Segment Routing Deployments and Demonstrations at Interop Tokyo ShowNet

Ryo Nakamura†*, Teppei Kamata‡*

†The University of Tokyo

‡Cisco Systems

*ShowNet NOC Team

Interop® 23 Tokyo

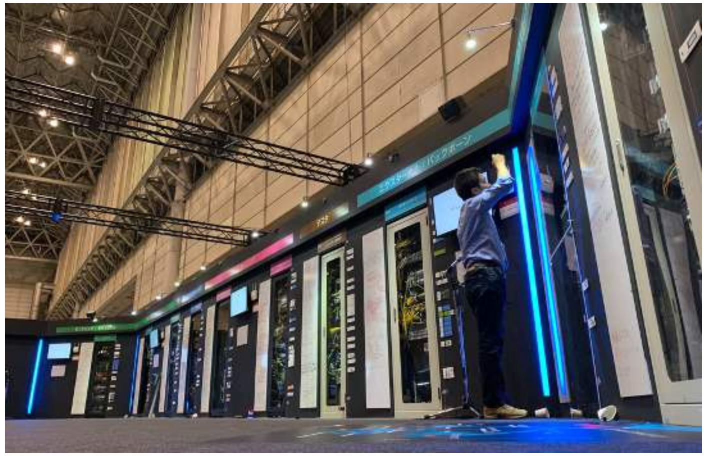
JUNE 14 - 16
MAKUHARI MESSE, JAPAN

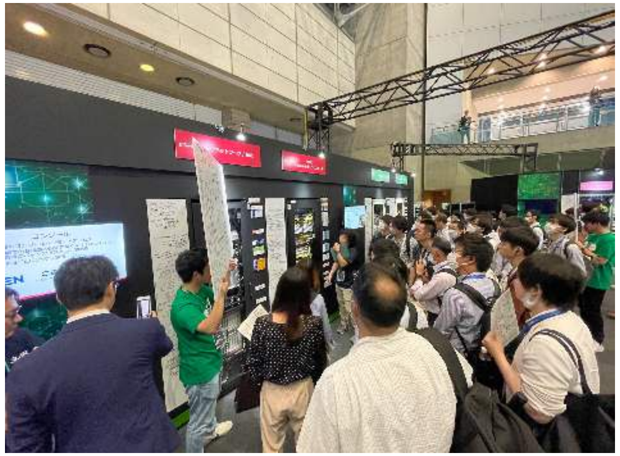
- **Interop Tokyo**
  - The largest annual exhibition of Internet Technologies in Japan
  - Over 200 booths and about 120,000 visitors for 3 days
  - About 270 sessions in the exhibition and Interop conference

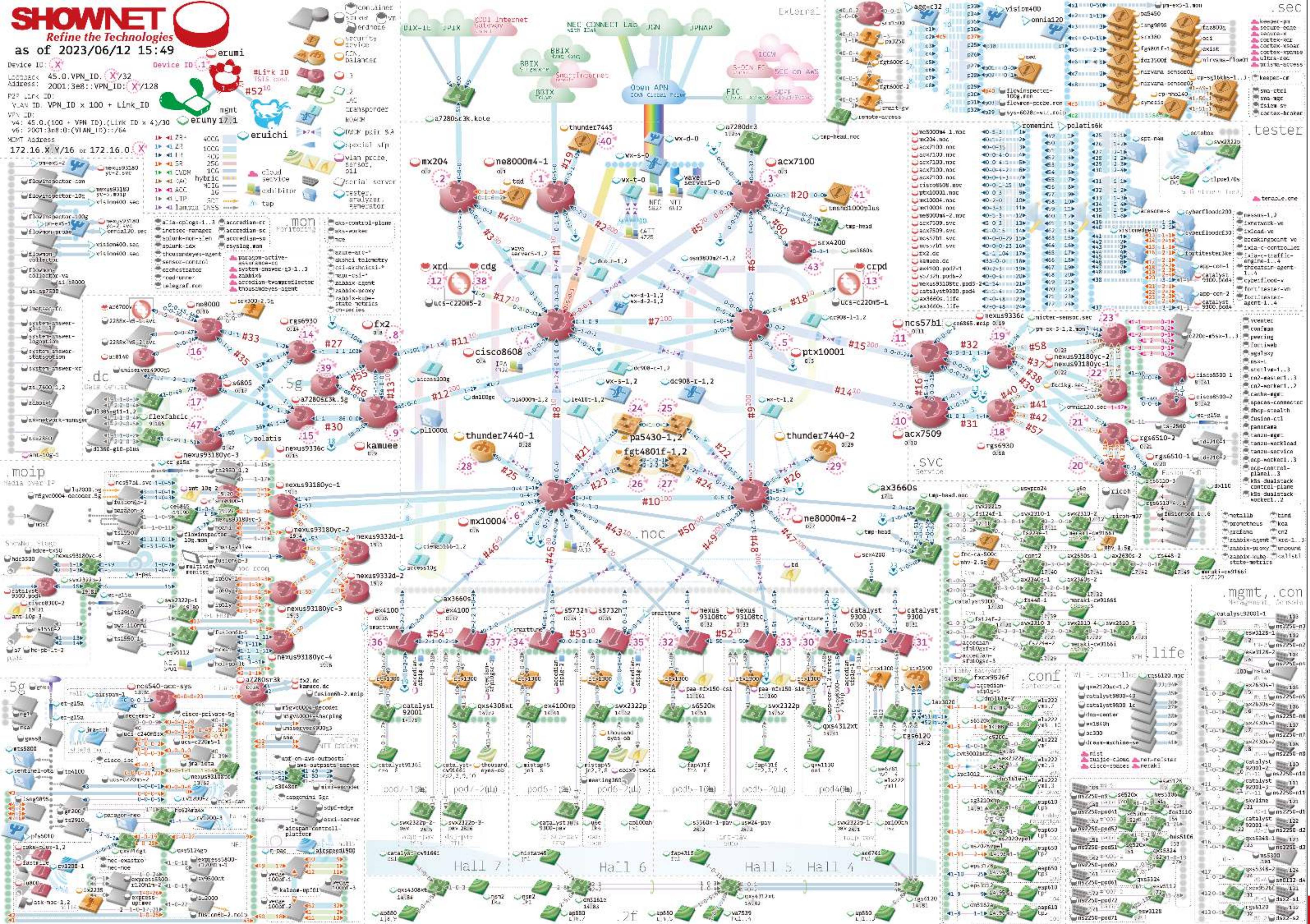
SHOWNET Refine the Technologies

ShowNet (AS290)

- **Live Demonstration Network built at Interop Tokyo**
  - Over 20 full-height racks with contributed networking devices
    - 6 months for design, 12 days for construction, and 3 days for the exhibition
  - Providing internet connectivity for Interop exhibitors and visitors
  - Conducting interoperability tests and demonstrations of the latest networking technologies every year







ShowNet 2023 Topology Diagram

Icons:

- Router
- Firewall
- Switch
- Wireless AP
- ROADM
- Server

and more

Who makes ShowNet?

- NOC Team
  - about 30 specialists from xSPs, vendors, and academia
- Contributors
  - over 600 product specialists from vendors
- Volunteers
  - about 40 persons from academia (students) and industry (junior staffs)
    - An educational aspect



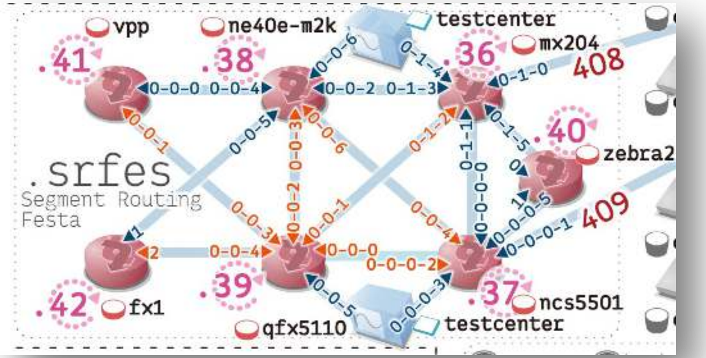

SHOWNET Refine the Technologies

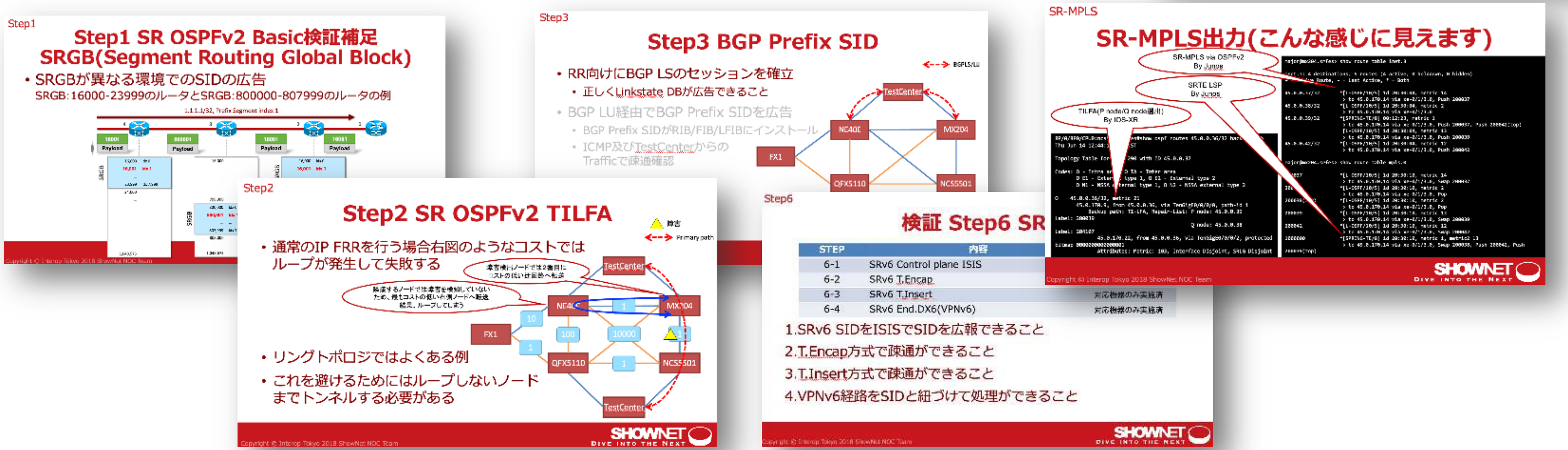
Segment Routing at Interop Tokyo ShowNet

- We have been continuously testing/showcasing SR at ShowNet
  - 2018: SR-MPLS and SRv6 Inter-operability test
  - 2019: SRv6 Service Chaining
  - 2020: Interop Tokyo was canceled due to COVID-19
  - 2021: SR-MPLS and SRv6 Backbone and an experiment with BGP-EPE
  - 2022: SRv6-L3VPN Backbone (IPv6 single stack)
  - 2023: SRv6 without L3VPN (Global Routing Table) Backbone

Segment Routing Interop in 2018

- Interop test with 8 device from 5 vendors
  - SR-MPLS: OSPFv2 Basic + TI-LFA, BGP Prefix SID, TE, SRTE&PCEP
  - SRv6: ISIS, T.Encap, T.Insert

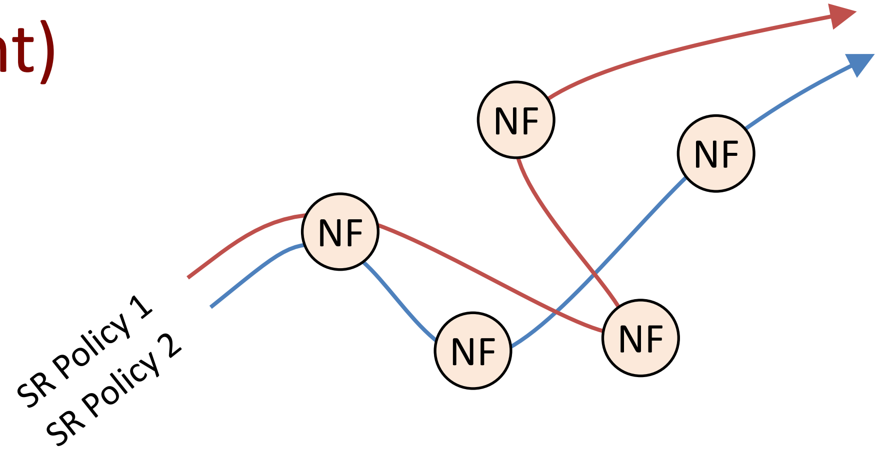




SHOWNET
Refine the Technologies

SRv6 Service Chaining in 2019

- Service Programming with Segment Routing
  - draft-ietf-spring-sr-service-programming
  - A SID represents a service (Service segment)
    (e.g., VNFs or physical appliances)
  - SID List represents a series of services
    that are applied to packets

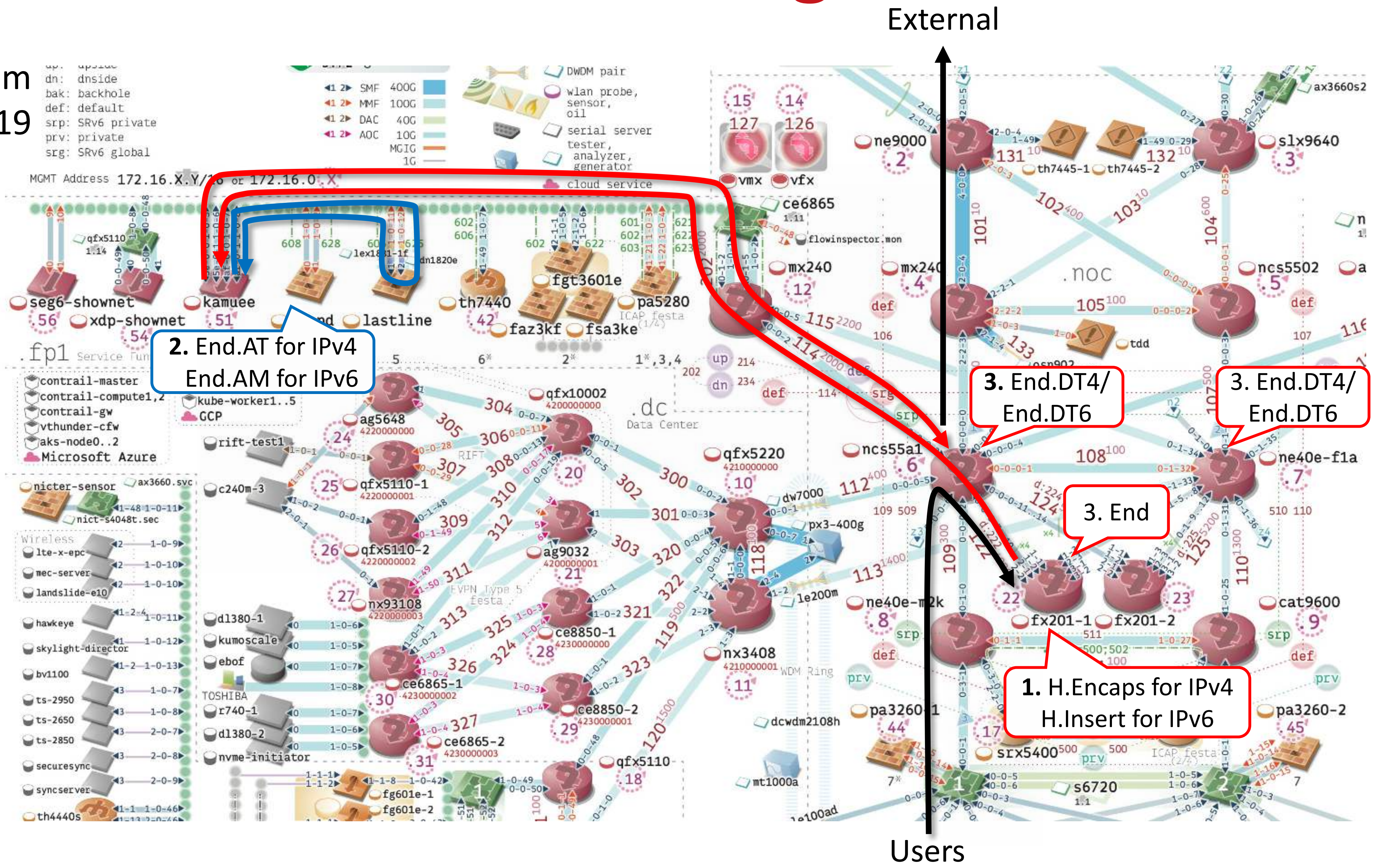


- In ShowNet 2019
  - We built SRv6-based Service Chaining
    with three SRv6-capable products
    and five SRv6 proxies (two products and three OSS)

SHOWNET Refine the Technologies

2019 SRv6 Service chaining Overview

Topology diagram of ShowNet 2019

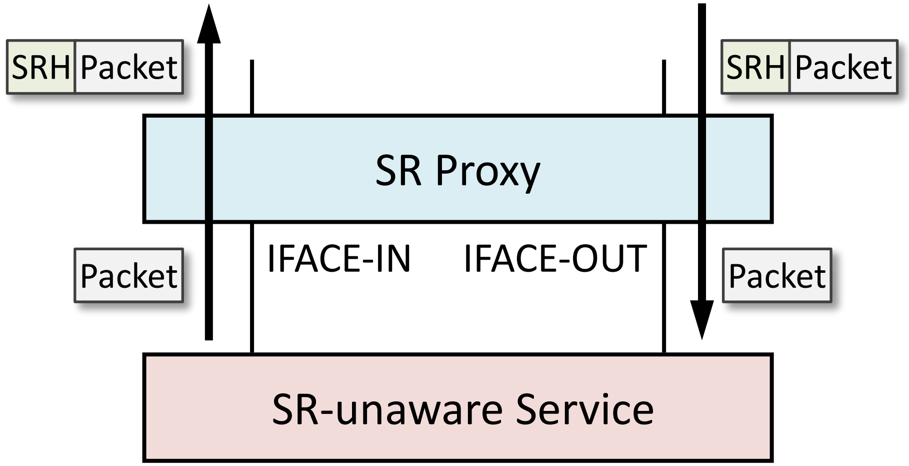


SHOWNET Refine the Technologies

SR-Proxy

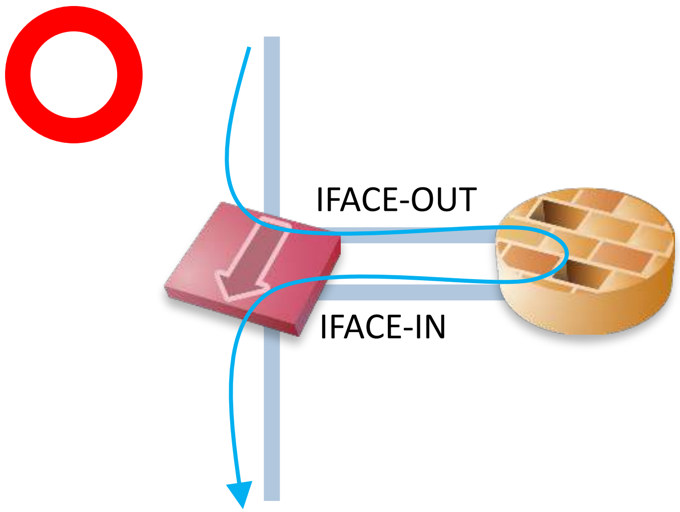
- Not all services (NFs) may be capable of SRv6 encapsulation
  - SR-unaware services
- SR-Proxy integrates SR-unaware services into SRv6 networks
  - SR-Proxy pops SRH and passes inner packets to a service (IFACE-OUT)
  - SR-Proxy pushes SRH to returned packets (IFACE-IN)

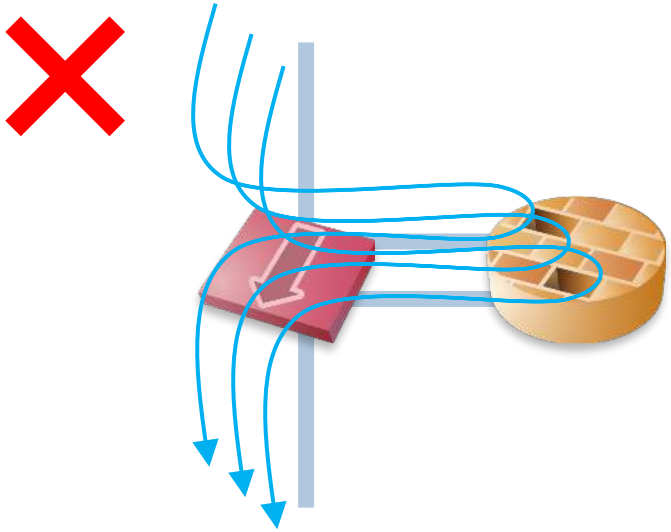
There are several mechanisms for SR-Proxies
as per draft-ietf-spring-sr-service-programming



End.AS, static proxy

- IFACE-OUT: Pop outer headers (SRH, IPv6)
- IFACE-IN: Push (static) outer headers (SRH, IPv6)
- Pros: easy to implement
- Cons: a single End.AS proxy cannot accommodate multiple service chains





End.AM, Masquerading proxy

- IFACE-OUT: Overwrite the IPv6 DA with the last SID (SL[0]) in the SID List
- IFACE-IN: Overwrite the IPv6 DA with the next SID (SL[1]).
- Pros: stateless and easy to implement
- Cons: not easy to applicable to IPv4-in-IPv6 packets

SRv6 Network

IPv6 DA	C::1
Segment List[0]	A::1
Segment List[1]	B::1
Segment List[2]	C::1
SegLeft	2

Translation at IFACE-OUT

SR-unaware Function

IPv6 DA	A::1
Segment List[0]	A::1
Segment List[1]	B::1
Segment List[2]	C::1
SegLeft	2

Translation at IFACE-IN

SRv6 Network

IPv6 DA	B::1
Segment List[0]	A::1
Segment List[1]	B::1
Segment List[2]	C::1
SegLeft	1

SHOWNET Refine the Technologies

End.AT, Caching Proxy for IPv4 in SRv6 traffic

- IFACE-OUT: Cache the IPv6 header before passing the packet to the service and embed the index of the cache entry into the ToS field of inner IPv4 header
- IFACE-IN: Restore the IPv6 header including SRH from the cache entry indicated by the ToS value of the inner IPv4 packet
- draft-eden-srv6-tagging-proxy (not maintained since then...)

IPv6 DA	C::2
Segment List[0]	A::1
Segment List[1]	B::1
Segment List[2]	C::2
SegLeft	2

Remove IPv6 Hdr at IFACE-OUT

(inner) IPv4 Packet

ToS	2

Cache the IPv6 header in cache entries[2]

Restore IPv6 Hdr at IFACE-IN

IPv6 DA	B::1
Segment List[0]	A::1
Segment List[1]	B::1
Segment List[2]	C::2
SegLeft	1

SRv6 Network | SR-unaware Function | SRv6 Network

SHOWNET Refine the Technologies

Lessons Learned in 2019

- SRv6 Data Plane with Service Segments certainly works!
- No need to implement hop-by-hop paths to achieve service chains
  - There is no states in the network unlike OpenFlow-like solutions
- The implementations lacked control plane capabilities in 2019
  - We configured everything statically, e.g., SR policies and End behaviors
  - In 2024, we have PCEP

The details are described in draft-upa-srv6-service-chaining-exp

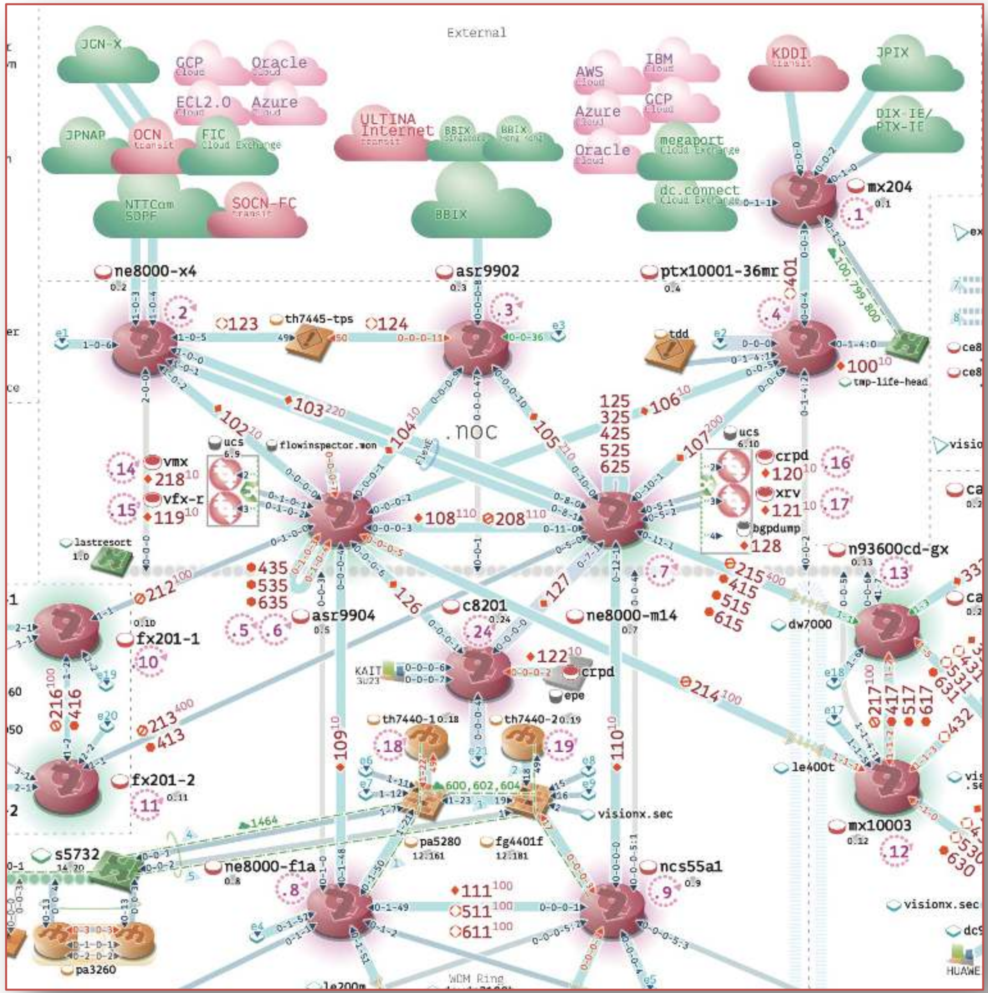
Table of Contents

1. Introduction 2
2. Terminology 3
3. SRv6 service chaining at Interop Tokyo 2019 ShowNet 3
4. Lessons Learned 5
   4.1. Transparency of SRv6 header 5
   4.2. Services that cannot co-exist with End.AM 6
   4.3. Service liveness detection and conditional advertisement of service segments 6
   4.4. TTL Decrement on SRv6 Proxies 6
   4.5. Control Plane Capabilities 7
   4.6. Match Condition for Applying SRv6 Functions 7
5. IANA Considerations 8
6. Security Considerations 8
7. Contributors 8
8. Acknowledgements 8
9. Normative References 8
Authors' Addresses 10

SHOWNET Refine the Technologies

Measurement Experiment with BGP-EPE in 2021

- SR-MPLS-based Backbone Network
  - Three external routers and
    Five core routers from three vendors
    run SR-MPLS with ISIS and BGP
  - Flex-Algo interop and demonstration
  - Measurement experiment with
    Egress Peer Engineering

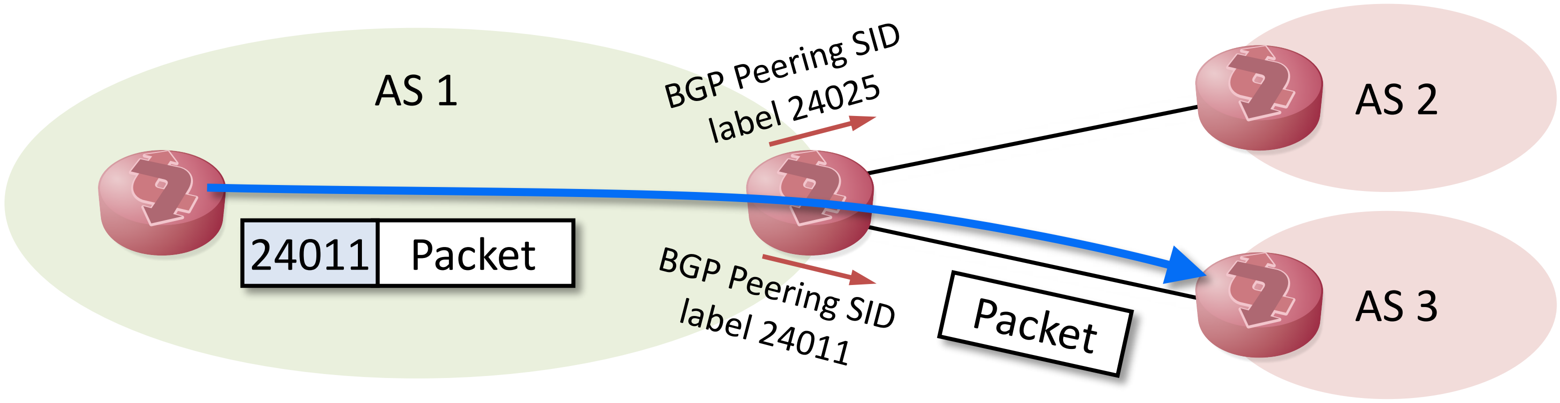


Topology diagram
of ShowNet 2021

Egress Peer Engineering defined in RFC9087

- Steering given packets to chosen egress ASes
  1. ASBR assigns BGP Peering SIDs to eBGP peers
  2. ASBR transmits packets with a BGP Peering SID to the peer corresponding to the SID regardless of underlying routing

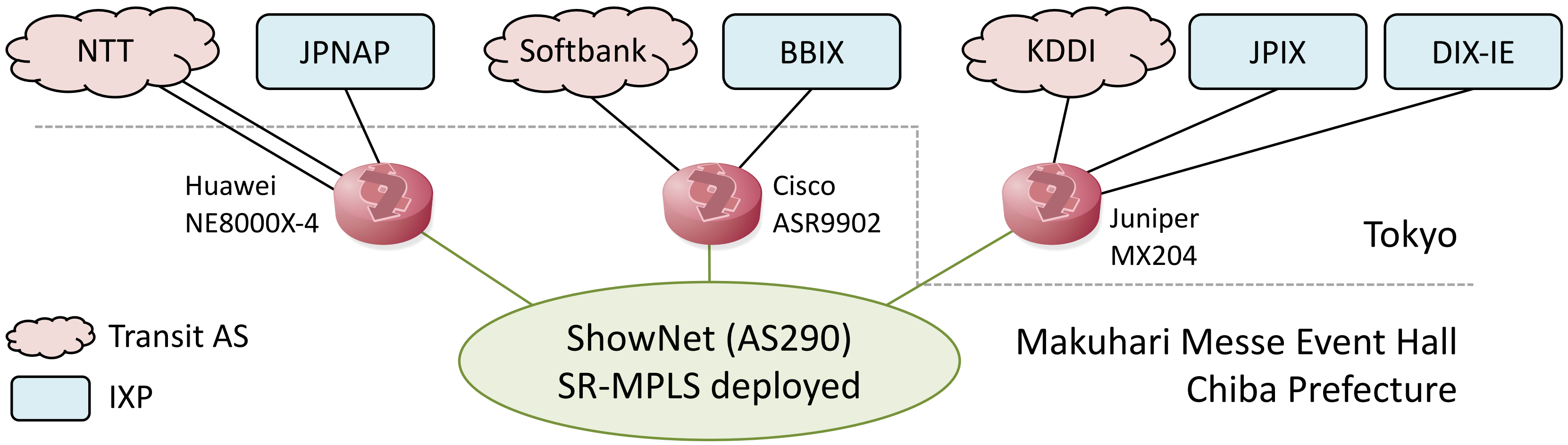
  ✓ No impact on other (non-encapsulated) traffic



SHOWNET Refine the Technologies

Measuring the Internet via different Egress ASes

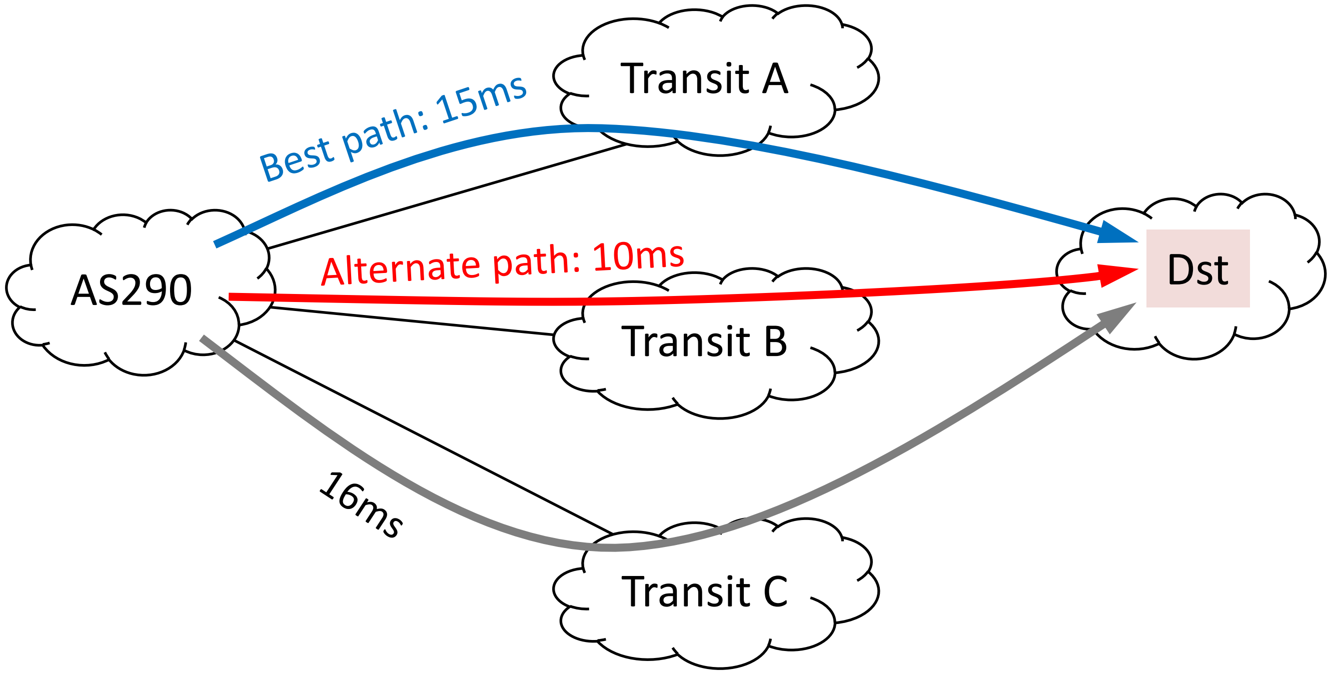
- ping and traceroute to 2.6M IPv4 addresses via 101 eBGP peers
  - Three transit ASes and 43 peer ASes who agreed to join the experiment
  - We obtained RTTs to a destination via all possible paths



SHOWNET Refine the Technologies

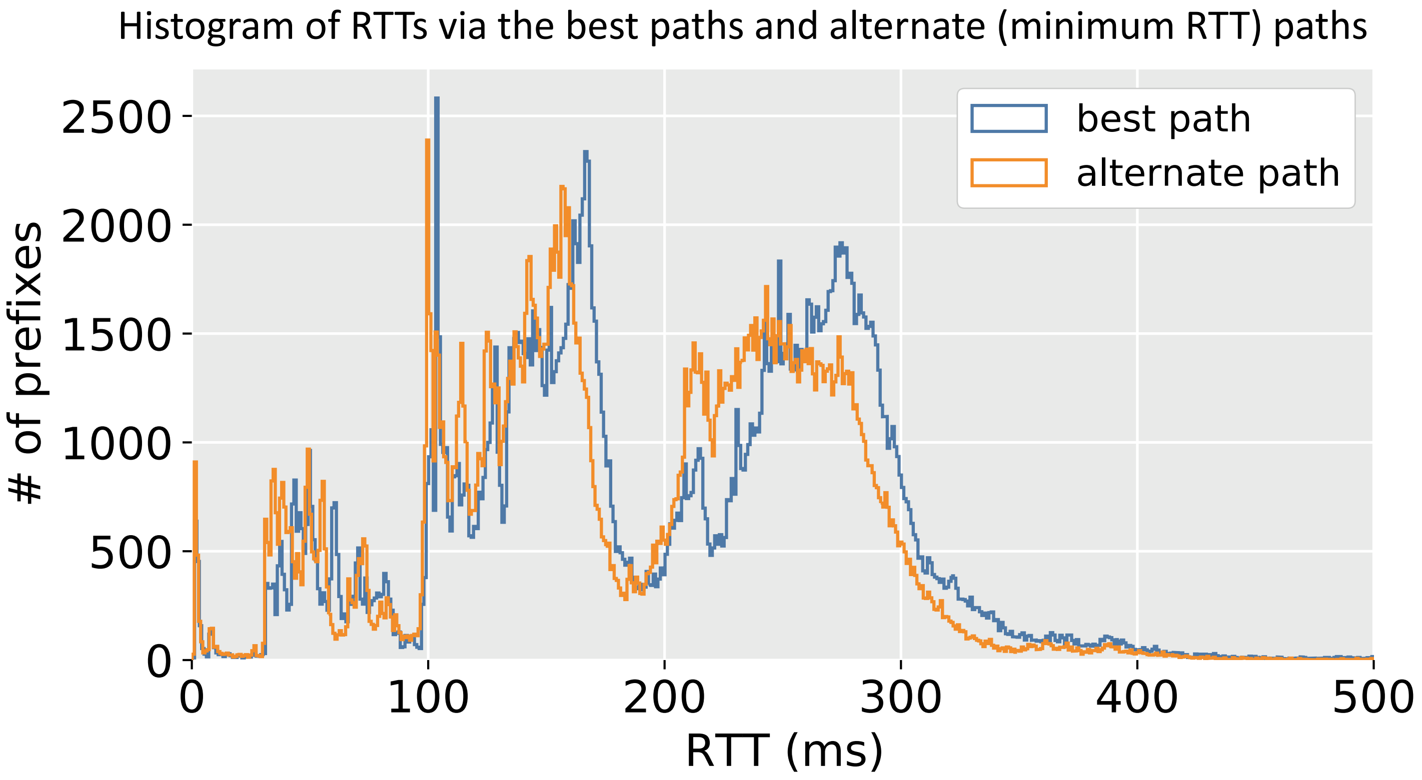
How best are the best paths?

- We revealed *how best are the best paths* from the obtained data
  - ✓Comparing RTTs via best paths and alternate paths for a destination



How best are the best paths?

- There are shorter latency paths than the BGP best paths
  - This graph shows the histogram of # of BGP prefixes per latency

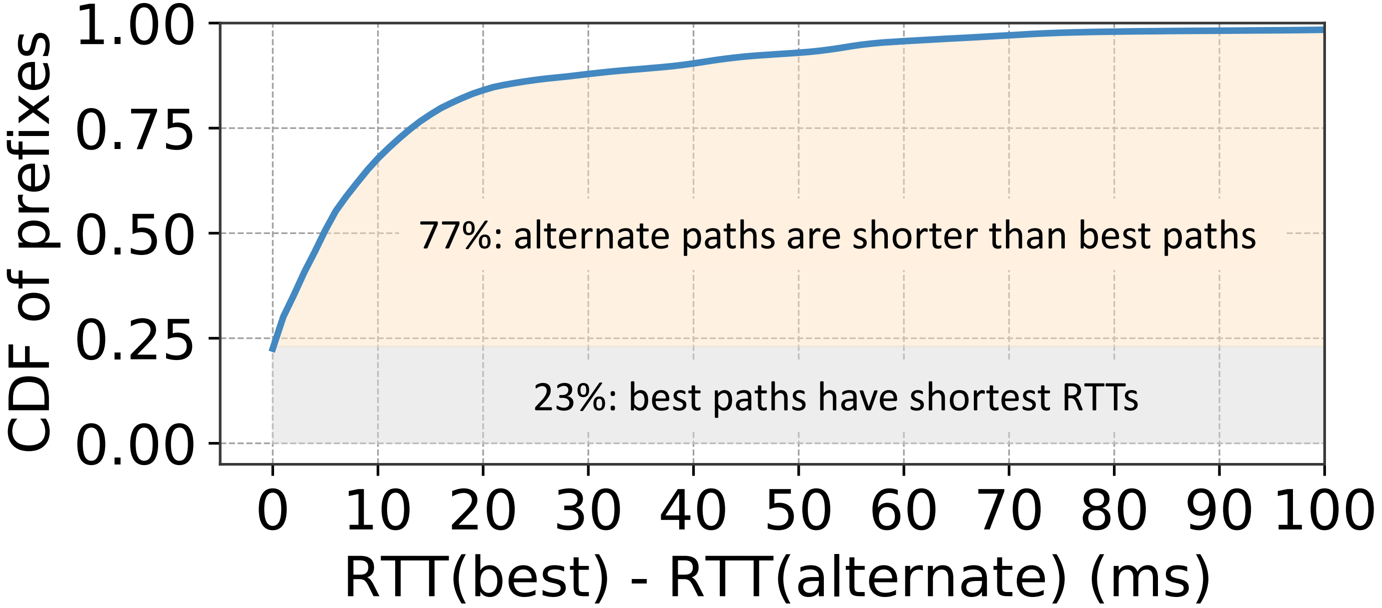


How best are the best paths?

- There are shorter latency paths than the BGP best paths
  - 77% of measured prefixes have alternate paths of shorter latency than their best paths
  - 17% of the prefixes got 10-20 msec latency improvement by alternate paths

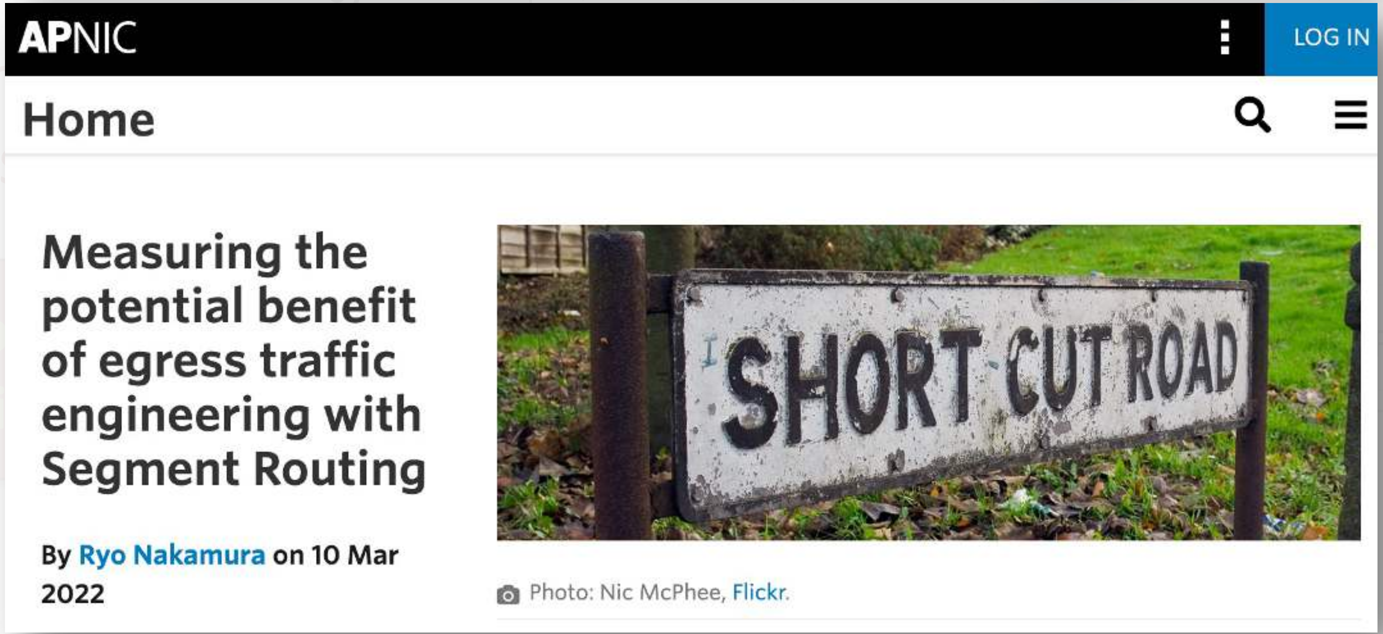
There is room to improve latency by BGP-EPE with SR!

CDF of latency improvement by alternate paths:
RTT(alternate) - RTT(best) for each prefix



SHOWNET Refine the Technologies

More results



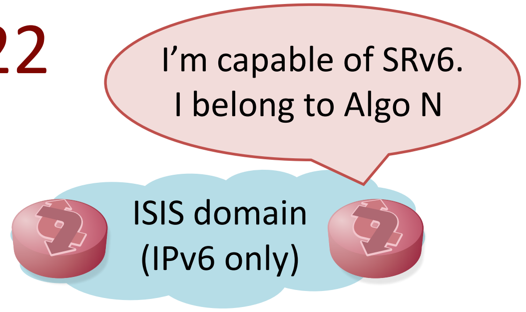
SRv6 Deployment in ShowNet 2021

- It was the first time to deploy SRv6 in ShowNet Backbone
  - 5 products from Cisco, Furukawa, Huawei, and Juniper
- Struggling to interop SRv6 L3VPN... draft-ietf-bess-srv6-services
- How to encode IPv6 next-hops for IPv4 NLRI
  - draft 01 (2019/11) references RFC5549
  - draft 02 (2020/2) - 05 (2020/11) references draft 5549revision
  - draft 06 (2021/3) references RFC8950
  - We faced both types of routers implementing RFC5549 and RFC8950

SHOWNET Refine the Technologies

SRv6 Deployment in ShowNet 2022

- The second time
  - 9 products from Cisco, Furukawa, Huawei, and Juniper
  - draft-ietf-bess-srv6-services was published as RFC9252 (2022/7)
- SRv6 L3VPN as per RFC9252 worked well
- We further tried to interop SRv6 Flex Algo over ISIS
  - draft-ietf-lsr-isis-srv6-extensions-18 (published as RFC9352 in 2023/2)
  - draft-ietf-lsr-flex-algo-20 (published as RFC9350 in 2023/2)
  - Not tried in ShowNet 2021, and it worked well in 2022 (except for delay metric)

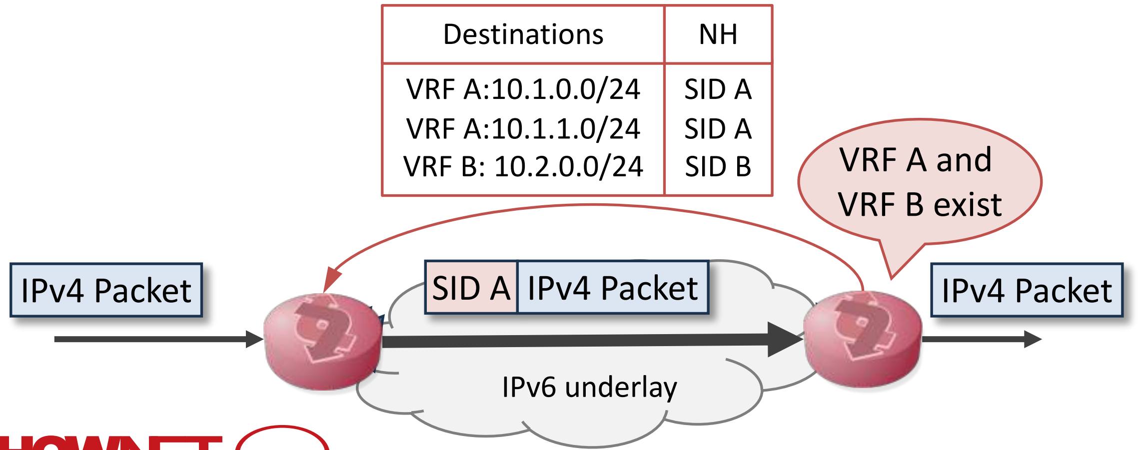


SHOWNET Refine the Technologies

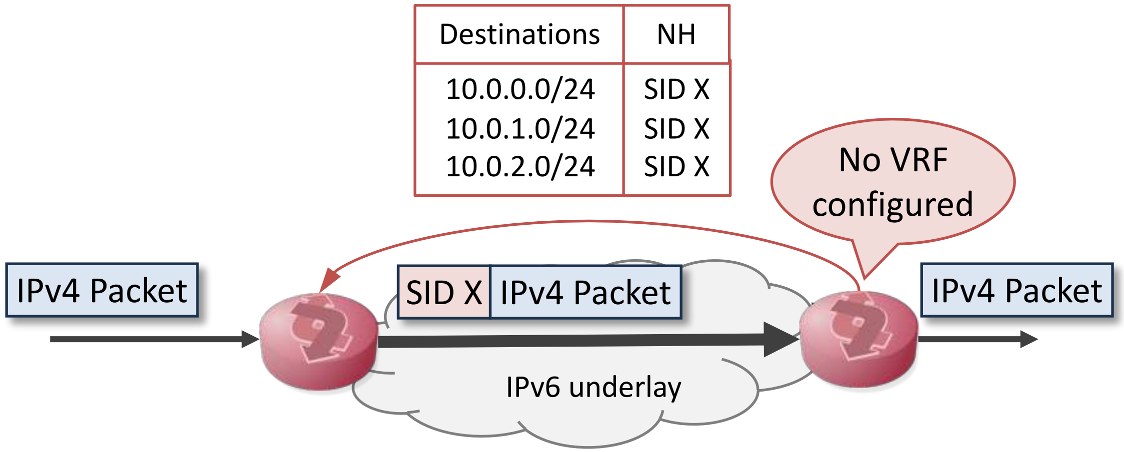
SRv6 Deployment in ShowNet 2023

- The third time
  - 11 products from Cisco, Furukawa, Huawei, Juniper, and NTTcom
- We tried to interop SRv6 without L3VPN: BGP AFI 1, SAFI 1
  - Egress PE advertises IPv4 NLRI with SRv6 Service SID as its next-hop
  - Ingress PE installs routes to IPv4 (or IPv6) destinations via SIDs in the default table
  - Deriving the benefit from SRv6 transport without IP-VPNs

NLRI AFI 1, SAFI 128 (VPNv4)



NLRI AFI 1, SAFI 1



```
inet.0: 917099 destinations, 1834194 routes (916664 active, 838429 holddown, 866 hidden)
130.69.0.0/16 (2 entries, 1 announced)
        State:
        +BGP    Preference: 170/-151
                Next hop type: Unusable, Next hop index: 0
                Address: 0x5639ce9148dc
                Next-hop reference count: 2202742, key opaque handle: (nil), non-key opaque
handle: (nil)
                Source: 2001:3e8::12
                State:
                Local AS:   290 Peer AS:   290
                Age: 10:51:42        Metric: 0
                Validation State: valid
                Task: BGP_290.2001:3e8::12+48625
                AS path: 2907 2501 I  (Originator)
                Aggregator: 2501 133.11.255.160
                Cluster list:  0.0.0.12
                Originator ID: 45.0.0.3
                Communities: large:290:1000:2 large:290:2023:11 large:290:2023:330
large:290:2023:3301 large:45686:1000:2 large:45686:1001:1
                Accepted MultiNexthop RecvNextHopIgnored
                SRv6 SID: 2001:3e8:fa00:3:4:: Service tlv type: 5 Behavior: 19 BL: 40 NL: 24
FL: 16 AL: 0 TL: 0 TO: 0
                Localpref: 150
                Router ID: 45.0.0.12
                Thread: junos-main
```

This route is in the default table (not VRF)

Its next-hop is an SRv6 SID

Looking back at the chronicle since 2019

- Implementations are steadily improving each year
  - In 2019, there were only limited data plane features for SRv6
  - Since 2021, SRv6-based L3VPN works
- IPv6 single-stack transport is very simple and easy to operate
  - No need to assign IP addresses on links because of IPv6 LL Address
  - It means there is no need to manage /30 (or /31) IPv4 prefix per link
- Scalability with Address summary
- Flexible use cases,
  e.g., service chaining and EPE, but require extra effort

```
interface FourHundredGigE0/0/0/24.14
 description fhg-0-2-0.ptx10k.noc
 mtu 9021
 ipv6 enable
```

SHOWNET Refine the Technologies

Conclusion

- We have demonstrated SR in Interop Tokyo ShowNet since 2018
- SR can build networks simply and easily, and it has the potential to realize remarkable features, e.g.,
  - Service Chaining
  - Traffic Engineering
  - Egress Peer Engineering

Interop Tokyo 2024 is scheduled from June 12 to 14.

See you in Tokyo!

SHOWNET Refine the Technologies