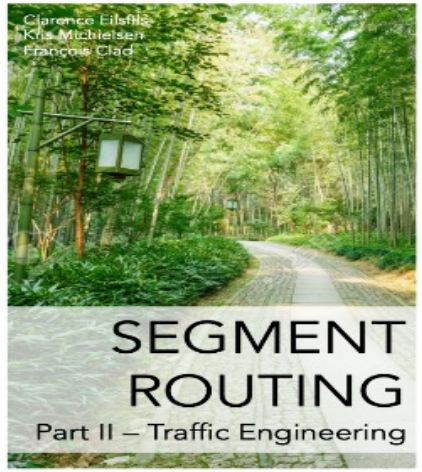
By Nelson Jose dos Santos Junior, Telecom Specialist

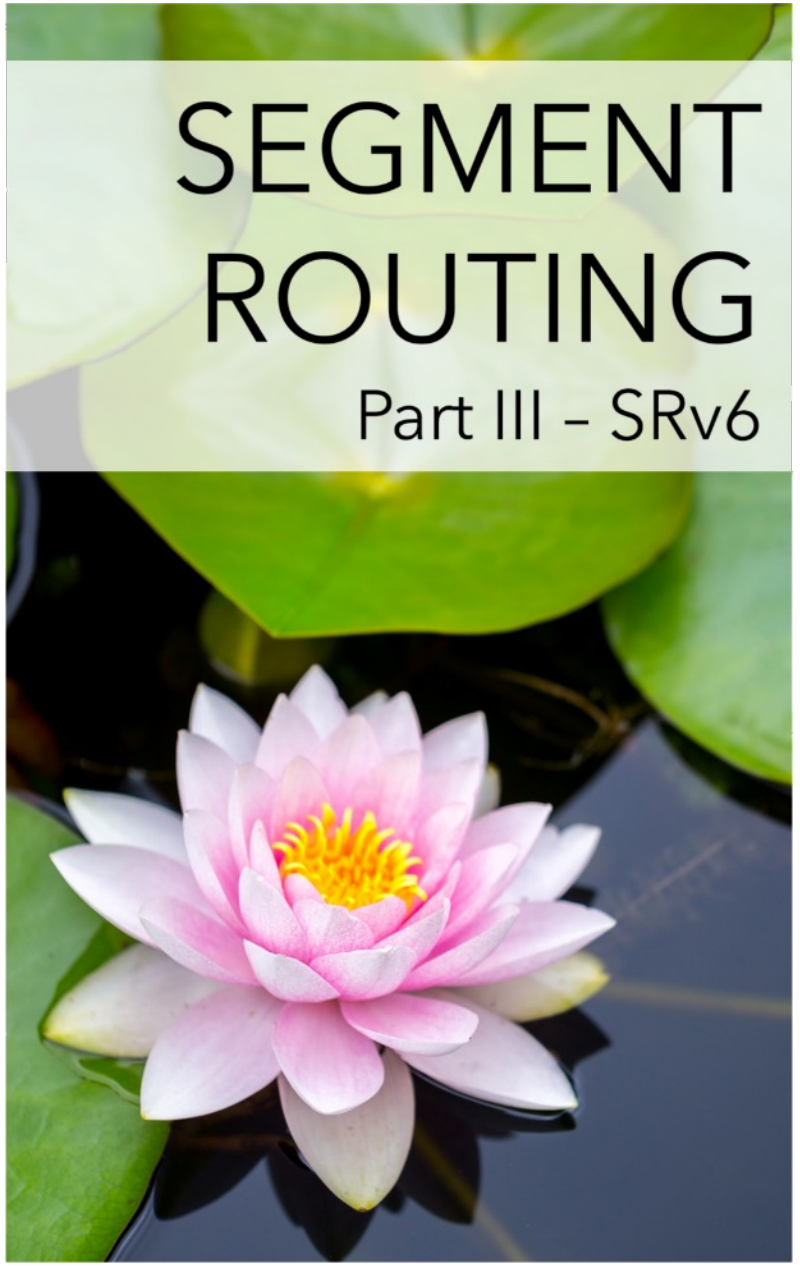
Simplicity Always Prevails

Join our next uSID/IPM event – 9 & 10 October 2024

Stay up-to-date

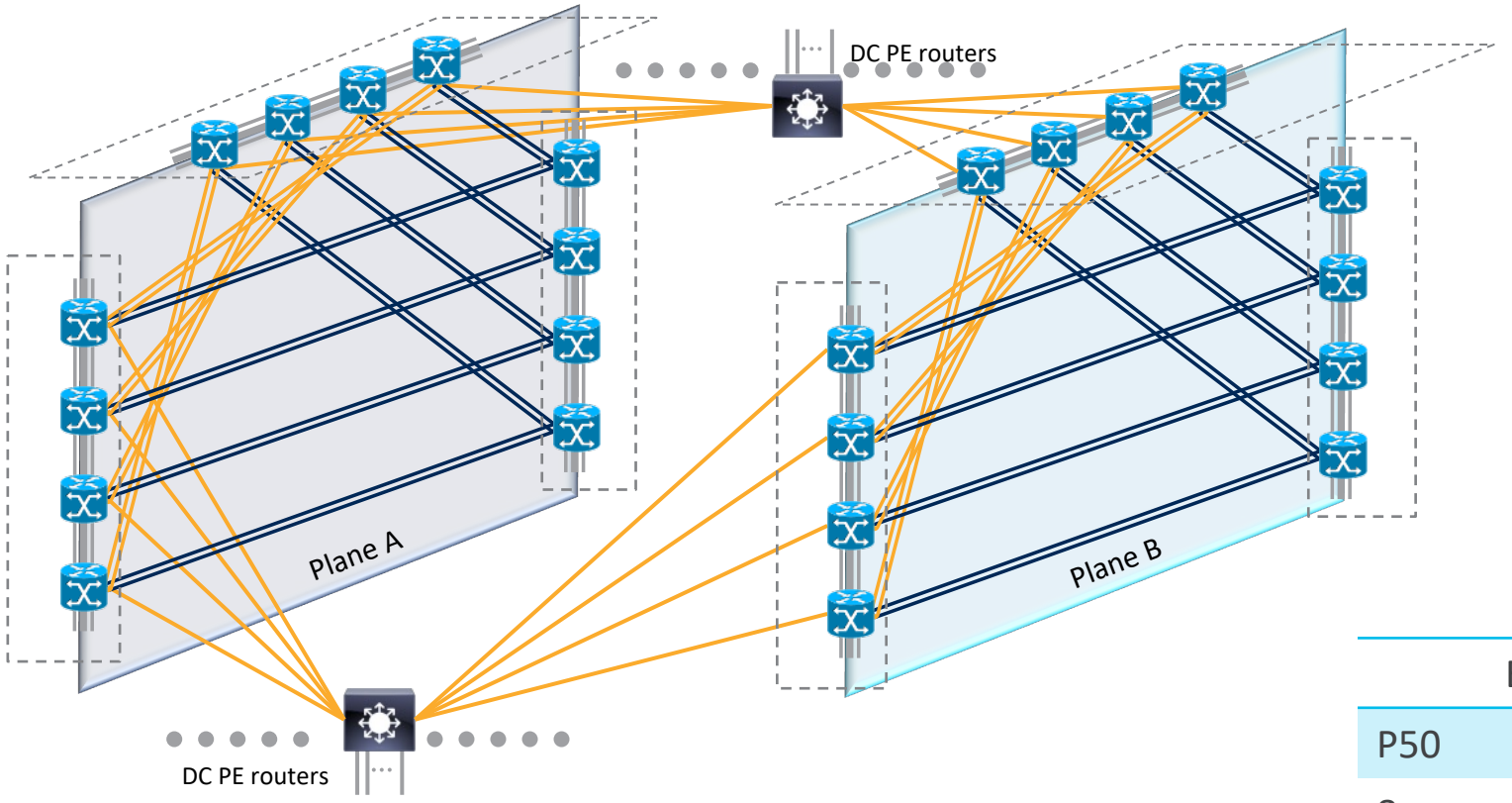






segment-routing.net

The nature of IP is ECMP



ECMP Paths		
P50	P75	P90
8	16	32

- Legacy solutions do not have the scale to measure all ECMP paths

The experience of all clients must be measured

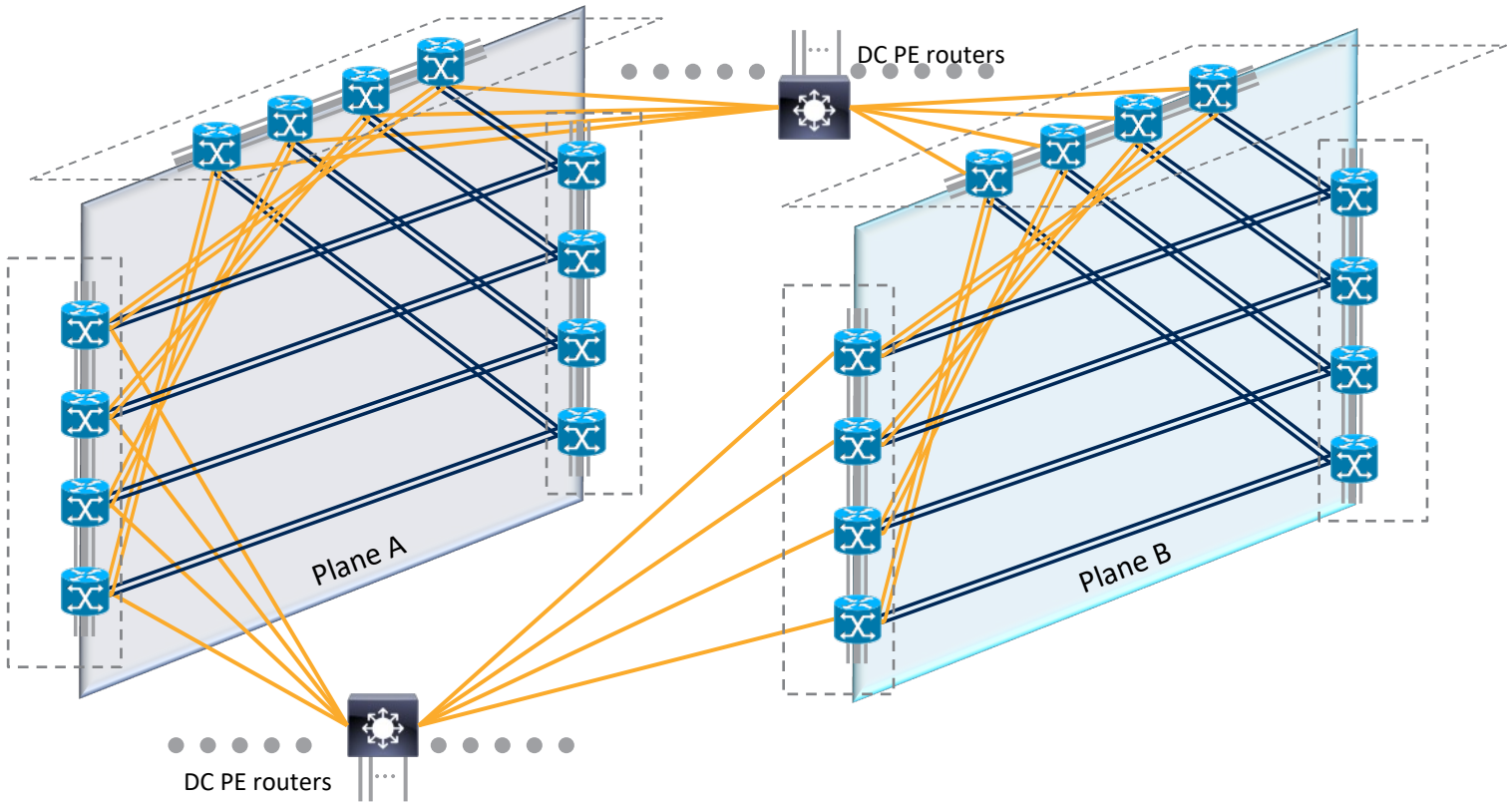
Would a bank accept to monitor < 0.1% of its access?

Legacy solutions are typically 1000 to 10000 times not scalable enough

Legacy coverage is < 0.1%

Operators learn outages from clients

Silicon One provides 14M probes per sec

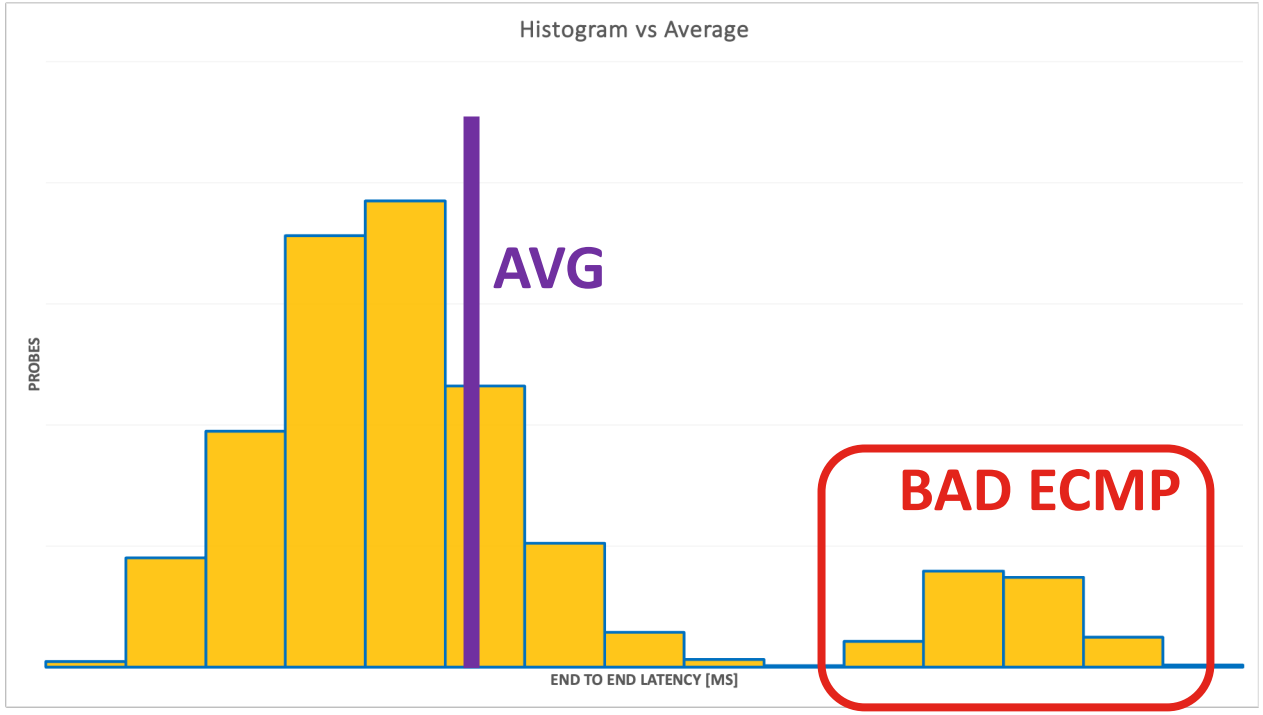


- 1 measurement every msec
- 500 edges
- 16 ECMP paths

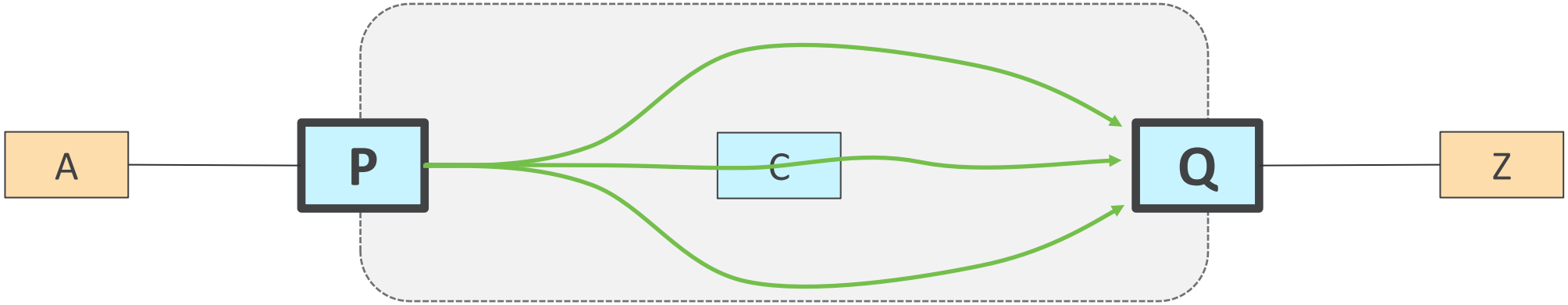
8M probes per sec (57% of Silicon One capability)

Richer Metrics

- 1 bad path out of 8 ECMP
- 12.5% of the clients impacted
- Average hides the issue
- IPM Histograms reports the experience of the whole population



Any IP Fabric, Any Edge to any Edge, Any ECMP Path

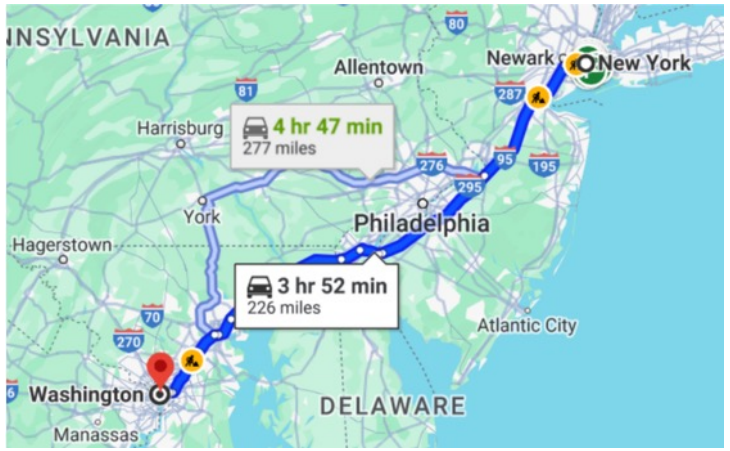


- Absolute Loss
- One-Way Latency (20nSec)
- Liveness (sub-2msec)
- Standard: STAMP (RFC 8762 & RFC 8972)

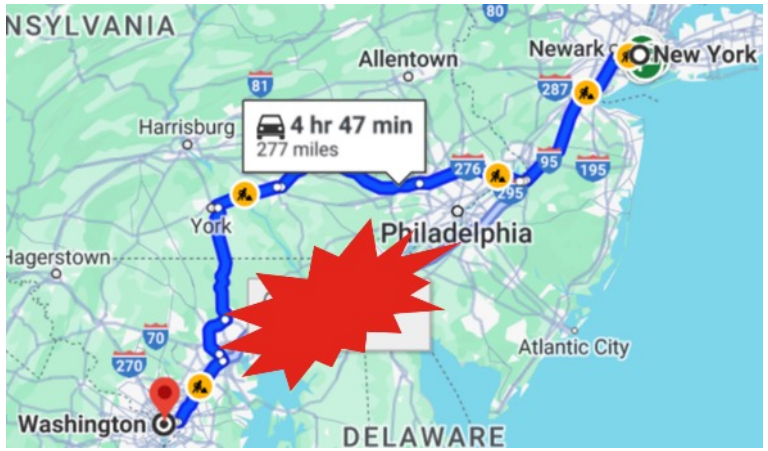
Much Cheaper through Silicon Integration

- Capex Elimination
  - SLA Appliance
  - Router port to appliance
- Opex Elimination
  - Rack Space
  - Power

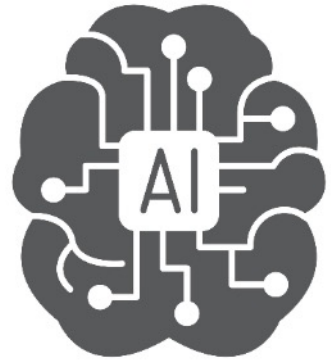
Continuous Correlation to Routing



Measured Latency
compared to best topology



Measured Latency
compared to current topology



- Time-series of Measurements from any P to any Q along any ECMP path
- Time-series of ECMP routed paths from any P to any Q

Inference

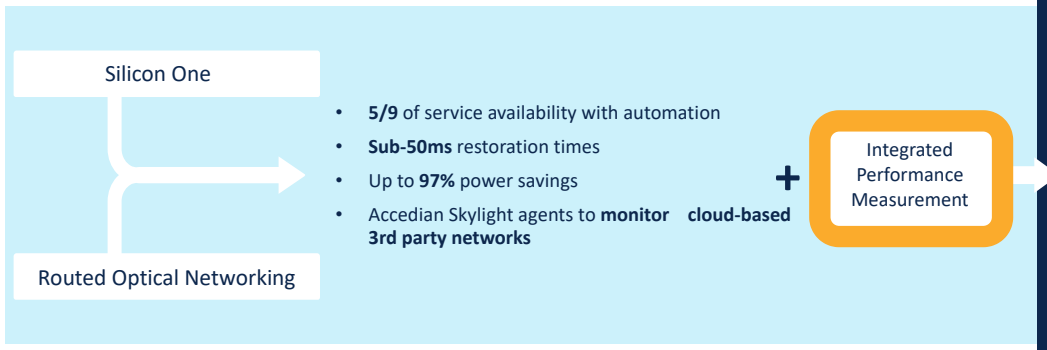
- Measurement (**PAR**, **MAD**) report SLA degradation (e.g., loss)
- Without any additional measurement, Routing Correlation allows to infer other (SRC, DST) pairs that are also impacted
  - BRU to MAD/LIS/SEV is impacted
  - LON to MAD/LIS/SEV is impacted

Michael Valentine, Technology Fellow, Network Architecture, Goldman Sachs

- uSID and IPM Use-case
- IPM Silicon Integration and Metrics
- Link

Bart Janssens, Senior Specialist Packet Architecture, Colt Technology Services

- Routing Analytics
- Accedian Skylight
- Deployment and Use-Cases
- Link



Helping Colt and its customers **lead in performance and sustainability**

Next phase

Accedian Skylight

End to end service visibility

Insights with micro-second granularity

Faster to launch and monetize SLA-backed services

New on-demand services based on power and SLA based metrics

Gyan Mishra, Associate Fellow
Verizon

- DC use-case
- uSID and IPM
- Lightweight Host Routing (LHR)
- Link

Eddie Ruan, Senior Staff Engineer
Alibaba

- uSID - Deployment Experience
- SONIC Experience
- Link

IP is better than ever

Build Anything End-to-End

Measure Everything


CISCO