EANTC

EUROPEAN ADVANCED NETWORKING TEST CENTER

Multi-Vendor Interoperability Test Report SDN, Segment Routing, EVPN, and Time Synchronization

Spring 2024

MPLS SD&AI NET WORLD ★ 9/11 APRIL 24 | 25TH EDITION palais des congrès de paris

Table of Contents

Editor's Note

Carsten Rossenhövel, Co-Founder & CTO, EANTC

At the 25th anniversary of Upperside's World Congress, the interoperability event truly exceeded all of our expectations. The hot staging test was even more intense than usual, with 80 senior experts on-site from 14 participating vendors. More than 130 routers, switches, test tools, and other networking devices were installed and configured, taking a total of 400 rack units in the EANTC lab in Berlin, Germany, in February. All vendors collaborated seamlessly as a unified force, tirelessly creating a mind-boggling number of 1597 interoperability test results in just two weeks together on-site—after five months of joint in-depth planning.

The result is the biggest MPLS & SDN interoperability test report we have created so far. Meanwhile, I abstain from claiming records because the reports keep growing every year, thanks to the outstanding support from participating vendors. This report summarizes the "2024 State of the Union" for SDN, Segment Routing, EVPN, and Time Synchronization multi-vendor interoperability, including all leading network equipment manufacturers.

What are the lessons learned? From the bird's eye point of view, the participating solutions are solid and well interoperable regarding SDN, Segment Routing, EVPN services, and Time Synchronization. Progress has been made in both ways: a) More implementations from more vendors are interoperable than in the previous year, and b) more advanced standards and more options are now interoperable. These aspects confirm that customers have more choices for robust SDN, Segment Routing, EVPN services, and Time Synchronization deployment.

Reaching the next plateau each year is far from trivial: In a mature, standardized environment, each next level creates incremental implementation complexity: backwards compatiblity and correct functions of the more basic software must still be maintained. (For this reason, we always include basic regression tests.)

The most important takeaways of this year's test are:

- On the way towards Autonomous Networks (AN), live network performance monitoring is a mandatory key component for self-healing, self-optimizing networks. TWAMP and Seamless BFD tests were combined with SR policies to check for SR Policy Liveness for the first time by seven vendors.
- Segment Routing policies and Flexible Algorithms played a central role. Seven vendors supported SR traffic engineering steering per destination, and eight vendors supported FlexAlgo-based path calculation. For the first time, we evaluated "Exclude Affinity" policies, where paths can be excluded based on affinity attributes.
- SRv6 completed the move to compressed segment IDs (μSIDs). Now, all SRv6 tests used only μSIDs.
- Many test cases in SR-MPLS and SRv6 focused on advanced routing scenarios: Inter-Autonomous Systems, Multi-Homing, RT5, Global IP routing tables and route summarization, and TI-LFA redundancy; confirming maturity of versatile routing functions across all Segment Routing variants.
- Multicast received renewed interest, both with Bit-Indexed Explit Replication (BIER) and MVPN over Multicast Source Routing over IPv6 (MSR6).
- EVPN testing included a nearly ultimate collection of E-Line, E-LAN, and E-Tree service options tests, covering port-active redundancy, Integrated Routing and Bridging (IRB), Proxy services, MAC mobility, and multicast service.
- The interworking between SR-MPLS, SR-VXLAN, and SRv6 continues to evolve on SR and EVPN levels. These are important to avoid technology lock-ins.
- Time synchronization tests were dominated by the move to Class D (250 ns precision), which is required for Open RAN, next-gen 5G and 6G networks. The success rate of Class D clock testing increased tremendously, and long chains of Class D boundary clocks were successfully tested for the first time publicly. Vendors increasingly focused on production readiness validations (holdover using Enhanced SyncE, port monitoring, and boundary clock interworking).
- Multi-vendor interoperability of 400G ZR and ZR+ coherent pluggable optics were tested in context of Time Sync transport, and in conjunction with a DWDM system carrying 400G line rate.

- Finally, in the network management area, there was a promising number of test cases focusing on provisioning of network slices, L3VPN, and L2VPN. We also continued PCEP tests from previous years. For the purpose of autonomous network operations, service-aware optimization, PCEP association groups, service provisioning, and transport path comuptation are key elements. The vendors who were involved this year have shown consistent support over the past years and have made good progress, but we hope that the industry will adopt AN principles in SDN more widely in the future.

Overall, the test coverage was amazing and the vendor device participation and success rates have become excellent. Troubleshooting usually takes place when very complex scenarios are configured, which take last-level experts from vendors to get them right in the first place. When reading this report, just imagine (if you are a certified expert for any of the participating equipment) whether you would be able to configure these scenarios off the top of your head. It's not only a short-term fashion that the industry needs to move towards Autonomous Networks—it's a necessity for managing the complexity of advanced configurations.

Speaking of AI… All conferences are humming with presentations related to the topic of the year. It's quite complex and time-consuming to define standards, implement them across the industry, and prove the technical benefits in real multi-vendor scenarios. We will surely see network-level AI testing in future EANTC interop test events, but it will take some time.

But enough said about future endeavours: This is the intro to a super-extensive test report packed with hundreds of diagrams, tables, and thousands of first-hand, new test results. We hope that it will provide insightful takeaways regarding your respective interest areas, whatever aspect of transport networks you are focusing on as a vendor, service provider, other network operator, or simply for educational reasons!

Carsten Rossenhoevel, CTO & Co-Founder, EANTC

EANTC's Mission

Since 1991, EANTC has validated the interoperability, performance, robustness, and security of network solutions, platforms, and applications. Our goal is to provide vendor-neutral, objective assessments in a transparent and reproducible way. At Upperside's conferences, we have coordinated the MPLS and SDN interoperability testing of world-leading vendors since 2003. Our mission is to help the industry to validate interoperability through standards-compliance at the earliest feasible stage, and to ensure performance, scalability, and security before switching on production services. Our testing services help accelerate technology development and improve the stability of vendor solutions, lowering the operational risks.

Test Area Selection

The test areas were introduced by EANTC and subsequently discussed with participating vendors, aiming to encompass all aspects of service provider networks. Vendors contributed several new test cases; we are fortunate to get the attention of many IETF RFC and draft editors as part of the vendor team. In the end, the test plan is usually way too extensive; test cases are prioritized that receive implementation support from the largest number of vendors.

The EANTC team usually eliminates any test cases that are implemented only by a single vendor because our focus is on multi-vendor testing. There is only one exception: If multi-vendor testing of a previously confirmed test case is attempted but fails during the hot staging, and only one vendor remains that can demonstrate a working and standards-compliant implementation, we value that commitment and report the result.

Working Process

Preparations for the MPLS/SDN interoperability event began in September 2023. We initiated discussions about test areas and test case ideas with all interested vendors during several rounds of technical calls per technology area. In these calls, we thoroughly discussed test case details, new testing ideas, and the applicable (draft) standards, to ensure that the test plans reflect the latest industry developments.

The **Hot Staging Event** took place in Berlin in the second half of February. Newest hardware with latest software versions had already arrived at the EANTC lab from all over the world, waiting for the starting signal. Two weeks of non-stopping testing, deep and extensive on-site discussions, racing time to solve some emerged issues, resulted in great results for all our vendors.

EANTC engineers observed and verified all test combinations and results in detail, following the test procedures and pre-defined test steps. This test report contains only results that have been submitted consistently by each vendor participating in a test run, have proven and logged results, and have been verified by an EANTC test specialist assigned to the respective test area. We take this huge manual effort to avoid misinterpretations and false positives.

Participating Vendors and Devices

The tables on this page list all devices that vendors installed and tested with during the interoperability event. In some cases, there were multiple fixed configurations of the same product families tested - often, to explore different interface types or other hardware options. This explains the long equipment list for some vendors.

Participants	Devices
Arista	7050SX3 7280CR3A 7280R, 7280R2, 7280R3 7280R3E
Calnex	Paragon-neo SNE Ignite Sentinel Sentry
Ciena	5169 ELS Navigator NCS
Cisco	8011-4G24Y4H 8201-24H8FH ASR-9901, ASR-9902 Crosswork Network Controller N3K-C36180YC-R N540-24Q8L2DD N540-28Z4C N540X-12Z16G N540X-16Z4G8Q2C N9K-C93180YC-FX3 N9K-C93240YC-FX2 N9K-C93400LD-H1 N9K-C93600CD-GX NCS-57B1-6D24 NCS-57C1-48Q6
Ericsson	Router 6673 Router 6676 Router 6678
H3C	CR16010E-F S12500R-2L S12500R-48C6D S6850-56HF
highstreet technologies	ht.Connect
Huawei	ATN910C-G ATN910D-A NetEngine 8000 F8 NetEngine 8000 M8 NetEngine 8000 X4 iMaster NCE-IP
Juniper	ACX7024 ACX7100-32C, ACX7100-48L ACX7332 ACX7509 MX204, MX304 Native Cloud Router Paragon Applications PTX10001-36MR PTX10002-36QDD
Keysight	IxNetwork Time Sync Analyzer
Microchip	TimeProvider 4100
Nokia	Network Service Platform (NSP) 7250 IXR-e2 7750 SR-1
Ribbon	NPT 2300
ZTE	ZXR10 M6000-4SE ZXR10 M6000-8SE

Table 1: Participating Vendors and Devices

In some cases, vendors brought multiple units of each device type to parallelize some efforts (e.g., when a device was included in tests with different IGPs or SR versions). For this reason, the total number of devices (140 units) was much larger than the number of device types (63).

Interoperability Test Results

As usual, this test reports documents only positive results (passed test combinations) individually with vendor and device names. Failed test combinations are not mentioned in the diagrams; they are referenced anonymously in the report to describe the state of the industry. Our experience shows that participating vendors quickly proceed to solve interoperability issues after our test so there is no point in punishing them for their willingness to learn by testing. Confidentiality is vital to encourage manufacturers to participate with their latest - often beta - solutions and enables a safe environment in which to test and learn.

Terminology

We use the term "tested" when reporting on multi-vendor interoperability tests. The term "demonstrated" refers to scenarios where a service or protocol was evaluated with equipment from a single vendor only.

Test Equipment

We thank Calnex and Keysight for their test equipment and support throughout the testing.

IxNetwork from Keysight was used to generate traffic for all test areas, along with the following devices from Calnex and Keysight, which were specifically used for clock synchronization.

As in previous events, several Calnex instruments were used in the Time Synchronization test cases. Paragon-neo was used to generate and measure PTP and 1PPS signals with sub-nanosecond (ns) accuracy and 250 picosecond (ps) resolution, enabling characterization of devices to Class D clock and networks up to level 6Cm at line rates from 1GbE to 400GbE.

Calnex SNE Ignite was used to insert delay for link asymmetry-based testing. With its integrated transparent clock function, impairments were applied in a timing-aware network without causing sync issues, allowing configurable non-ideal conditions to be created as required.

Calnex Sentry was again used for network tests to measure up to four 1PPS signals simultaneously, enabling synchronization to be monitored across a network or multiple tests to be run simultaneously.

In the O-RAN tests, measurement of the RF OTA signal from an O-RU using the Calnex Sentinel allowed the complete end-to-end sync functionality and performance to be evaluated, as well as PTP and 1PPS from a network node or the end clock.

Keysight participated with the Time Sync Analyzer (TSA), a scalable multiport clock quality test platform. It was used to generate and measure PTP, SyncE, and 1PPS signals up to Class D. TSA allows up to six PTP (and SyncE) measurements and up to four 1PPS measurements concurrently for monitoring and comparing synchronization performance across network clock chain.

EVPN Test Results

Ethernet VPN (EVPN) is an advanced networking technology that helps service providers and enterprises extend their local Layer 2 network services across the WAN. At its core, EVPN leverages the familiarity and ubiquity of Ethernet technology, extending its capabilities to create a flexible and dynamic virtual network environment. EVPN supports both Layer 2 and Layer 3 services, making operating and optimizing resource utilization easy. The technology is not limited to specific types of traffic or applications, providing a comprehensive and seamless solution for connecting diverse resources across distributed locations.

EVPN utilizes the Border Gateway Protocol (BGP)-based control plane, providing a robust foundation for managing large-scale networks while ensuring efficient and reliable communication between sites. At the transport layer, EVPN can use several data plane protocols, such as MPLS, SRv6, SR-MPLS, or VXLAN, to encapsulate and transport the traffic over a shared physical infrastructure.

Aligned with the evolution of EVPN service requirements in the industry and new standards defined in the IETF, we evolve the test coverage at each of our annual multi-vendor interoperability test events. We maintain several basic tests, validating the foundation of EVPN services with new participating implementations or functioning as a regression test for vendors that have participated previously. Beyond the basics, we aim to focus on newly standardized EVPN functionalities or ones supported by more vendors than in the previous interop event.

We built the typical spine-leaf architecture for our test. The spine serves as Route Reflector (RR), too. Arista 7280R3 and Cisco IOS XRd provided RR services for most of the SR-MPLS testbed, and Arista 7280R was the RR for the VXLAN testbed. In some test cases, a customer edge router (CE) was required to terminate services with dynamic routing. For the SR-MPLS testbed, the Arista 7280R3 took the CE role, and an Arista 7280R

served as CE for the VXLAN testbed. The traffic generator was always Keysight IxNetwork whenever traffic had to be generated for test result verification.

In the SR-MPLS area, we introduced the following new test cases this year:

- Preference-based Designated Forwarder (DF) election and port-active redundancy
- L2 attributes extended community
- Weighted multipath
- IGMP proxy
- IP prefix route resolution to gateway IP.

In the VXLAN area, the following new test cases were added:

- Preference-based DF election
- Interconnect Solution for EVPN Overlay and Multi-Site Solution for EVPN Overlay interworking

E-Line Test

E-Line is a traditional point-to-point service that can be implemented as an EVPN service type. RFC 8214 introduces the support of Virtual Private Wire Service (VPWS) in EVPN. EVPN provides the Border Gateway Protocol (BGP) control plane and multi-homing single-active or all-active redundancy to VPWS service, making VPWS more robust and scalable.

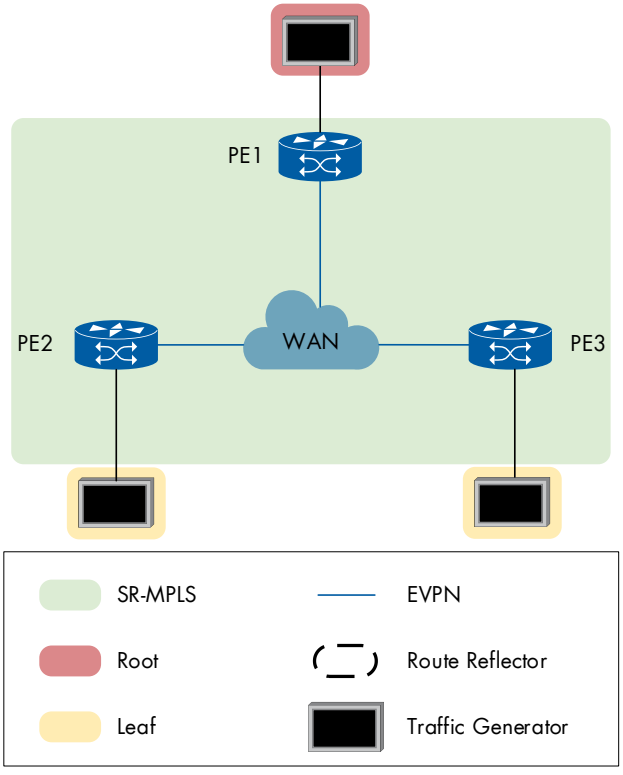


Figure 1: E-Line All-Active Multihoming

In our test, we used the multi-homing all-active mode and generated unicast traffic. We confirmed that the traffic was initially balanced between the multi-homing PEs. After that, we performed a failover by temporarily disabling one of the interfaces between PE and CE. The test was run sequentially with both links. Finally, we reactivated both links to verify the multihoming was functioning correctly.

There were four test combinations of PEs shown in Figure 1. The following devices participated successfully in the respective roles.

First combination:

- PE1: Ciena 5169
- PE2: Ribbon NPT 2300
- PE3: H3C CR16010E-F

Second combination:

- PE1: Arista 7280R3
- PE2: Nokia 7750 SR-1
- PE3: Cisco NCS-57C1-48Q6
- PE4: Huawei NetEngine 8000 M8

Third combination:

- PE1: Ciena 5169
- PE2: Ribbon NPT 2300
- PE3: Juniper MX304
- PE4: H3C CR16010E-F

Fourth combination:

- PE1: Arista 7280R3
- PE2: H3C S12500R-2L
- PE3: Huawei NetEngine 8000 M8
- PE4: Juniper MX304

E-Tree Test

E-Tree is a Layer 2 service that enables rooted-multipoint connections, for example, between the company's headquarters and its branches. EVPN E-Tree in RFC 8317 inherits the rooted-multipoint service feature and utilizes the EVPN BGP control plane to offer more flexibility and redundancy features.

We ran a test with bi-directional full-mesh unicast traffic three times. Each device was assigned a specific role, either as a root or leaf, in each run. Our findings indicated that the traffic successfully passed between the root and the leaves, while no traffic passed between the leaves as intended.

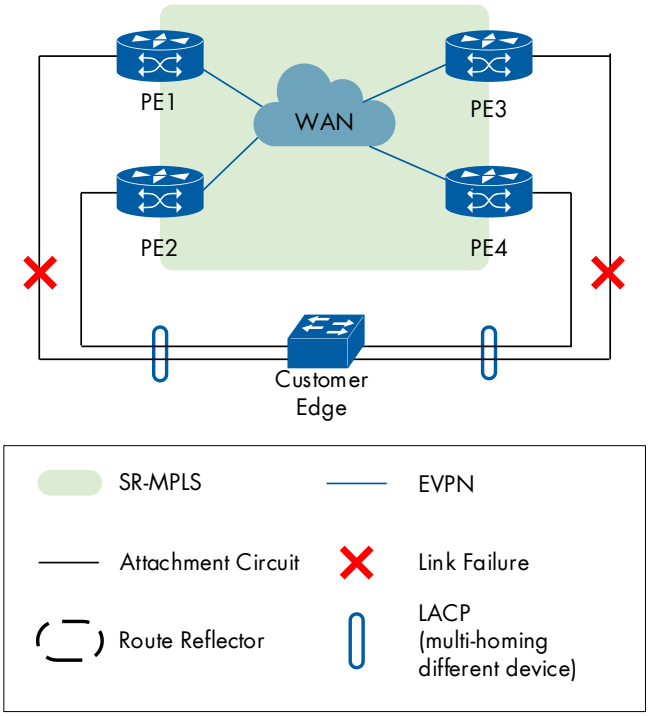


Figure 2: E-Tree Test Topology

There were three test combinations of PEs shown in Figure 2. The following devices participated successfully in the respective roles.

First combination:

- PE1 (Root): Juniper MX304
- PE2 (Leaf-Attached): Huawei NetEngine 8000 M8
- PE3 (Leaf-Attached): Arista 7280R3

Second combination:

- PE1 (Root): Huawei NetEngine 8000 M8
- PE2 (Leaf-Attached): Juniper MX304
- PE3 (Leaf-Attached): Arista 7280R3

Third combination:

- PE1 (Root): Arista 7280R3
- PE2 (Leaf-Attached): Huawei NetEngine 8000 M8
- PE3 (Leaf-Attached): Juniper MX304

E-LAN Test

E-LAN is the third of the classic Layer 2 VPN services, providing a multipoint-to-multipoint virtual LAN service. When implemented as an EVPN service type, it uses BGP as the control plane and all the multihoming functions of EVPN.

We performed three test runs with different combinations and traffic profiles in our test. In the first run, we tested with bi-directional unicast traffic and five devices multihoming under all-active mode. We also performed a link failure to simulate the switchover. All devices worked properly in this test.

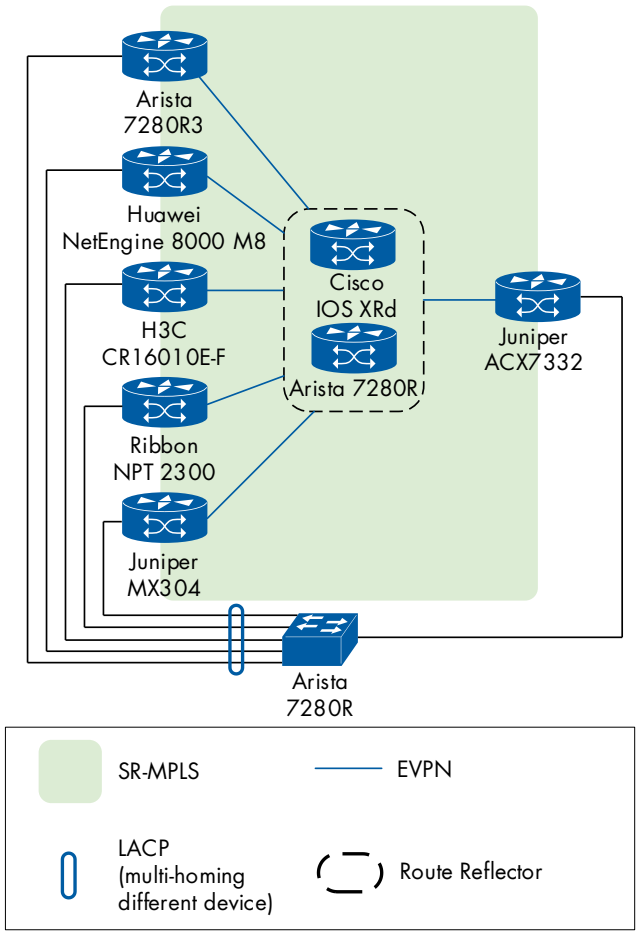


Figure 3: E-LAN All-Active Multihoming—Unicast

The following devices participated successfully:

- Single homed Provider Edge (PE): Juniper ACX7332
- Multihomed Provider Edge (PE): Arista 7280R3, Huawei NetEngine 8000 M8, H3C CR16010E-F, Ribbon NPT 2300, Juniper MX304

Then, we performed another two runs with bi-directional broadcast traffic and different device multihoming combinations. We confirmed that only one broadcast copy was received on the remote endpoint.

Devices participating in the first run (Figure 4):

- Single homed PE: Ciena 5169
- Multihomed PE: Arista 7280R3, H3C CR16010E-F, Ribbon NPT 2300

Devices participating in the second run (Figure 5):

- Single homed PE: Juniper MX304
- Multi-homed PE: Arista 7280R3, Juniper ACX7332

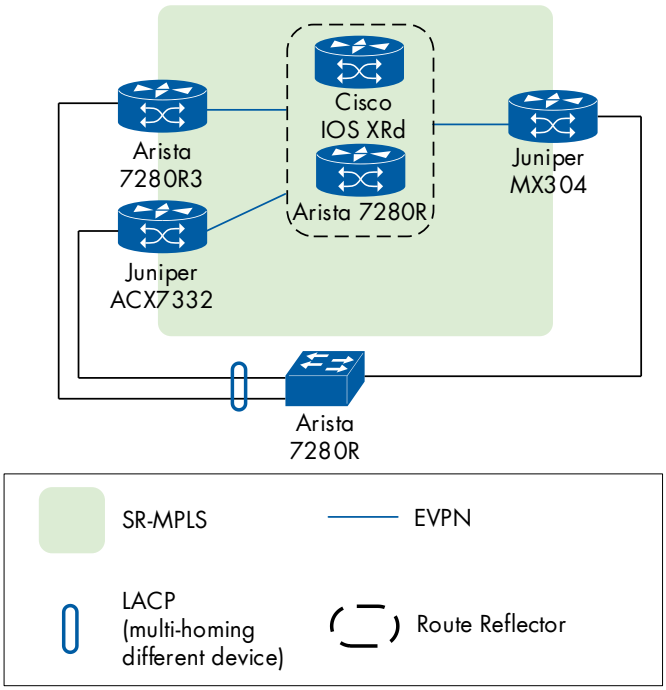


Figure 4: E-LAN All-Active Multihoming, Broadcast Combination 1

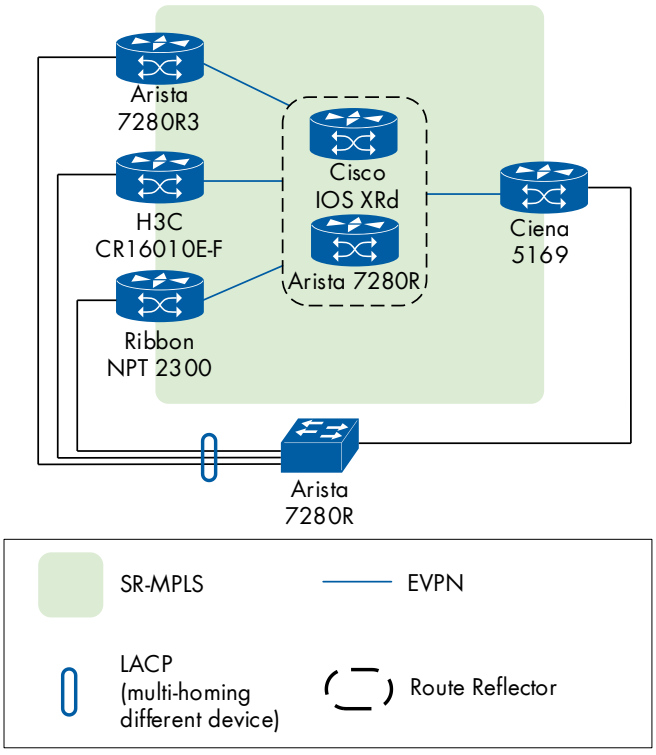


Figure 5: E-LAN All-Active Multihoming, Broadcast Combination 2

Flexible Cross-Connect Service

The Flexible Cross-Connect (FXC) solution (draft-ietf-bess-evpn-vpws-fxc) provides a flexible solution for bundling multiple attachment circuits across various Ethernet segments and physical interfaces into a single EVPN VPWS service tunnel. It is admirable that this feature still maintains Single-Active and All-Active multi-homing offered by the EVPN BGP control plane.

We conducted a test with two bi-directional unicast in two different VLANs and used one Pseudowire (PW) in the EVPN core to transport the data. BGP was used for MAC learning, and the traffic flowed seamlessly as expected. We had three test runs in total, including multihoming and singlehoming. With the multihoming setup, we observed the traffic was balanced between PEs. We also simulated failover by disabling and enabling the interface between PE and CE in the multihoming setup, and the behavior was as expected.

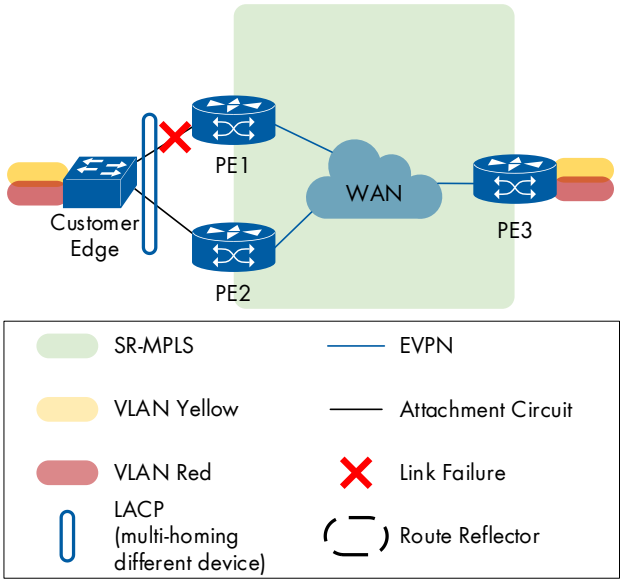


Figure 6: Flexible Cross-Connect Service

The following devices participated successfully.

First test combination:

- PE1: Arista 7280R3
- PE2: None
- PE3: Juniper ACX7332

Second test combination:

- PE1: Arista 7280R3
- PE2: None
- PE3: Cisco N540-24Q8L2DD

Third test combination:

- PE1: Arista 7280R3
- PE2: Juniper ACX7332
- PE3: Cisco N540-24Q8L2DD

EVPN-VPWS with Pseudowires (PWs) virtual Ethernet segment (vES)

In the EVPN network, the physical link is typically the preferred choice when building an Ethernet Segment (ES). However, in some cases, the complexities of the network environment may make it less optimal. To provide greater flexibility in network setup, a virtual Ethernet segment (vES) has been introduced in draft-ietf-bess-evpn-virtual-eth-segment. This can be created using a set of Ethernet Virtual Circuits (EVCs), such as VLANs, MPLS Label Switch Paths (LSPs), or Pseudowires (PWs).

We used PWs to build a port-active redundancy group with two PWs to two PEs with the EVPN-VPWS service on top of it. We then sent bi-directional unicast traffic and performed a switchover between the two PWs. The result was in line with our expectations, and the test was successful.

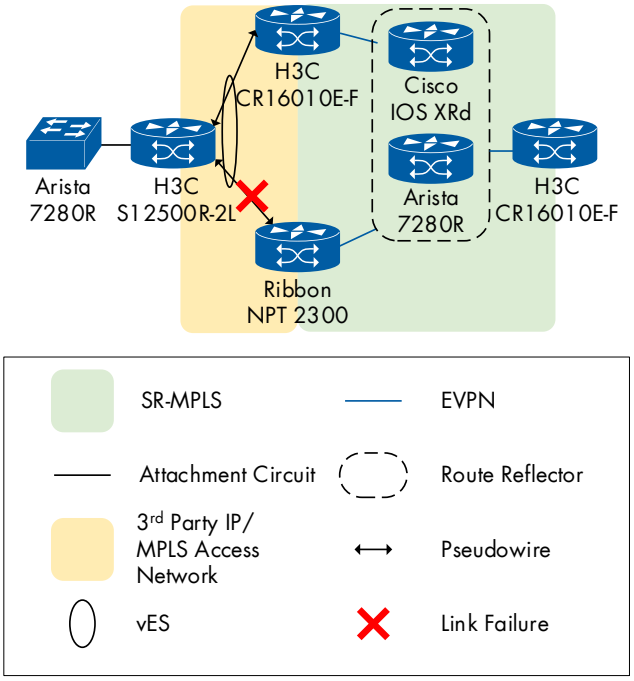


Figure 7: EVPN-VPWS with PWs vES

The following devices participated successfully:

- Single homed PE: H3C S12500R-2L, H3C CR16010E-F
- Multihomed PE: H3C CR16010E-F, Ribbon NPT 2300

Proxy MAC-IP Advertisement

In implementing multihoming, it is essential to note that the Customer Edge (CE) device balances the traffic between multiple Provider Edge (PE) routers using various mechanisms. However, it has been observed that this can sometimes result in the MAC learning not being synchronized between the multihomed Pes. To address this issue, the proxy MAC/IP advertisement in draft-rbickhart-evpn-ip-mac-proxy-adv is designed specifically for such scenarios.

During our testing, we confirmed that all Devices Under Test (DUTs) sent out a MAC/IP route with a "P" flag set when they were not the direct recipients of the MAC addresses. We then performed a link shutdown to verify that both DUTs were able to send and receive the proxy MAC/IP advertisement message, ensuring smooth and uninterrupted communication between the devices.

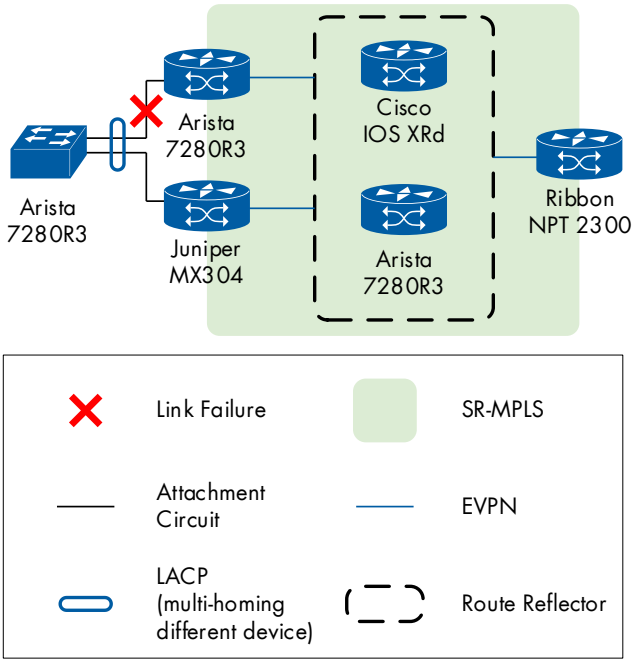


Figure 8: Proxy MAC-IP Advertisement

The following devices participated successfully:

- Single homed PE: Ribbon NPT 2300
- Multihomed PE: Arista 7280R3, Juniper MX304

Preference-based EVPN DF Election

In EVPN networks, the Designated Forwarder (DF) is crucial in forwarding Broadcast, Unknown unicast, and Multicast (BUM) traffic within an Ethernet Segment (ES). By default, the DF is selected based on a modular-based Election algorithm that efficiently handles different Ethernet Tags in the ES. However, circumstances require a more deterministic and user-controlled approach, such as during regular maintenance or software upgrades. This is where the preference-based DF can be helpful. It is introduced in draft-ietf-bess-evpn-pref-df. The election is based on the value we configure on the interface, allowing us to control the DF by configuration rather than by link failure or other unexpected behavior.

We conducted a test using the "Highest-Preference algorithm" and "Don't preempt" disabled. We used single-active and unicast traffic with EVPN-VPWS as a service during the test. We observed the switchover during the test to verify that the DF was working as expected. Our test confirmed that the preference-based DF election and the settings we used worked correctly.

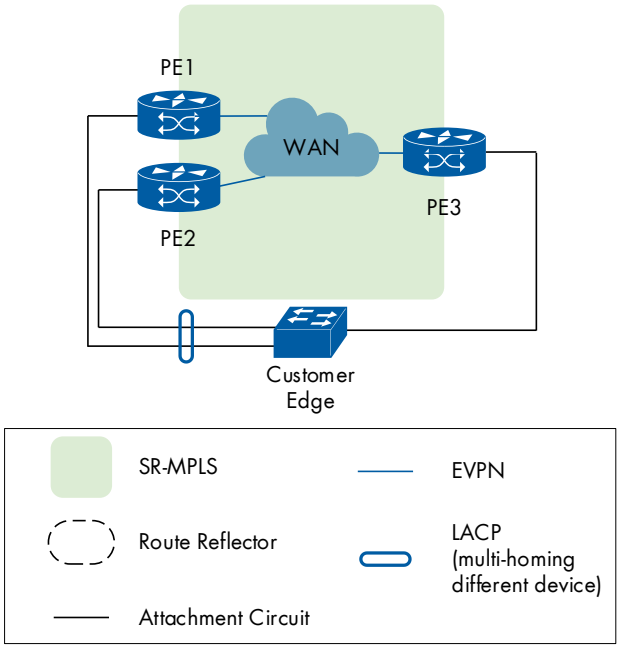


Figure 9: Preference-Based EVPN DF Election, SR-MPLS

See Table 2 for a list of successful device combinations.

We performed the same test on the VXLAN testbed as well. In the VXLAN testbed, we verified that a DF is elected based on the preference algorithm in an all-active multi-homing scenario. After changing the preference of the DF to a lower preference, the new DF election has signed the new DF to the one with the higher preference value.

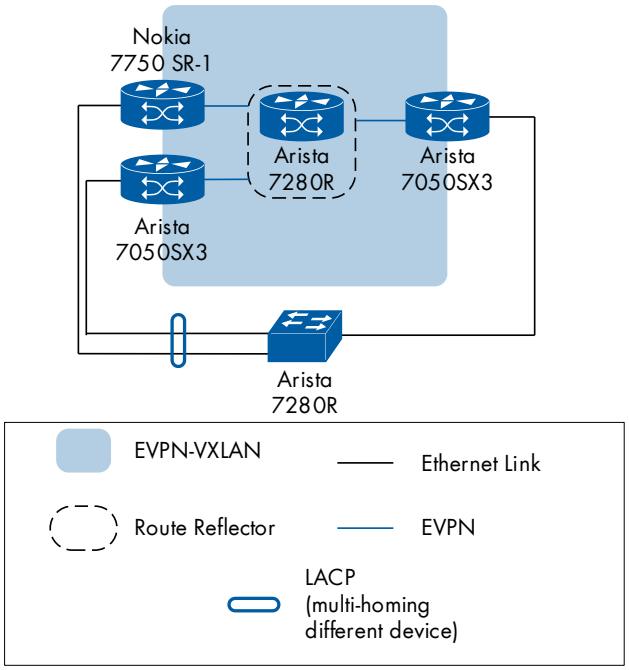


Figure 10: Preference-based EVPN DF Election, VXLAN

The following devices participated successfully:

- Single homed PE: Arista 7050SX3
- Multihomed PE: Arista 7050SX3, Nokia 7750 SR-1

EVPN Port-Active Redundancy

The Port-Active redundancy mode in draft-ietf-bess-evpn-mh-pa is a highly dependable system that adheres to open standards and is fully compatible with RFC 7432. One of its key advantages is its ability to work with any underlying technologies and services, making it a versatile solution. Additionally, it supports various Designated Forwarder (DF) election algorithms, including modulo, HRW, preference, and others.

Juniper ACX7332	Nokia 7750 SR-1	Huawei NetEngine 8000 M8
Ciena 5169	Juniper ACX7332	Juniper ACX7509
Ribbon NPT 2300	Juniper ACX7332	Juniper ACX7509
Huawei NetEngine 8000 M8	Cisco 8201-24H8FH	H3C CR16010E-F
Arista 7280R3	Cisco 8201-24H8FH	H3C CR16010E-F
Arista 7280R3	Juniper ACX7509	H3C CR16010E-F
H3C CR16010E-F	Ribbon NPT 2300	H3C CR16010E-F
H3C S12500R-2L	Ribbon NPT 2300	H3C CR16010E-F

Table 2: Preference-based EVPN DF Election, SR-MPLS DUT combinations

We used EVPN-VPWS as the service and a preference-based DF algorithm during our testing. We tested the system with one active and one standby port and executed the switch-over by configuring a higher preference number. We sent traffic simultaneously to confirm that the switch-over was executed correctly.

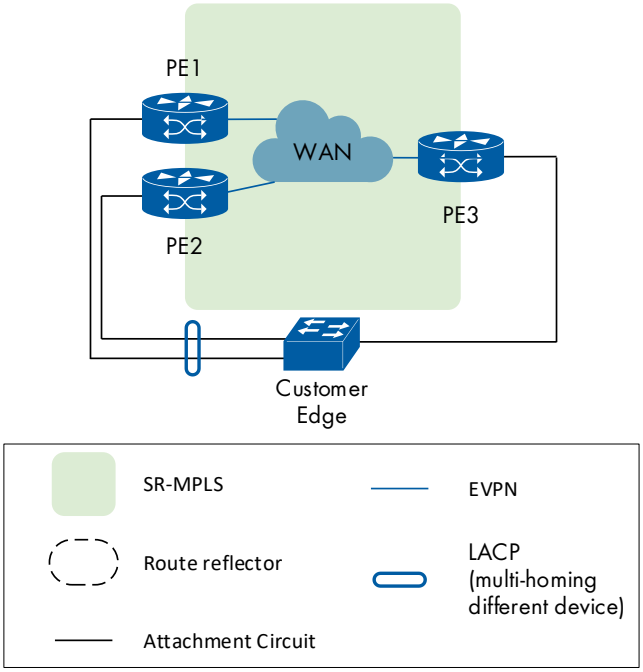


Figure 11: EVPN Port-Active Redundancy

The following devices participated successfully:

First iteration:

- PE1: Ciena 5169
- PE2: Ribbon NPT 2300
- PE3: H3C CR16010E-F

Second iteration:

- PE1: Ciena 5169
- PE2: Juniper ACX7332
- PE3: H3C CR16010E-F

Third iteration:

- PE1: Cisco 8201-24H8FH
- PE2: Huawei NetEngine 8000 M8
- PE3: H3C CR16010E-F

EVPN Layer 2 Attributes Extended Community

The new RFC7432bis draft defined a new EVPN Layer2 Attributes Extended Community. This community defines the attributes of Maximum Transmission Unit (MTU), Control Word (CW), and flow label, which are all fundamental Layer 2 attributes that can enhance L2 fault tolerance.

Our test confirmed that the BGP routes for EVPN RT-1 and RT-3 on the DUT were correct. Both EVPN-VPWS and EVPN-ELAN services had MTU, CW, and Flow Label enabled. Additionally, we verified that the host connected to the service was reachable via ICMP ping packets when the MTU, CW, and flow labels were set.

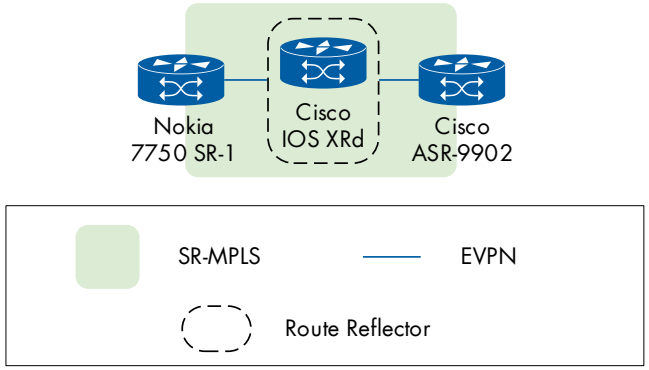


Figure 12: EVPN Layer 2 Attributes Extended Community

Weighted Multipath Procedures for EVPN Multi-Homing

RFC 7432 defines equal bandwidth distribution between CE and egress PEs, leading to equal load balancing of remote traffic. However, this can be limiting when adding/removing links or when there are link failures. To address this, a new EVPN Link Bandwidth extended community is introduced in draft-ietf-bess-evpn-unequal-lb, providing greater flexibility.

During the test, we verified the value unit 0x01, which indicates the weight of the link rather than its bandwidth. We had three links on one DUT and two on another. The weight number was 3 and 2, respectively, when all the links were up. We shut down one link on the three-link DUT, and a BGP update message was sent out, updating the weight to 2. Similarly, when we shut down one link on the 2-link DUT, the weight was updated to 1. Once we recovered the link, the weight values became 3 and 2, respectively.

The test topology is shown in Figure 13 on the following page.

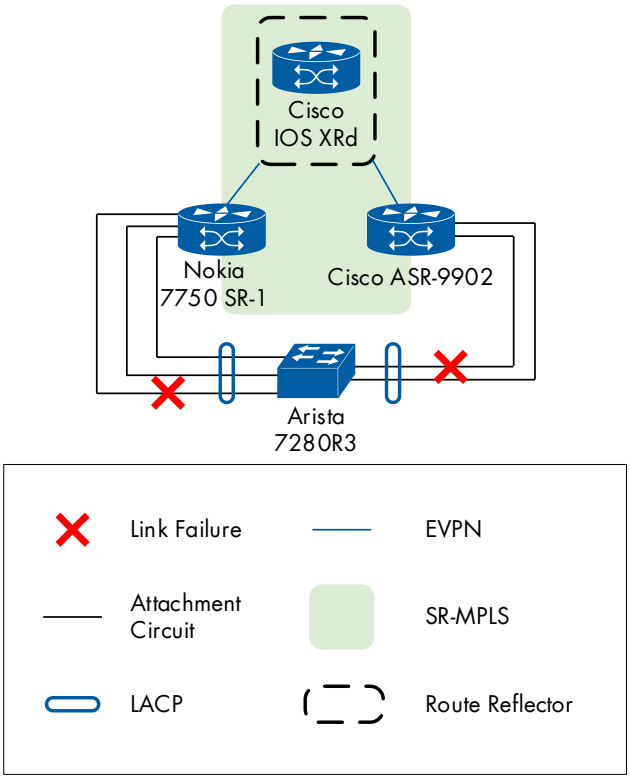


Figure 13: Weighted Multipath Procedures for EVPN Multi-Homing

IP Prefix Resolution to Gateway IP

Section 9.2 of RFC 9135 introduces a new use case for inter-subnet forwarding, which is achieved by using EVPN RT-5 to advertise the subnet behind a Tenant System (TS) and performing recursive route resolution to resolve destination endpoints.

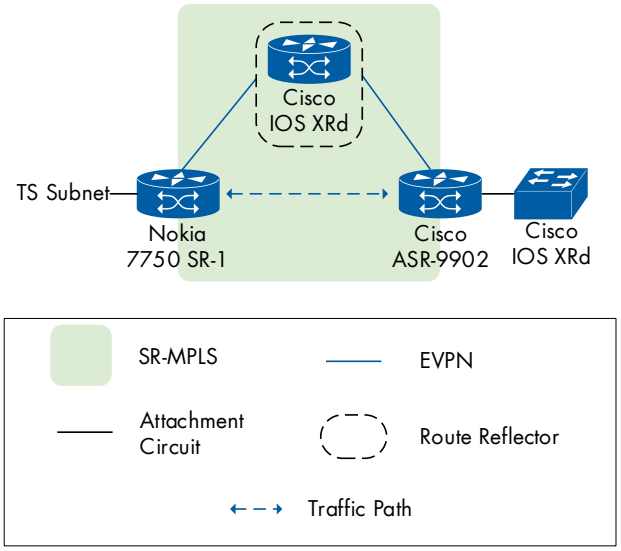


Figure 14: IP Prefix Resolution to Gateway IP

During our testing, a TS subnet behind the left side PE was advertised with an RT-5 destined to its own gateway IP address. The route server then advertised it to the right site PE, and the right site PE performed recursive route resolution of the gateway IP it received. As a result, the right side PE directly sent traffic destined for the subnet behind the left side PE without involving the route server. It makes the route server a pure control plane and the efficiency forwarding path on the data plane without route server involvement.

EVPN IGMP-Proxy

RFC 9251 defines IGMP/MLD proxy for the EVPN network. It utilizes RT-6, RT-7, and RT-8 to effectively manage multicast traffic. RT-6 is responsible for selective multicast forwarding, while RT-7 and RT-8 handle the IGMP/MLD join/leave message synchronization issue for multihoming multicast forwarders. IGMP proxy has been tested for years in the VXLAN testbed, but it's the first time we tested it in the SR-MPLS testbed.

In our testing scenario, the right-side DUTs were under all-active multihoming and acted as the multicast traffic receiver. We sent an IGMPv3 join message from a VM connected to the right side and observed that RT-6 and RT-7 were generated, and they appeared in both right-side DUT's EVPN route table. Additionally, we noted that two RT-6s appeared on the left-side DUT's route table. We then sent pings from the host to the multicast group we had just joined, and the pings were successful. Finally, we sent an IGMPv3 leave message from the same VM and observed that RT-8 was displayed on both right-side DUTs.

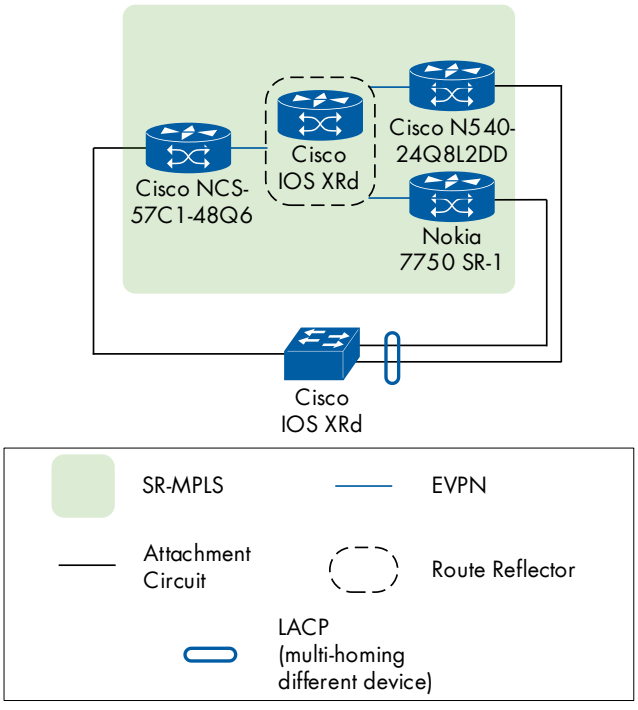


Figure 15: EVPN IGMP-Proxy

Integrated Routing and Bridging (IRB) Section

Symmetric IRB

When EVPN is deployed on a large scale, it becomes important to have both bridging (within the same subnet) and routing (between different subnets) functions in place. Two solutions were created to address this challenge: RFC 9135 defines the Integrated Routing and Bridging (IRB) solution, and RFC 9136 further improves IRB with the IP Prefix Advertisement (RT-5) solution.

We defined two service interfaces in RFC 7432 to test these solutions: VLAN-based and VLAN-aware bundle service interfaces. We then sent full mesh traffic to all DUTs involved in each scenario to ensure that IRB works between everyone in the same topology.

VLAN-Based service enables one-to-one mapping of a single bridged domain to a single bridged domain. Each VLAN is associated with a single EVPN Instance (EVI), resulting in a separate bridge table for each VLAN.

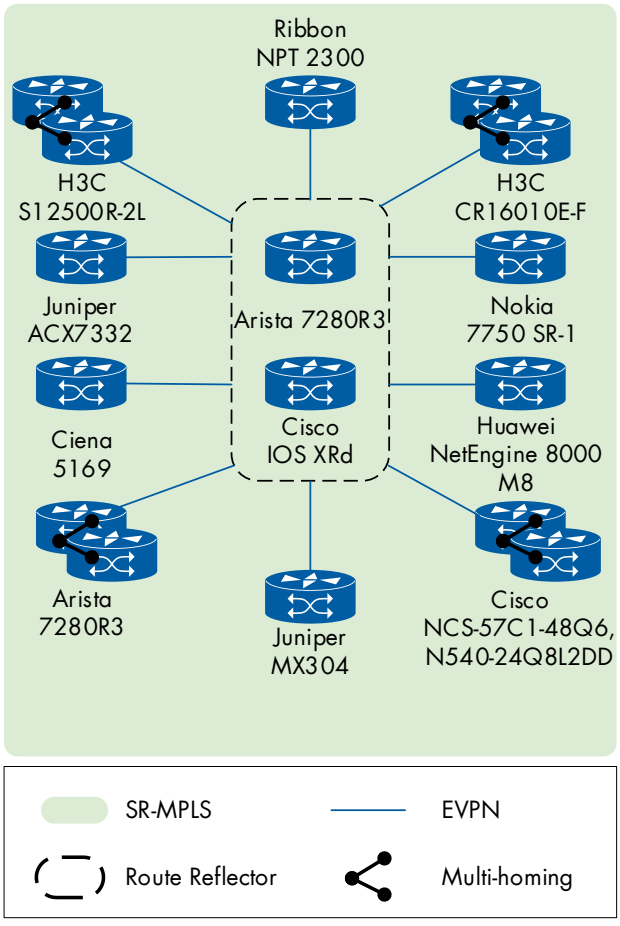


Figure 16: Symmetric IRB, SR-MPLS, VLAN-Based

With the VLAN-Aware bundle service interface, an EVPN instance corresponds to multiple broadcast domains (e.g., multiple VLANs) with each VLAN having its own bridge table, which means multiple bridge tables (one per VLAN) are maintained by a single MAC-VRF corresponding to the EVI.

Figure 16 contains the combinations successfully validated in the SR-MPLS testbed.

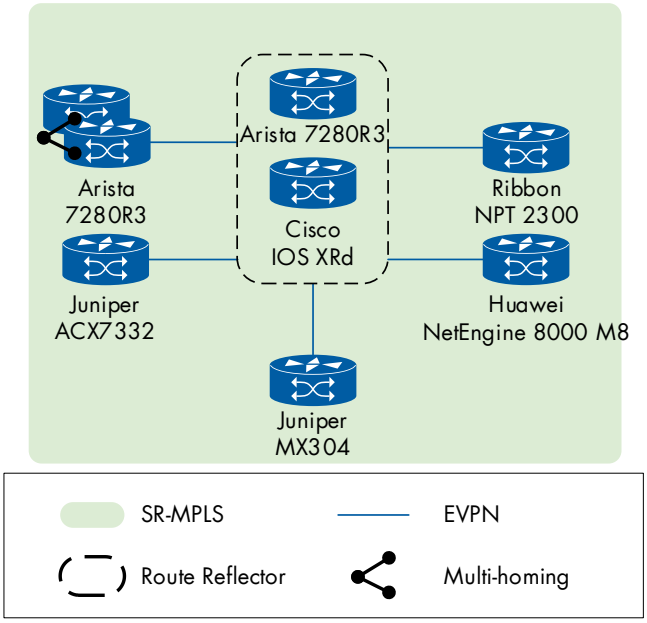


Figure 17: Symmetric IRB, SR-MPLS, VLAN-Aware Bundle Combination 1

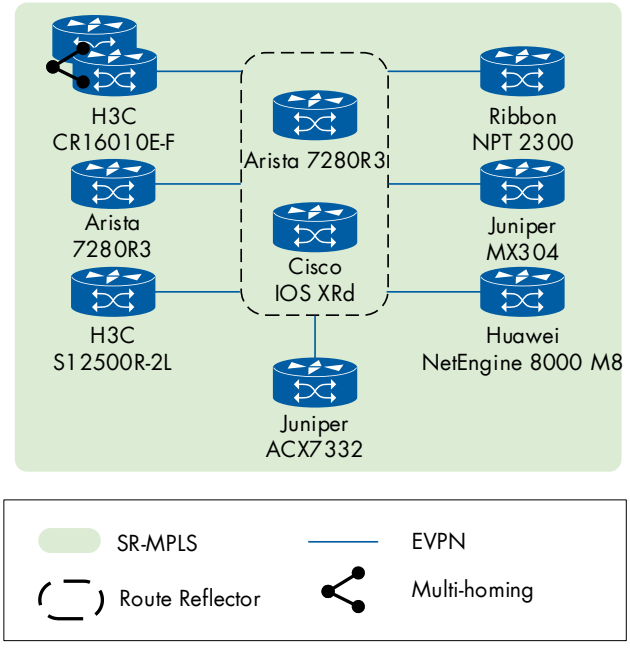


Figure 18: Symmetric IRB, SR-MPLS, VLAN-Aware Bundle Combination 2

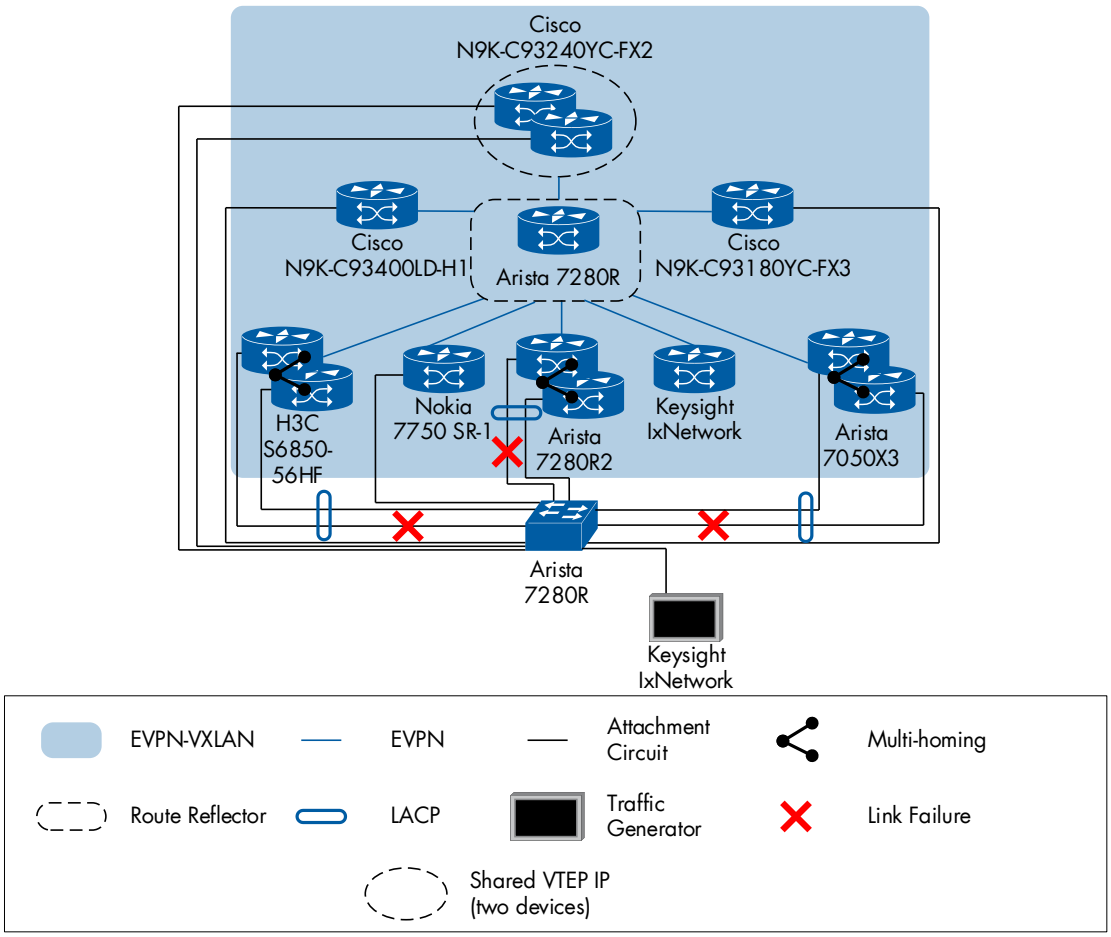


Figure 19: Symmetric IRB, VXLAN, VLAN-Based

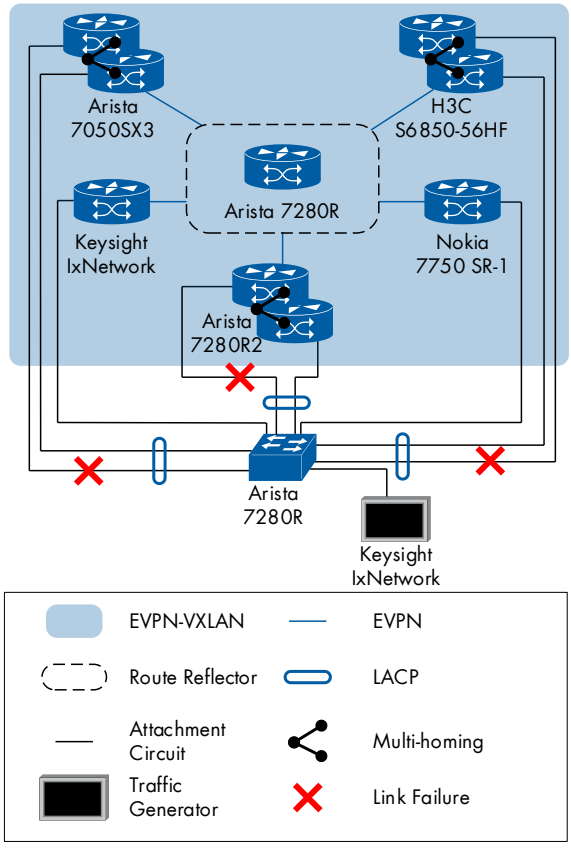


Figure 20: Symmetric IRB, VXLAN, VLAN-Aware Bundle

We performed the same test in the VXLAN testbed.

In the VXLAN test, we verified the Symmetric IRB interoperability with VLAN-Based and VLAN-Aware Bundle service and redundancy. We sent bidirectional inter- and intra-subnet unicast traffic and observed no traffic loss, with no failover. We also checked the traffic load balance between the all-active multihomed PEs. Redundancy was proven for the multihomed PEs with a failover. We noticed some traffic loss during the failover, as expected, but no more traffic loss after the failover.

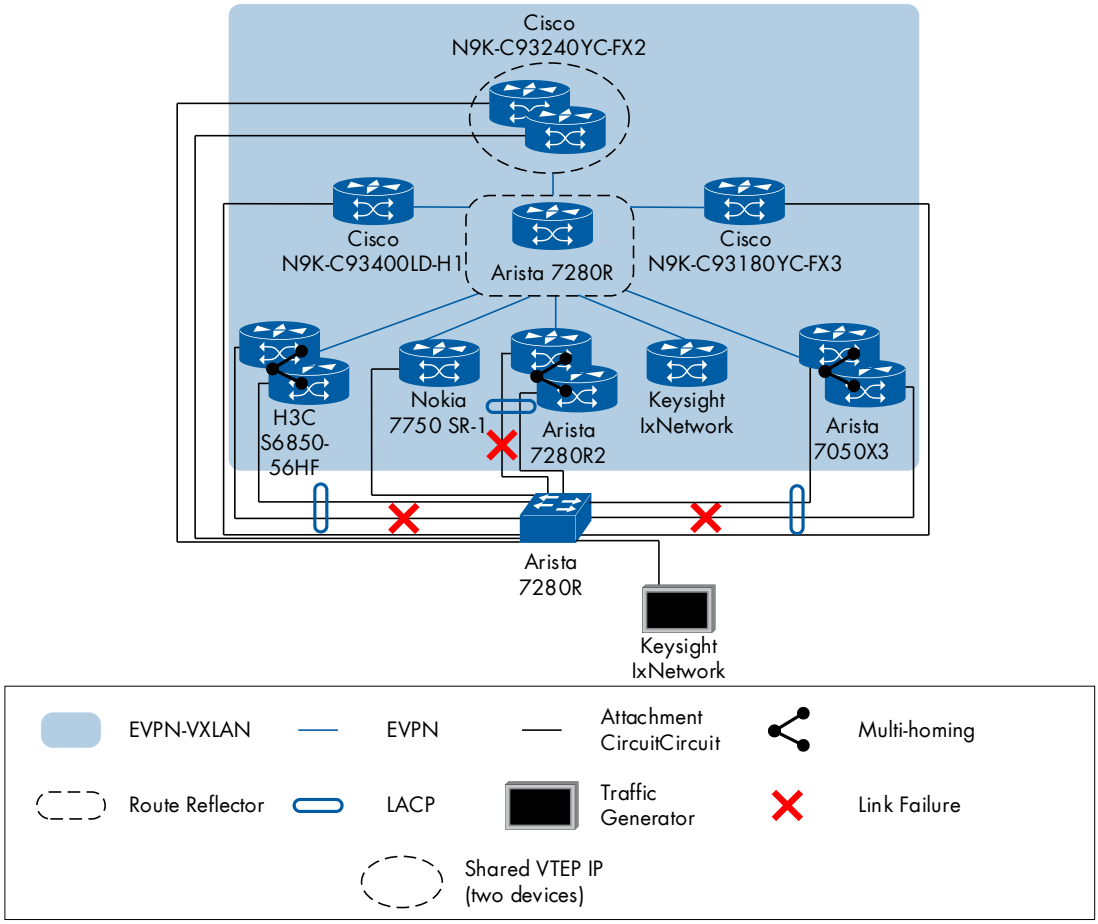


Figure 21: Asymmetric IRB, VXLAN, VLAN-based (identical to Figure 19)

Asymmetric IRB

We verified the interoperability of Asymmetric Integrated Routing and Bridging (IRB). Compared to Symmetric IRB, Asymmetric IRB requires both IP and MAC lookups at the ingress Network Virtualization Edge (NVE), while only a MAC lookup is needed at the egress NVE. However, under Symmetric IRB mode, both IP and MAC lookups are required at the ingress and egress NVE. Asymmetric IRB may have scaling issues in a network with a large number of Broadcast Domains (BDs) and Supplementary Broadcast Domains (SBDs). However, using it in a smaller network provides lower latency and more straightforward configuration.

In this test, we verified the Asymmetric IRB under VLAN-Based and VLAN-Aware Bundle services. We sent bi-directional inter- and intra-subnet unicast traffic to confirm that there was no packet loss and that bridging and routing were functioning as expected. The multi-homed Provider Edges (PEs) were able to balance the traffic load, and redundancy was confirmed by simulating link failures at one of the PEs in each Ethernet Segment (ES) if they are multihomed.

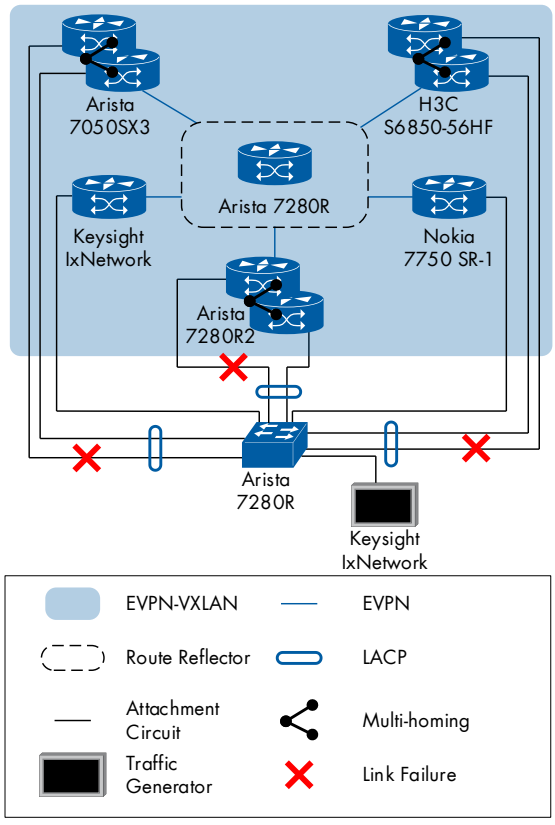


Figure 22: Asymmetric IRB, VXLAN, VLAN-Aware Bundle

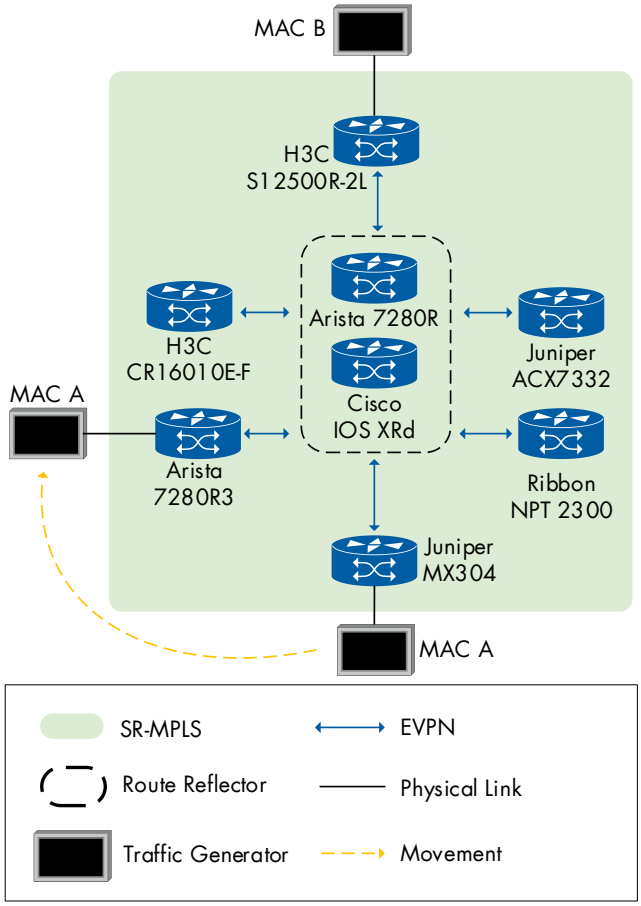


Figure 23: MAC Mobility, SR-MPLS

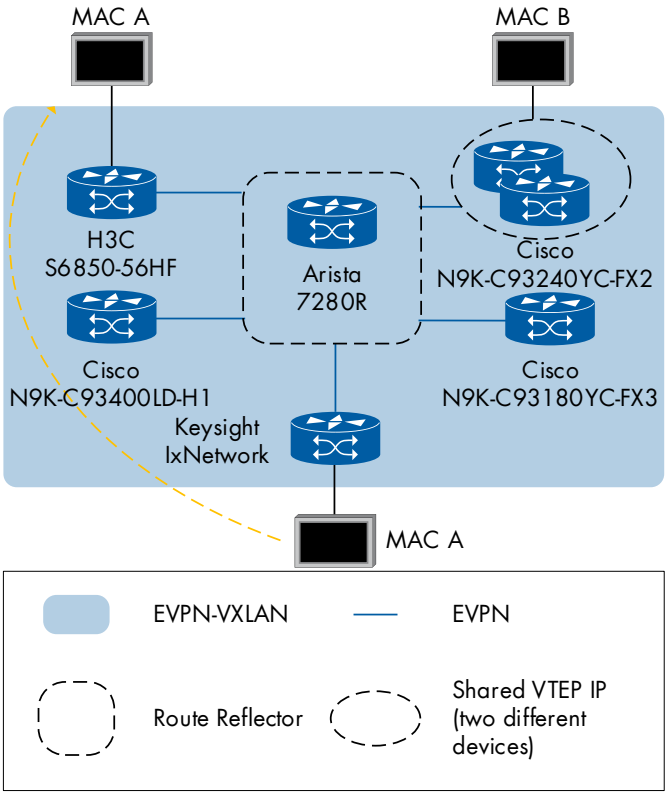


Figure 24: MAC Mobility, VXLAN

MAC Mobility

In today's data center networks, it is necessary to move VMs/tenants frequently due to maintenance, deployment changes, and other requirements. For this reason, MAC mobility is vital in the EVPN area.

In our test, we thoroughly verified that a mobile MAC address was successfully moved between DUTs, and the sequence number in RT-2 was increased as specified in RFC 7432. We sent bi-directional unicast traffic throughout the testing process, which was forwarded to the latest MAC location as expected. We performed the same test on both SR-MPLS and VXLAN testbeds.

EVPN-VXLAN to EVPN VXLAN Tunnel Stitching for DCI

In large EVPN-VXLAN fabrics, it is important to effectively manage the number of VXLAN tunnels between leaf devices in data centers and between data centers to avoid the overwhelming capacity of the gateway devices. One solution that has been found to be effective in optimizing the number of VXLAN tunnels between two data centers is the VXLAN to VXLAN stitching solution.

During our test, we focused on the "Integrated interconnect solution," which integrates the NVE Gateway and WAN Edge functions into a single system. To ensure end-to-end reachability, we generated bridging and routing unicast traffic at the same time, and we were able to observe zero packet loss. To avoid any potential loops, we configured the D-path. Moreover, we ensured redundancy by using all-active multihoming. To this end, we performed link failure and recovery tests, which were successful. The load balancing for multihomed devices also worked as expected.

(See Figure 25 on the next page for the test topology)

Interconnect Solution for EVPN Overlay and Multi-Site Solution for EVPN Overlay Interworking

Modern data centers must cater to various customer requirements and ensure redundancy across different sites. As a result, the current data center is exploring the use of IP-only networks in the WAN area, which offers a cost-effective and streamlined solution compared to traditional DCI technologies like MPLS/VPLS.

In this test, we validated bridged and routed data plane traffic between EVPN Domain 1 and an All-Active EVPN GW based on RFC 9014 and EVPN Domain 2 with an EVPN GW based on the Sharma-

draft ("draft-sharma-bess-multi-site-evpn"). To ensure gateway resiliency, the RFC 9014 Gateways were interconnected using an I-ESI to provide All-Active multi-homing. The EVPN gateways based on the Sharma-draft ("draft-sharma-bess-multi-site-evpn") provided resiliency via overlay ECMP. We verified the implementation by sending end-to-end unicast traffic and observed no packet loss. (Test topology in Figure 26)

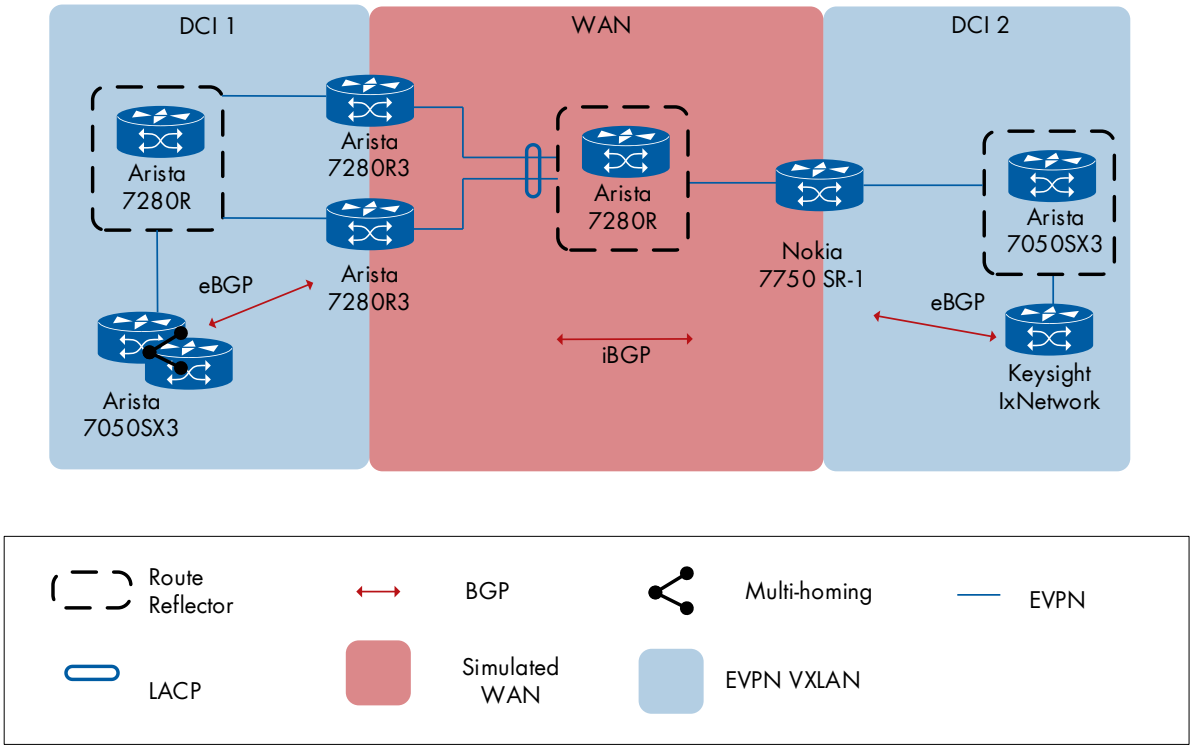


Figure 25: EVPN-VXLAN to EVPN VXLAN Tunnel Stitching for DCI

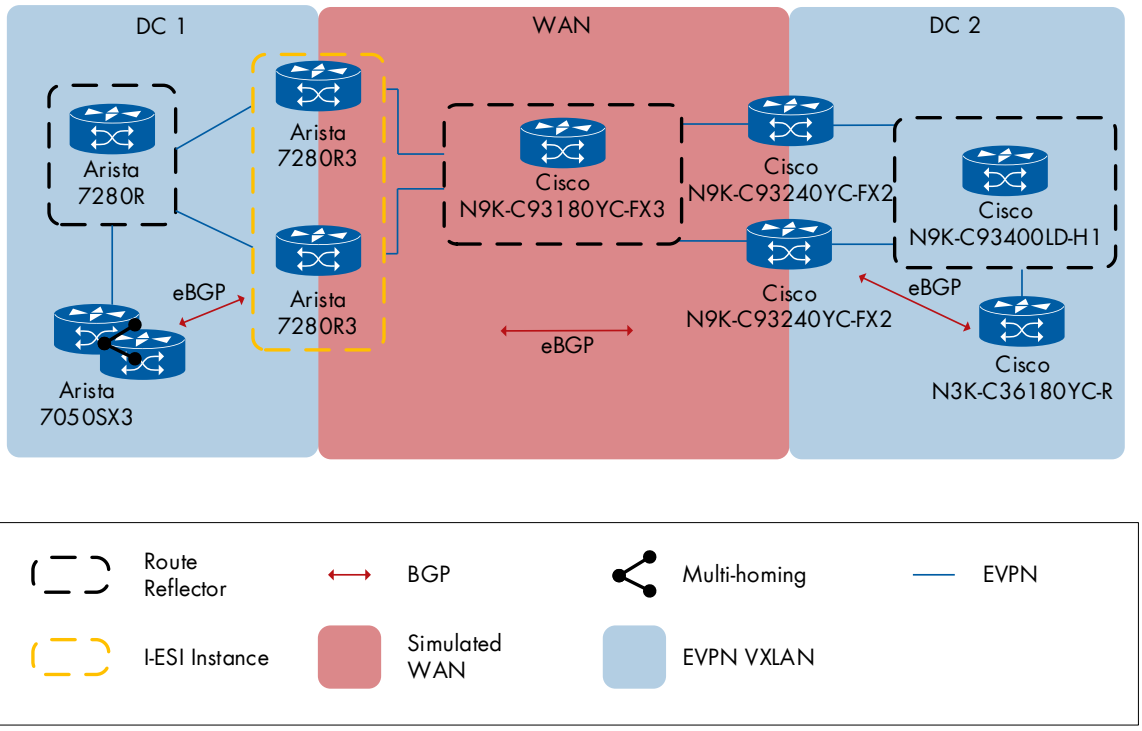


Figure 26: Interconnect Solution for EVPN Overlay and Multi-Site Solution for EVPN Overlay Interworking

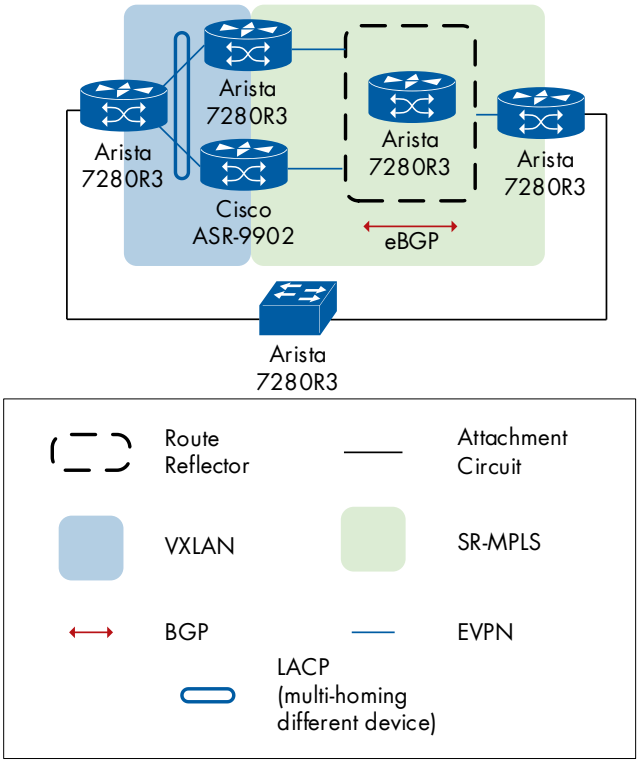


Figure 27: EVPN-VXLAN and EVPN-SR-MPLS Interworking Combination 1

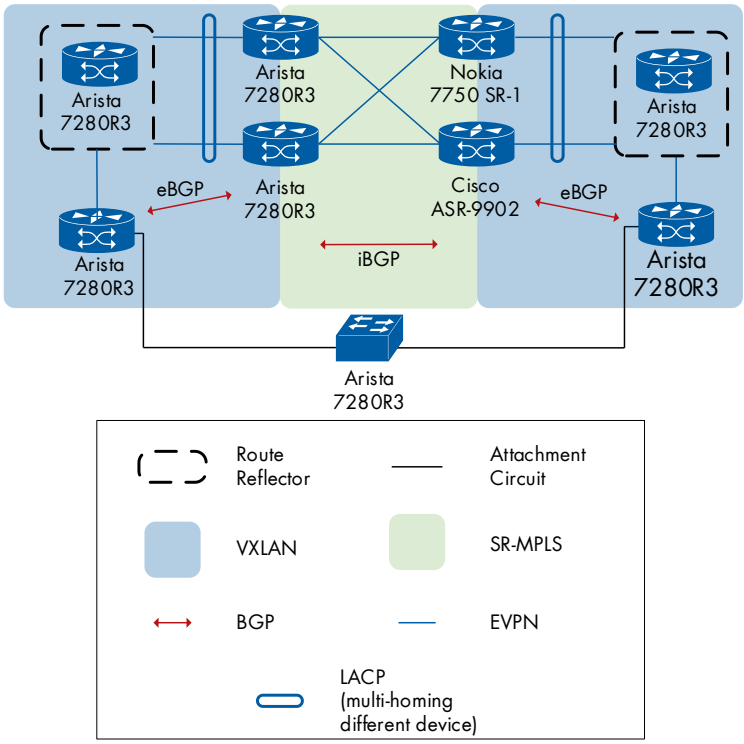


Figure 28: EVPN-VXLAN and EVPN-SR-MPLS Interworking Combination 2

EVPN-VXLAN and EVPN-SR-MPLS Interworking

In real-world networks, EVPN may need to pass multiple WAN transportation domains to reach the destination. Therefore, multi-domain interworking gateways play a crucial role in networks. To support the interworking gateway, the Interconnect Ethernet Segment and Interconnect Ethernet Segment Identifier (I-ES and I-ESI) should be supported to achieve the interworking target.

During our test, we conducted two runs with bi-directional unicast traffic. The traffic passed through the EVPN-VXLAN and EVPN-SR-MPLS domains in both runs. MAC addresses were learned from both domains and presented in a single EVI MAC table. Traffic was balanced between the multihoming PEs as expected.

The test topologies are shown in Figures 27/28.

EVPN and IPVPN Interworking

EVPN/IP-VPN Interworking is commonly required. EVPN can operate across multiple domains, while BGP serves as the universal control plane for the overlay. To ensure loop avoidance and determination of appropriate paths, EVPN must interwork with IP-VPN using EVPN RT-2 and RT-5's information.

During our test, we conducted three runs. The first run (see Figure 30 below) was a combination of EVPN-VXLAN and IP-VPN-SR-MPLS. Both sides used SR-MPLS transport for the second and third runs (see Figure 31 and Figure 31 below), with EVPN and IP-VPN overlay interworking. We utilized a standard BGP community

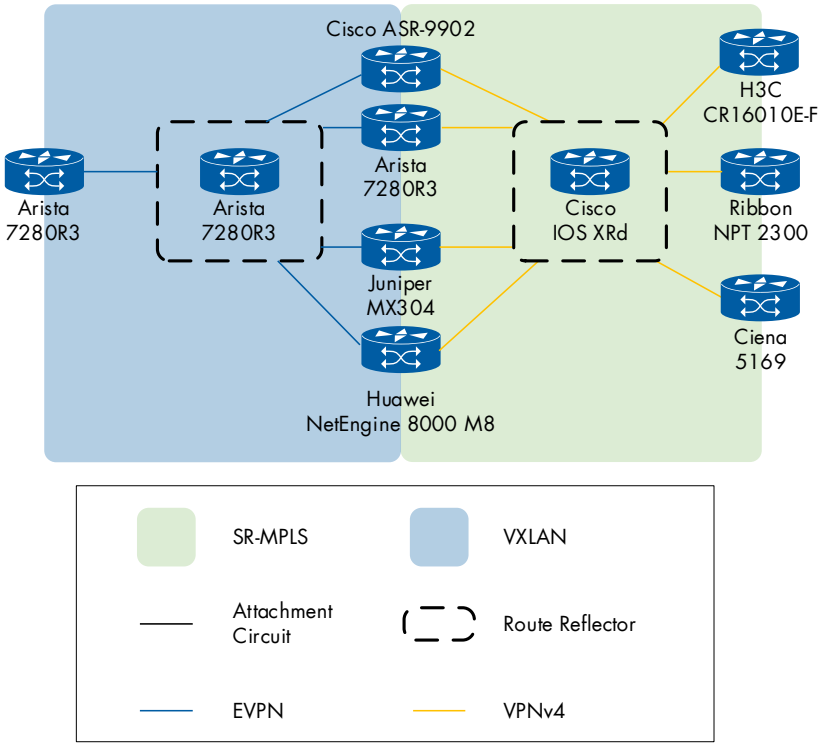


Figure 29: EVPN-VXLAN/IP-VPN-SR-MPLS Interworking

for loop detection and did not test the D-path community. We transmitted bi-directional unicast packets, which were successfully delivered end-to-end.

Figure 30 shows the first of two test combinations for this interworking scenario.

The second combination is not displayed here because its topology was completely identical, with the only exception that the Huawei router (bottom gateway in the diagram) was replaced by a H3C S12500R-2L router.

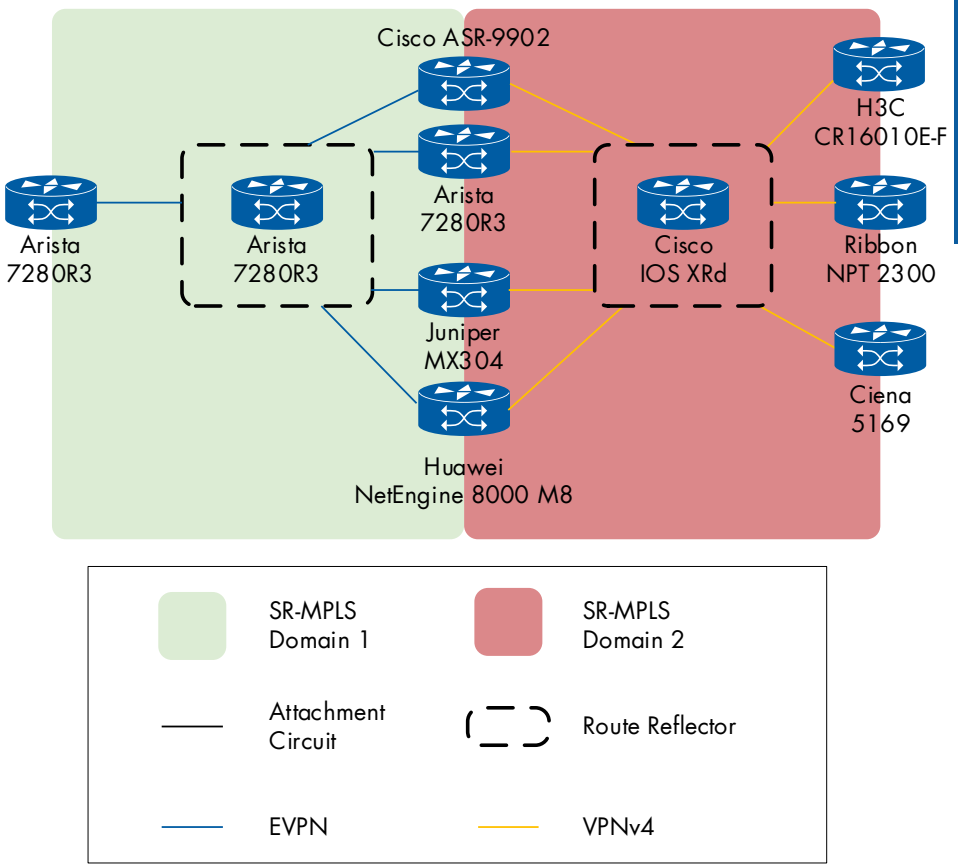


Figure 30: EVPN-SR-MPLS and IPVPN-SR-MPLS Interworking Combination 1

EVPN VXLAN with IPv6 VTEPs

As the global public IPv4 address pool is depleted, ISPs are gradually adopting IPv6. As a result, data center networks are transitioning to IPv6 as well. By leveraging IPv6 unnumbered underlay, we can significantly simplify the configuration process of the IPv6 underlay and enable BGP to establish peering sessions without explicit IP address configuration on the interfaces.

Last year, we conducted tests of IPv6 unnumbered underlay, overlay, and VTEPs but with IPv4 hosts. This year, we demonstrated a pure IPv6 network, where hosts were IPv6 as well. We conducted two tests with the pure IPv6 network, including symmetric IRB with VLAN-based and VLAN-Aware bundle service interfaces. We sent full mesh unicast traffic to verify end-to-end connectivity and observed no packet loss.

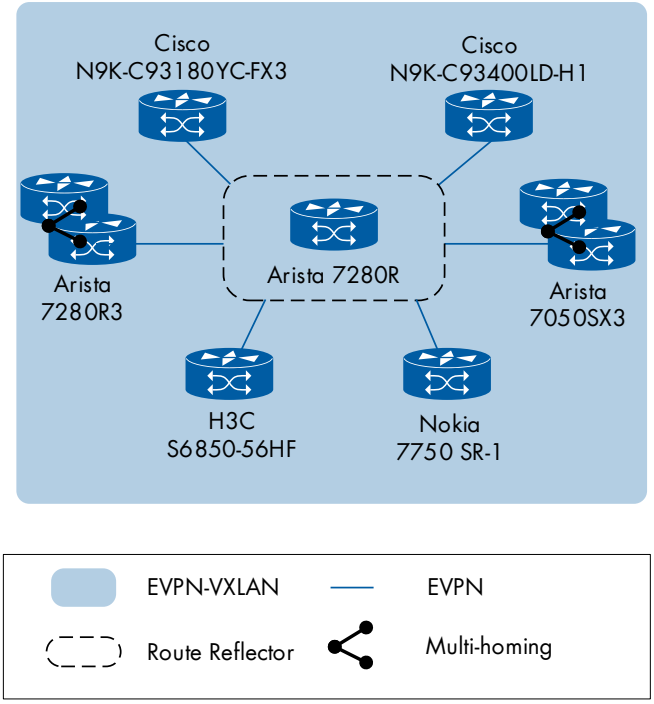


Figure 31: IPv6 Symmetric IRB, VXLAN, VLAN-Based

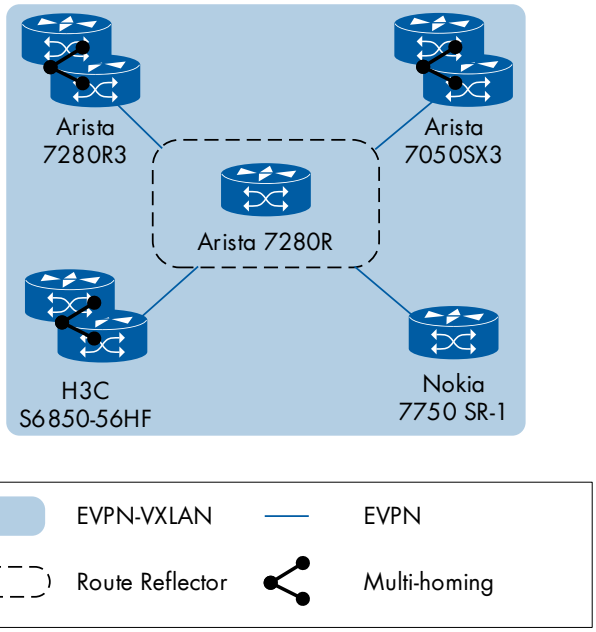


Figure 32: IPv6 Symmetric IRB, VXLAN, VLAN-Aware Bundle

Optimized Inter-Subnet Multicast (OISM) Selective multicast Ingress Replication (IR) and PIM/EVPN Gateways (PEG) election

Multicast is a crucial technology that helps conserve bandwidth and reduce network load. It is used for real-time data and multimedia. Typically, it works alongside the Protocol Independent Multicast (PIM) protocol, which aims to minimize the number of multicast copies. However, PIM is not always the best solution for the Network Virtualization Overlay (NVO) core network. This is where the optimized Ingress Replication (IR) is introduced in draft-ietf-bess-evpn-optimized-ir, providing a more tailored approach to multicast traffic transportation within the NVO core network.

First, we verified the IR function with the topology in Figure 33. We proved the IR worked as expected. We sent IGMPv3 join messages from the simulated hosts and observed SMET (RT-6) in DUT's routing table. We then generated multicast traffic from the source to receivers and saw no packet loss.

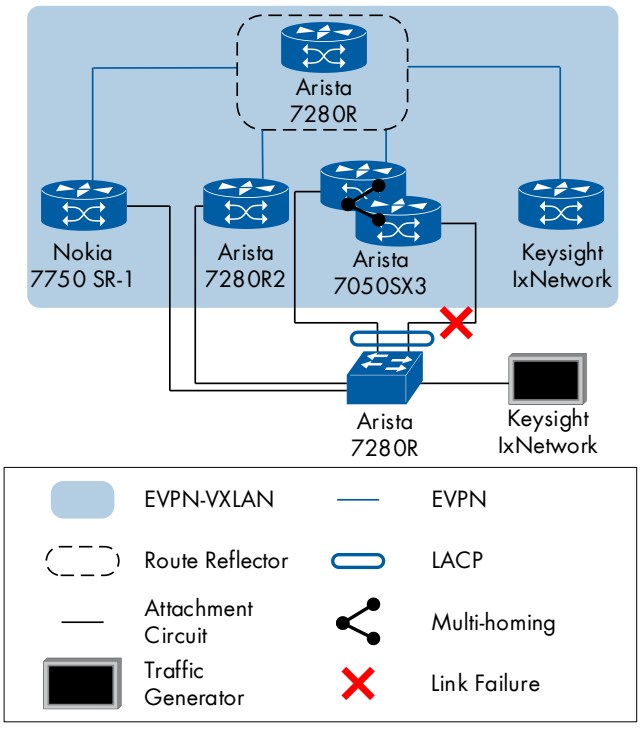


Figure 33: OISM Selective Multicast IR

One of the challenges in the NVO network is the suboptimal scenario of inter-subnet multicast traffic. Sometimes, the inter-subnet multicast traffic traverses the entire network, even if the destination is physically close to the source. This inefficiency underscores the need for a solution like EVPN OISM-based forwarding solution in draft-ietf-bess-evpn-irb-mcast, designed to handle IP multicast traffic in complex L2/L3 network environments. It capitalizes on the IRB interface to streamline the multicast traffic forwarding process.

We then conducted the second PIM/EVPN Gateways (PEG) election test. In this test, we used a minimal but comprehensive topology. The setup included an external PIM gateway (with PIM-SM configuration), 2 PEGs multihomed, and the IR was used inside the NVO network. The PEG Election used a modulus-based DR election. We sent bidirectional multicast traffic from emulated hosts connected to the external PIM router and the Provider Edges (PEs) in the EVPN VXLAN data center, and the multicast traffic passed end-to-end as expected.

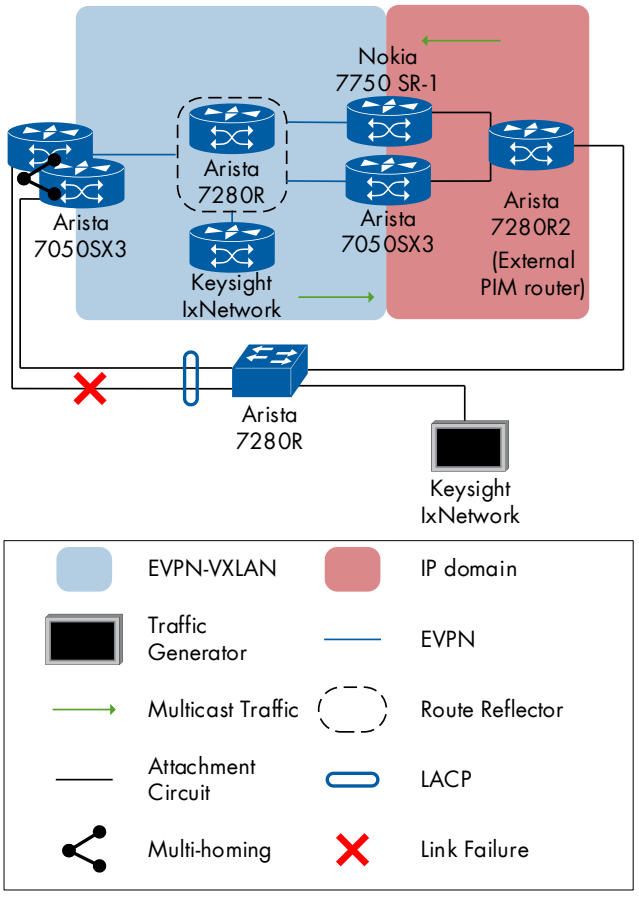


Figure 34: OISM-based L3 Multicast PEG

Continuing with the OSIM testing, we previously tested the same OISM-based IR using IPv4. Therefore, this time we have tested it with a pure IPv6 implementation, both in overlay and underlay, with participation from two vendors.

We assessed the control plane interoperability by checking the IMET and SMET routes in respective routing tables. The multicast FIB on the DUT was populated with multicast groups and outgoing interfaces. However, the IPv6 multicast traffic didn't flow as expected, and the vendor didn't have time to troubleshoot further because the event was about to close.

SR-MPLS Test Results

Segment Routing (SR) over Multiprotocol Label Switching (MPLS) has emerged as a cornerstone transport network technology, driving efficiency, scalability, and flexibility across network architectures. SR MPLS addresses new network demands (caused by the exponential growth in data traffic, increased cloud computing, and the rollout of 5G networks) that require more intelligent, robust, and agile routing methodologies.

This year's EANTC interop event testing included essential evaluations such as L3VPNs service, Topology Independent Loop-Free Alternate (TI-LFA), and SR Traffic Engineering (TE) policies and explored new territories. For the first time, we included tests focused on an IPv6 control plane and Bit Index Explicit Replication (BIER) and executed critical scenarios for inter-AS communication using BGP-LU (labeled Unicast) with prefix SID and employing anycast SID.

The incorporation of an IPv6 control plane is a recognition of the global shift towards IPv6, leveraging the established SR MPLS technology (not to be confused with SRv6, which is covered in a separate section of this report). Additionally, the BIER functionality underscores a move towards more efficient multicast routing solutions, which are crucial for bandwidth optimization and the enhancement of multicast traffic delivery. Moreover, our examination of inter-AS scenarios using Anycast addresses the indispensable need for seamless, efficient routing across autonomous systems, a key requirement for today's interconnected large networks with private peering connections beyond the default paths to Internet backbones.

L3VPN over SR-MPLS

The basic interoperability test for Layer 3 Virtual Private Networks (L3VPN) is fundamentally important as it is an initial step to ensure that the test bed is fully operational and ready for in-depth evaluation. This procedure began with vendors verifying the control plane, including Border Gateway Protocol (BGP), Interior Gateway Protocol (IGP), and the proper configuration of Segment Routing (SR) MPLS labels. Once these configurations were confirmed to function correctly, vendors deployed L3VPN services, and IPv4 traffic was sent end-to-end to confirm proper forwarding paths in each test run.

Our test architecture was designed around two distinct environments, each based on a different IGP: one leveraging the IS-IS protocol and the other the OSPF protocol, structured within a spine-and-leaf topology. Later, we examined the routers' forwarding tables to ensure that all routes were accurately installed.

Following this, we initiated a traffic flow across all network nodes and observed zero packet loss over all streams.

To reduce cabling overhead, all systems participating in the bulk of the SR-MPLS tests were connected to a main hub router. This year, the vendors agreed to select a Cisco router for the IS-IS topology and a Juniper router for the OSPF topology, based on logistics considerations. All tests were carried out full-mesh with pairs of all participating systems; Cisco or Juniper did not have a special role.

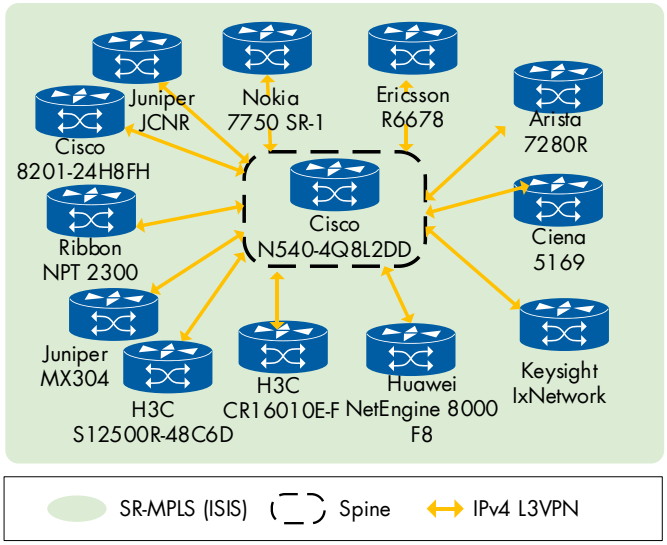


Figure 35: SR-MPLS over IS-IS

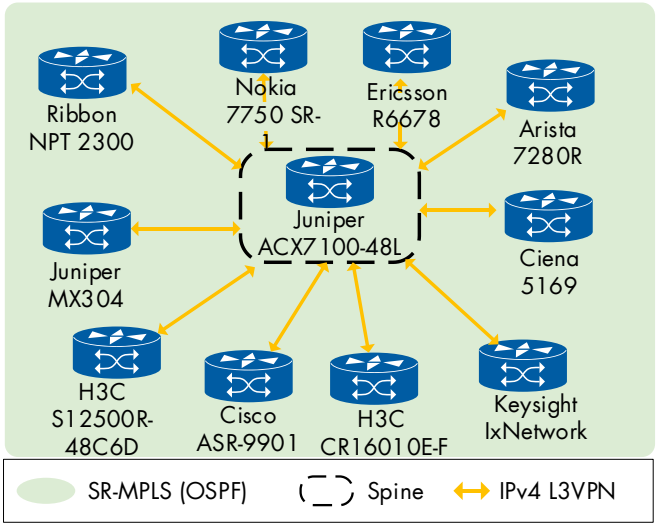


Figure 36: SR-MPLS over OSPF

Topology Independent Loop-Free Alternate (TI-LFA) over SR MPLS

TI-LFA is a fast reroute mechanism designed to work over Segment Routing (SR) MPLS networks. Its primary function is to protect against link or node failures, guaranteeing a fast recovery of a working path in case of such infrastructure failures. TI-LFA pre-installs a backup forwarding entry for each protected destination ready to be activated instantaneously upon detecting the failure of a link used to reach the destination.

Testing this capability across all participating devices is essential to confirm operational reliability. Doing so guarantees that TI-LFA functions as intended, safeguarding network integrity by promptly responding to failures and maintaining uninterrupted service delivery. This is vital for network operators relying on TI-LFA to reroute traffic and minimize the impact on network performance and user experience.

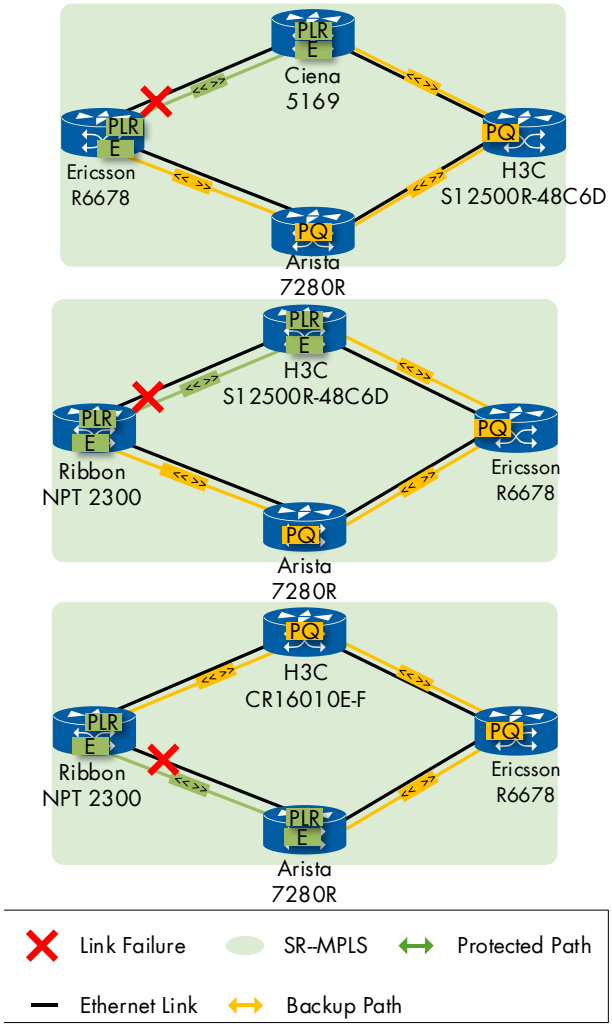


Figure 37: TI-LFA over SR-MPLS (Local Link Protection)

We included the TI-LFA test case in our annual interop event scope for many years, and it is kept as a regression test and basic test for newly participating equipment because of its fundamental importance.

Each multi-vendor configuration involved a logical network of four nodes arranged in a diamond topology, where we established an L3VPN service. Traffic was initially routed by the Point of Local Repair (PLR) directly to the adjacent node. During the setup, we configured TI-LFA on the node responsible for the protected link, establishing both primary and backup paths within the forwarding plane. To test the resilience of this configuration, we simulated a link failure by physically disconnecting the link. This process allowed us to observe the activation of the backup path, which was successfully implemented in under 50 milliseconds.

All participated nodes demonstrated accepted failover time ranging between 4.6 ms and 39 ms.

The following devices participated successfully:

PLR node	PQ node
Ribbon NPT 2300	Ericsson R6678
H3C S12500R-48C6D	Arista 7280R3
Arista 7280R3	H3C S12500R-48C6D
Ciena 5169	Arista 7280R3
Ericsson R6678	H3C S12500R-48C6D

Table 3: TI-LFA over SR-MPLS Pairs

Inter-AS option C

Inter-AS Option C represents a key approach for scalability and expanded connectivity across multiple autonomous systems (AS). This option is required by service providers with extensive routing and label information exchange needs; it helps to avoid overburdening the core network infrastructure.

In our testing, we looked into Inter-AS Option C's integration with BGP Labeled Unicast (BGP-LU), focusing on how it establishes Label Switched Paths (LSPs) for services such as L3VPN. These services require a robust interconnection between separate AS, a task traditionally managed by an IGP within a single AS, but which must transition to an Exterior Gateway Protocol (EGP), like BGP, when spanning multiple AS.

The use of BGP Prefix-SID (defined in IETF RFC 8669) in these tests is particularly interesting. It marks a significant evolution from conventional MPLS operations, as it enables the conveyance of SIDs as an attribute within BGP. We thoroughly evaluated the performance and reliability of two types of TLVs within

the BGP Prefix-SID: the Label-Index TLV, which is essential for indexing the SID, and the Originator SRGB TLV, which optionally provides SRGB details from the originating node for label calculation.

In our test configuration, we established two ASs linked by one or more Autonomous System Border Routers (ASBRs) serving as inline Route Reflectors (RRs).

At the core of this configuration, a PE initiated a BGP-LU update to advertise its own network prefix. This advertisement was marked with an implicit null label, which signifies that this PE is the endpoint of the MPLS path. Additionally, this BGP LU update was augmented with a BGP Prefix-SID attribute that included a Label Index TLV, and an SRGB TLV

Following this, the ASBR received the BGP LU update and was responsible for propagating this information to the PE located in the second AS, but first, the next hop was changed to self. This control plane setup was then tested by forwarding L3VPN traffic between PEs from different ASs.

We conducted tests using a single ASBR and then with multiple ASBRs, all successful. However, one test run, which is not included in the results, involved a scenario where the router serving as an ASBR, upon receiving an egress advertisement of label 3 (implicit null) and the removal of the PHP bit on the Node SID, eliminated all transport labels in its capacity as an ASBR and consequently dropped the transit traffic.

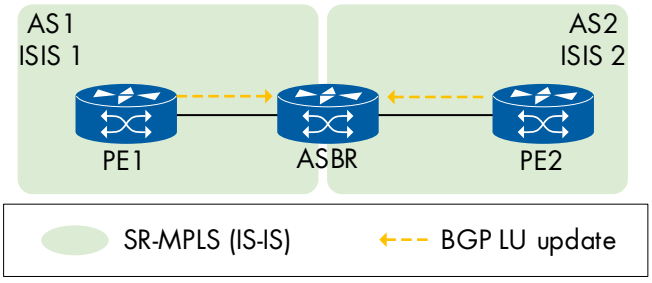


Figure 38: Inter-AS Option C with BGP-LU

Table 4 explains which devices were successfully tested in the roles of ASBR and/or PE.

Inter-AS Option C Using Anycast Next Hops

Following the establishment of the Inter-AS option C setup, our examination proceeded to deploy Anycast SID within a BGP-SR framework.

This involved configuring multiple ASBRs to broadcast the same prefix, accompanied by an identical SID.

Device	ASBR	PE
Arista 7280R3	X	X
Ciena 5169		X
Cisco 8201-24H8FH	X	
Cisco ASR-9901		X
Ericsson R6678		X
H3C CR160010E-F	X	X
H3C S12500R-48C6D		X
Huawei NetEngine 8000 F8	X	X
Juniper ACX7100-48L		X
Juniper MX304	X	
Nokia 7750SR-1	X	
Nokia 7250 IXR		X
Ribbon NPT 2300	X	X

Table 4: Inter-AS Option C with BGP-LU Results

Anycast routing was utilized to direct traffic towards several advertising nodes, ensuring that packets aimed at an Anycast address were routed to the closest node regarding network topology.

We verified that the label stack imposed by the PEs for routing to the egress included three specific labels: service, egress pe, anycast asbr.

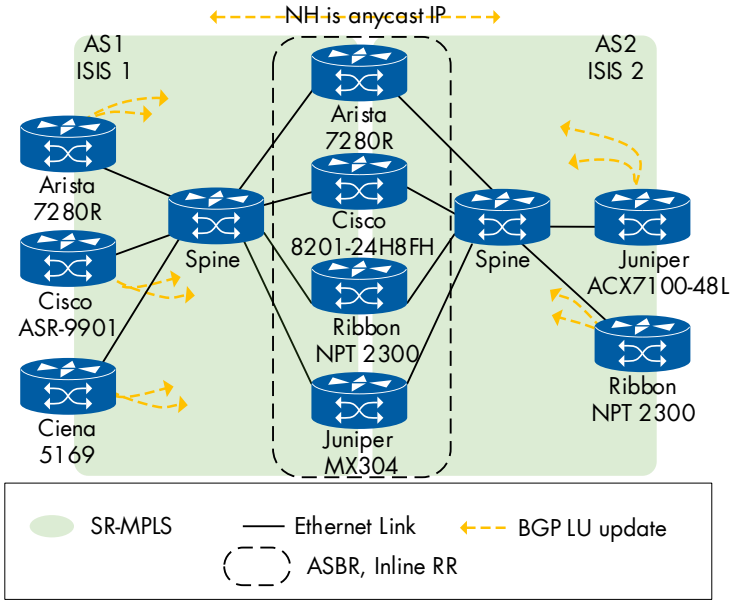


Figure 39: Inter-AS Option C using Anycast

Per Destination SR TE Traffic Steering

With SR policies, it is possible to design multiple pathways, known as candidate paths, each potentially having its own set of segments to navigate the network's topology. This flexibility allows customized routing strategies, such as load balancing across different links to manage network congestion or prioritize certain paths for critical services. These paths can be manually configured or be advertised using various protocols, such as PCEP or BGP, ..., which introduces a seamless way to implement SR-TE policies across big networks.

We began by connecting the two PE routers via two spine routers (see Figure 40 for logical test topology). This created two potential paths for data to travel between the PE routers. Then, we advertised BGP service routes with a color extended community from the tail-end router to the head-end router using BGP. At the head-end router, we had an SR policy in place that matched the endpoint address of the tail-end router and the color community value. This policy was designed to steer traffic based on the color tag it encountered.

We confirmed that traffic flowed through the specified SR-TE policy when applying the color tag. When the color tag was removed, the traffic was expected to be rerouted to an alternative tunnel or SR native path. Should there be no alternative route available, the traffic was to be dropped.

The following devices completed the test successfully:

Headend 1	Headend 2
Arista 7280R3	Ribbon NPT 2300
Nokia 7750 SR-1	Arista 7280R3
Ribbon NPT 2300	Nokia 7750 SR-1
Arista 7280R3	Huawei NE 8000 F8
Juniper MX304	H3C CR16010E-F
Juniper MX304	Nokia 7750 SR-1
Huawei NE 8000 F8	Nokia 7750 SR-1
Ribbon NPT 2300	H3C S12500R-48C6D
Nokia 7750 SR-1	H3C S12500R-48C6D
Ribbon NPT 2300	Ciena 5169

Table 5: SR-TE Steering per Destination Test Results

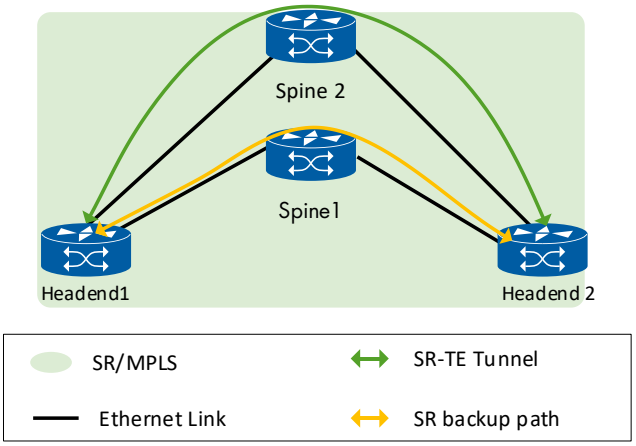


Figure 40: SR-TE Steering per Destination

SR Policy Liveness

Segment Routing policies should be checked for their "liveness" to detect faults quickly, identifying any failing part within the network elements and paths used by the respective SR policy. By recognizing these issues promptly, automated self-healing can be initiated immediately, ensuring minimal disruption to the affected data flows.

We focused on two primary liveness techniques: Seamless Bidirectional Forwarding Detection (S-BFD) and Two-Way Active Measurement Protocol (TWAMP) in loopback mode. Each method offers a unique approach to monitoring the network's health.

S-BFD operates on the principle of sending continuous test messages between two points within the network. These probes help detect discrepancies in the path, signaling potential issues if they fail to return within a predetermined interval.

TWAMP in loopback mode is employed for a simplified approach without the need for any signaling to bootstrap the performance monitoring session. Probes from a device contain specific labels sent using Segment List (s) of the SR Policy Candidate Path that guide them back along the same path. Configured with a time interval and a failure threshold of three missed probes, this method ensures a backup route is engaged promptly upon detecting path failures.

Our test was designed to assess the liveness of an SR-TE policy between two PE routers connected through dual spine routers, establishing two distinct paths. Initially, we set up an S-BFD session between the PEs to monitor the primary path's status. Subsequently, we introduced an Access Control List (ACL) on the spine link to drop the S-BFD packets selectively. This was executed without disconnecting the physical link,

aiming to simulate a failure scenario without triggering immediate link failure detection mechanisms such as TI-LFA.

The routers that were involved in the network were able to detect any faults that were induced through the absence of S-BFD or TWAMP probes. In response to the failure, the network switched the traffic to the secondary path promptly.

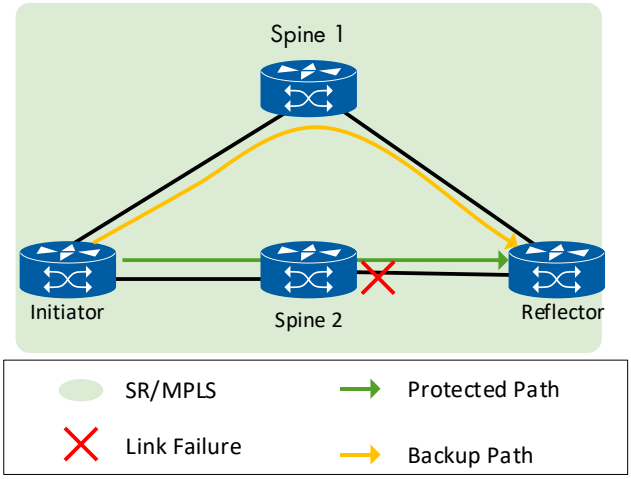


Figure 41: Liveness of SR Policy

The successful test combinations are shown in table 6 below.

Initiator \ Reflector	Arista 7208R3	Cisco 8201-24H8HF	H3C CR16010E-F	H3C S12500R-48C6D	Huawei NetEngine	Juniper MX304	Nokia 7750 SR-1	Ribbon NPT 2300
S = S-BFD T= TWAMP								
Arista 7280R3		S			S		S	S
Cisco 8201-24H8FH	T				T	T		T
H3C CR16010E-F					S			S
H3C S12500R-48C6D					S		S	
Huawei NE 8000 F8	S	S	S	S		S	S	
Juniper MX304		S			S		S	S
Nokia 7750 SR-1	S			S	S	S		S
Ribbon NPT 3200	S	S	S			S	S	

Table 6: Liveness SR Policy, Test Pairs

Optical Networks Using 400ZR/ZR+

Introducing 400G ZR and ZR+ long-range pluggable optics for routers marks a significant advancement in coherent optical technology. These modules utilize advanced modulation techniques and dense wavelength division multiplexing (DWDM) for 400 Gigabit Ethernet links. As a result, 400G ZR and ZR+ promise to reduce the costs associated with long-haul network links because routers can be interconnected directly without intermediate transmission network equipment.

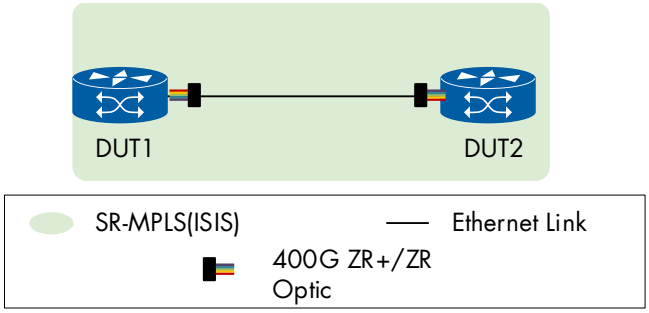


Figure 42: Optical Networks Test Scenario Using 400GbE ZR/ZR+ Optics

We conducted a series of tests on 400G ZR+ optics, engaging them across various operational modes that utilized open Forward Error Correction (OFEC). These tests included modes 1x400, 1x100, 2x100, 3x100, and 4x100, exclusively between ZR+ optics. In these cases, the technology requires agreeing on the DWDM carrier wavelength between vendors and following the right modulation (QPSK, 8QAM, or 16QAM modulation).

In a back-to-back setup, we confirmed interoperability across these modes without issues, except for the 3x100 test. In this case, one vendor needed to switch from their proprietary Enhanced mode to the standard ZR+ mode to achieve a successful outcome.

Additionally, we explored the interoperability between ZR+ and ZR optics in a 1x400 configuration. Vendors adjusted the frequency/wavelength for the latter to match ZR specifications while setting the FEC and modulation according to the OpenZR standards for the 1x400 mode. The participating devices and pluggables are listed in table 7.

Flexible Algorithms

For optimal 5G network performance, precise data routing is necessary to meet specific application requirements such as low latency, high bandwidth, and minimal packet loss.

Normally, network paths are determined based on the shortest distance between two points. However, this method falls short when multiple services compete for the same route, leading to potential bottlenecks. Traffic engineering has been the traditional solution to distribute digital traffic more evenly across different paths, but manual configuration is complex, resource-intensive, and error-prone.

This is where Flexible Algorithm (Flex-Algo), defined in RFC 9350, comes into play, introducing a key building block for SR-TE. Flex-Algo allows networks to be segmented into different planes, each governed by its own routing rules. This segmentation enables more nuanced control over data paths, enhancing network efficiency and service quality.

Our testing focused on four specific algorithms within this framework, targeting various operational scenarios:

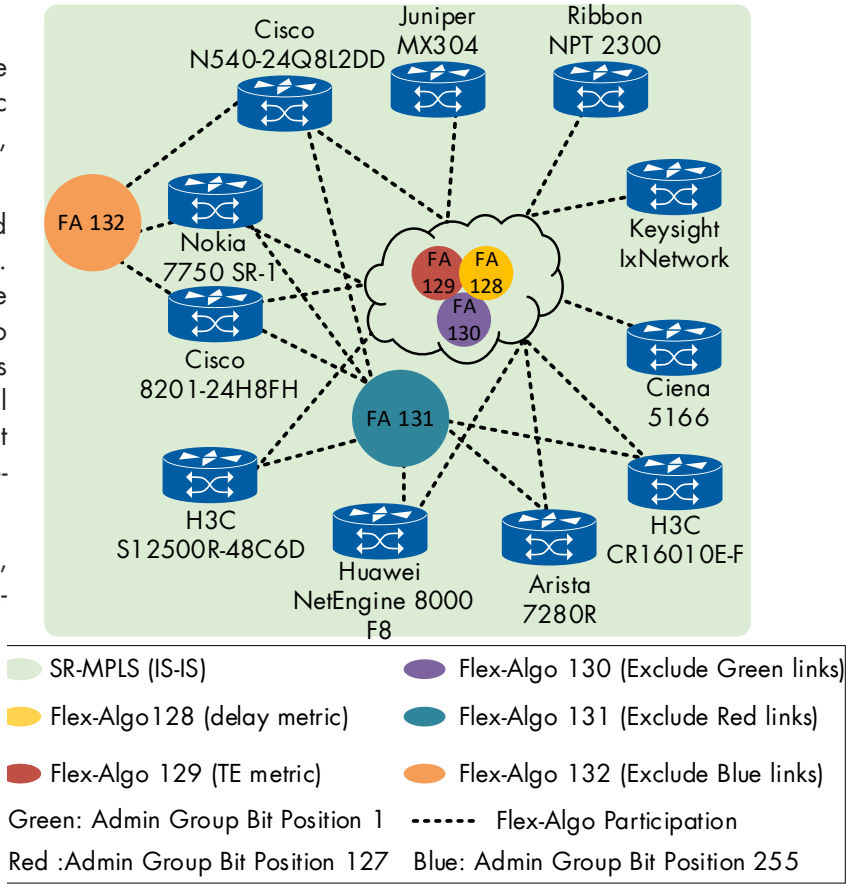


Figure 43: Flexible Algorithms over SR-MPLS

FA 128 Delay Metric: We tested the algorithm's capacity to prioritize paths with lower latency by utilizing Two-Way Active Measurement Protocol (TWAMP) or statically assigned delay values on links.

FA 129 TE Metric: We assessed the algorithm's ability to select paths based on Traffic Engineering (TE) metrics.

FA 130, 131, and 132 Exclude Affinity with Three Options: RFC5305 introduces a type of link TLV called the Administrative Group (AG), specifying that each link can have up to 32 Administrative Groups. These groups are advertised using fixed-length 32-bit bit-masks. Later, RFC 7308 expanded on this by introducing Extended Administrative Groups (EAGs) along with a sub-TLV for IS-IS and OSPF protocols. This development allows network operators to advertise more than 32 "colors" or categories within a network.

Our tests included the examination of the algorithm's functionality to exclude certain paths based on affinity attributes. By setting bit positions to 1, 127, or 255,

DUT1	Pluggable	DUT2	Pluggable	DWDM amplified?	Operating Modes
ZTE ZXR10 M6000-4SE	ZR/H3C	Juniper PTX10002-36QDD	ZR+/Juniper	Yes	1x400
H3C S12500R-48C6D	ZR/H3C	ZTE ZXR10 M6000-4SE	ZR/H3C	Yes	1x400
Cisco 8201-24H8FH	ZR+/Cisco	ZTE ZXR10 M6000-4SE	ZR/H3C	yes	1x400
Ribbon NPT 2300	ZR+	Juniper PTX10001-36MR	ZR+/Juniper		1x400
H3C S12500R-48C6D	ZR/H3C	Juniper PTX10002-36QDD	ZR+/Juniper	Yes	1x400
Cisco 8201-24H8FH	ZR+/Cisco	Ribbon NPT 2300	ZR+		1x400
Cisco 8201-24H8FH	ZR+/Cisco	H3C S12500R-48C6D	ZR/H3C	Yes	1x400
Juniper PTX10002-36QDD	ZR+/Juniper	Cisco 8201-24H8FH	ZR+/Cisco		1x400,4x100, 3x100, 2x100 & 1x100

Table 7: Optical Internetworking Tests 400GbE ZR/ZR+

we verified the algorithm's interoperability in advertising link affinity using the extended administrative group. This involved mapping bit positions to specify which links should be included or excluded in the flexible algorithm calculation.

We verified that the participating nodes announced their capabilities through sub-TLVs at the same IS-IS level. Additionally, they advertise prefix SID information that links these SIDs to particular algorithm IDs, enabling algorithm-specific routing decisions.

A unique logical topology for each Flex-Algo was then generated, considering only the nodes participating in that algorithm and adjusting link inclusion based on configured constraints like administrative groups or required metrics.

Utilizing this topology, nodes were expected to calculate optimal routes based on Flex-Algo's defined metrics and calculation methods and install the path calculation result into its MPLS forwarding table.

Flexible Algorithm Using New Constraints

IETF draft-ietf-lsr-flex-algo-bw-con-07 (work in progress) introduces an advanced framework for establishing and implementing a variety of administratively assigned metrics through generic metrics. It also offers new Flexible Algorithm Definition (FAD) constraints, strategically enabling network administrators to avoid low bandwidth or high latency links.

This year, our focus expanded to include testing these new constraints, alongside exploring reverse affinity constraints as outlined in IETF draft-ietf-lsr-igp-flex-algo-reverse-affinity (work in progress). For our testing, we connected devices back-to-back across two ports and introduced three new Flex Algorithms:

FA 140 (Minimum Bandwidth): Sets a minimum bandwidth threshold, effectively bypassing links below this bandwidth limit.

FA 141(Maximum Delay): Establishes a maximum delay limit to avoid links exceeding a latency threshold.

FA 142 (Exclude Reverse Affinity): Implements an exclusion for reverse affinity, steering clear of links marked with a specific affinity by the remote end.

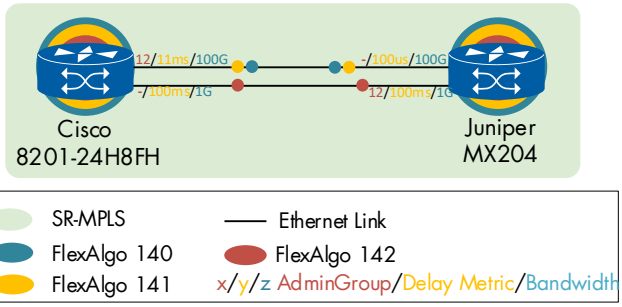


Figure 44: SR MPLS FlexAlgo with new constraints

We confirmed the advertisement of the new constraints through ISIS and ensured that the routes set up for the prefix SID associated with this algorithm comply with the designated constraints. If a link doesn't meet these specific constraints, it must be excluded from the Flex-Algo topology.

While testing the minimum bandwidth Flex-Algo, one of the devices incorrectly interpreted the bandwidth advertised by the other end. However, the issue was resolved after updating the software with the correct calculation method, and the test worked as intended.

	Flexible Algorithm Type					
PE Node	**128 TWAMP**	**128 Static**	**129**	**130**	**131**	**132**
Arista 7280R3	X		X	X	X	
Ciena 5169		X	X	X	X	
Cisco 8201-24H8FH	X	X	X	X	X	X
Cisco N540-24Q8L2DD	X	X	X	X	X	X
H3C CR16010E-F		X	X	X	X	
H3C S12500R-48C6D		X	X	X	X	
Huawei NetEngine 8000 F8		X	X	X	X	
Juniper MX304	X	X	X	X	X	
Keysight IxNetwork		X	X	X		
Nokia 7750 SR-1	X	X	X	X	X	X
Ribbon NPT 2300		X	X	X	X	

Table 8: Flexible Algorithms over SR-MPLS, Successful Test Results

TI-LFA with Flexible Algorithms

The IETF draft-ietf-rtgwg-segment-routing-ti-lfa-13 (work in progress) specifies that an implementation may optionally support Topology-Independent Loop-Free Alternate (TI-LFA) for protecting Node-SIDs linked to a specific Flexible Algorithm (Flex-Algo). It is mandated that an implementation must exclusively utilize Node-SIDs bound to the FlexAlgo and/or Adj-SIDs that are unprotected to build the repair list.

In our evaluation, the TI-LFA mechanism was applied within an FA topology. This involved setting up a testbed as depicted in Figure 45, with nodes configured under Flex-Algo 128.

Through this setup, we verified that in scenarios where a link failure disrupts the primary route, the TI-LFA algorithm efficiently utilized backup paths defined by Algo-128 directing traffic through it rather than the shortest IGP path.

All nodes exhibited the expected behavior, with failover times consistently below 50 milliseconds.

Nevertheless, an issue was encountered in a particular scenario involving a router that sent out two types of Routing Capabilities: a general type includes the Segment Routing Global Block (SRGB), and a second type with the Flexible Algorithm (FA). The receiving router did not process the SID from the general Routing Capability because its deployment solely acknowledged the second type associated with FA, which lacked SRGB details. Consequently, due to the absence of SRGB information in the FA-specific message, the router disregarded the Label Switched Path (LSP).

The vendor engineering team developed a patch that combined the two TLVs, successfully facilitating the correct installation of the Segment Identifier (SID).

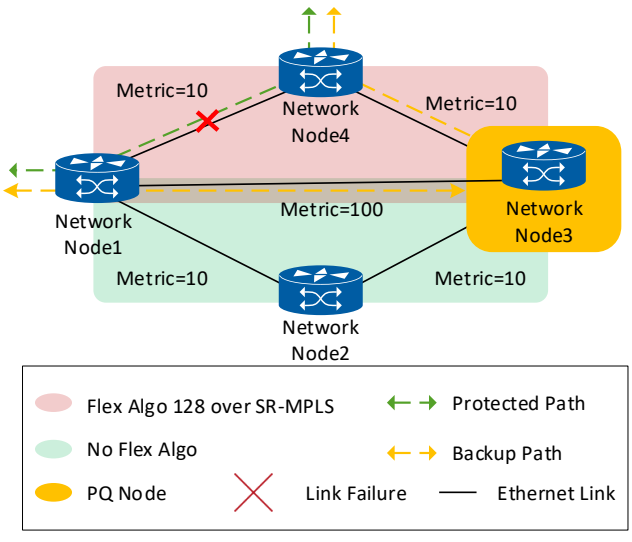


Figure 45: TI-LFA with FlexAlgo

There were four successful test combinations.

First test combination:

- Node1 (DUT): Arista 7280R3
- Node2: Arista 7280R3
- Node3: Ribbon NPT2300
- Node4: Nokia 7750 SR-1

Second test combination:

- Node1 (DUT): Juniper MX304
- Node2: Arista 7280R3
- Node3: Nokia 7750 SR-1
- Node4: Huawei NetEngine 8000 F8

Third test combination:

- Node1 (DUT): Nokia 7750 SR-1
- Node2: Arista 7280R3
- Node3: Arista 7280R3
- Node4: Ribbon NPT2300

Fourth test combination:

- Node1 (DUT): Ribbon NPT2300
- Node2: Arista 7280R3
- Node3: Ribbon NPT2300
- Node4: Nokia 7750 SR-1

SR-MPLS with IPv6 control plane

Implementing an IPv6-only control plane for SR-MPLS networks represents a significant and natural step forward in network architecture. This approach is in direct response to the worldwide transition to IPv6 due to the exhaustion of IPv4 addresses.

By leveraging IPv6, we benefit from its features over the well-established and matured SR-MPLS network.

The testing procedure included complete IPv6 addressing (Loopback and links), node SIDs, adjacency SIDs and Flex-Algo SIDs. Also the TWAMP light protocol was established using these IPv6 addresses.

The configuration of BGP VPN services utilized IPv6 next-hop addresses to route IPv4 address families over an IPv6 control plane. We also implemented BGP color communities to categorize VPN IPv6 and IPv4 prefixes, to deploy targeted traffic management policies.

In our testing scenarios, we employed color tagging for IPv4 and IPv6 traffic to direct data through specified routing paths based on different criteria. For IPv6 traffic, we applied color tags corresponding to a Flex-

Algo that prioritizes delay metrics. IPv6 traffic would be routed along the most delay-efficient paths, utilizing TWAMP's dynamic delay measurement capabilities.

On the other hand, IPv4 traffic was tagged with a different color intended to activate an on-demand-SR-TE policy. This policy was designed to guide the IPv4 traffic exclusively through paths associated with the 'blue' administrative group.

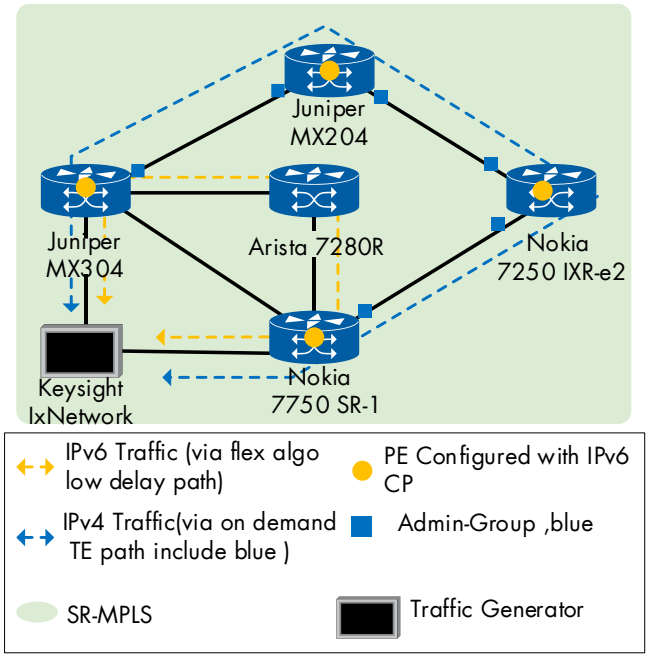


Figure 46: SR MPLS with IPv6 Control Plane

Bit Index Explicit Replication (BIER)

BIER (Bit Index Explicit Replication) is an architecture for multicast routing, designed to forward multicast data efficiently and efficiently. Instead of utilizing traditional multicast routing trees and storing state information at each router, BIER uses a bit-string, which is a sequence of bits, in the packet header to represent the destinations for a multicast packet. This approach simplifies the forwarding process significantly and reduces the amount of state information that needs to be maintained in the network.

The initial step involved ensuring that every edge node (PE) within a BIER sub-domain was assigned a BFR-ID unique to that sub-domain. Besides BIER Forwarding Router ID (BFR-IDs), additional data, such as the nodes' IP addresses, were propagated across the IGP. This process enables each network node to create its BIER forwarding information. We examined the BIER routing table and the BIER neighbors to confirm this.

Then, the BIER Forwarding Ingress Router (BFIR) sent multicast data, wrapped within a BIER header, to the BIER-Forwarding Egress Routers (BFERs) through transit BIER Forwarding Routers (BFRs). The header includes a BitString, among other elements, where each bit corresponds to the BFR-ID of a BFER. A set bit indicates that the associated BFER is a designated packet recipient. We verified that the transit BFRs reviewed the BitString and determined correctly which neighboring routers require packet replication, utilizing the Bit Index Forwarding Table (BIFT).

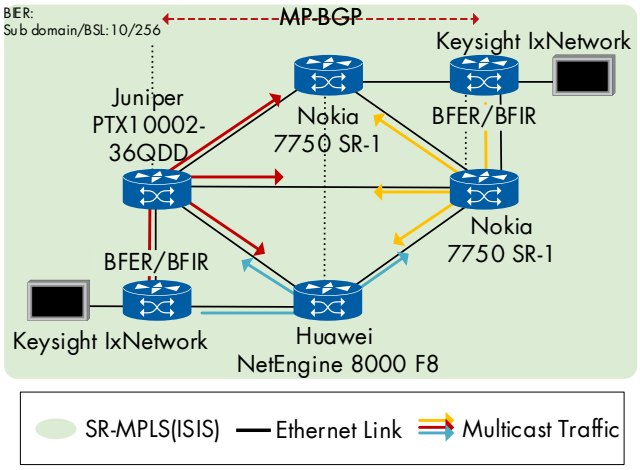


Figure 47: BIER Test Topology

The following devices participated in the test as:

- BFER/BFIR: Keysight IxNetwork
- Transit (Replicating) nodes:
  Huawei NetEngine 8000 F8,
  Juniper PTX10002-36QDD,
  Nokia 7750 SR-1

Segment Routing over IPv6 (SRv6) Test Results

SRv6 stands out in the evolving scene of network technologies for its innovative approach to simplifying network operations, enhancing programmability, and supporting the demands of modern network services, especially in the context of 5G and beyond.

This year marked a significant milestone in our testing procedures as, for the first time, we exclusively utilized micro Segment ID (μSID) across all our tests. We had started testing multi-vendor interoperability of μSIDs already in 2023, and this year, the concept has been adopted by all participating vendors for all SRv6 test scenarios. This move underscores a notable industry trend towards embracing this method. Moreover, we explored using SRv6 argument signaling for BGP service routes within the ELAN multi-homing test. This year also saw the first verification of multicast functionality using Msr6, alongside presenting a case for link resource slicing within an SRv6 framework.

It's also significant to highlight that our evaluations, including Layer 2/3 VPN services, SRv6 Locator Summarization, and Unreachable Prefix Announcement (UPA) frameworks and the rest of the tests, have drawn a wide array of participants this year. Such extensive involvement points to a considerable progression towards improved interoperability across the industry.

In all our testing scenarios, we connected every participating router to the traffic generator (Keysight IxNetwork) and initiated mesh traffic among all nodes. This was done over the service being evaluated or the service implemented, specifically to demonstrate the functionality of the feature under test.

L2/L3VPN Services over SRv6 Test

Layer 2/3 VPN services are basic constructs for transporting isolated and protected customer traffic across an SRv6 network – similar to other Segment Routing and MPLS network platforms.

Our testing began with constructing a comprehensive topology incorporating all participants, enabling us to conduct most of our tests using this setup. This topology featured two IS-IS levels linked by several Area Border Routers (ABRs). Every node was configured to establish a BGP connection with a route reflector and form IS-IS adjacencies with neighboring routers. The ABRs were set up to support both levels and were tasked with route leaking from level 2 to level 1, ensuring seamless connectivity across the entire network architecture.

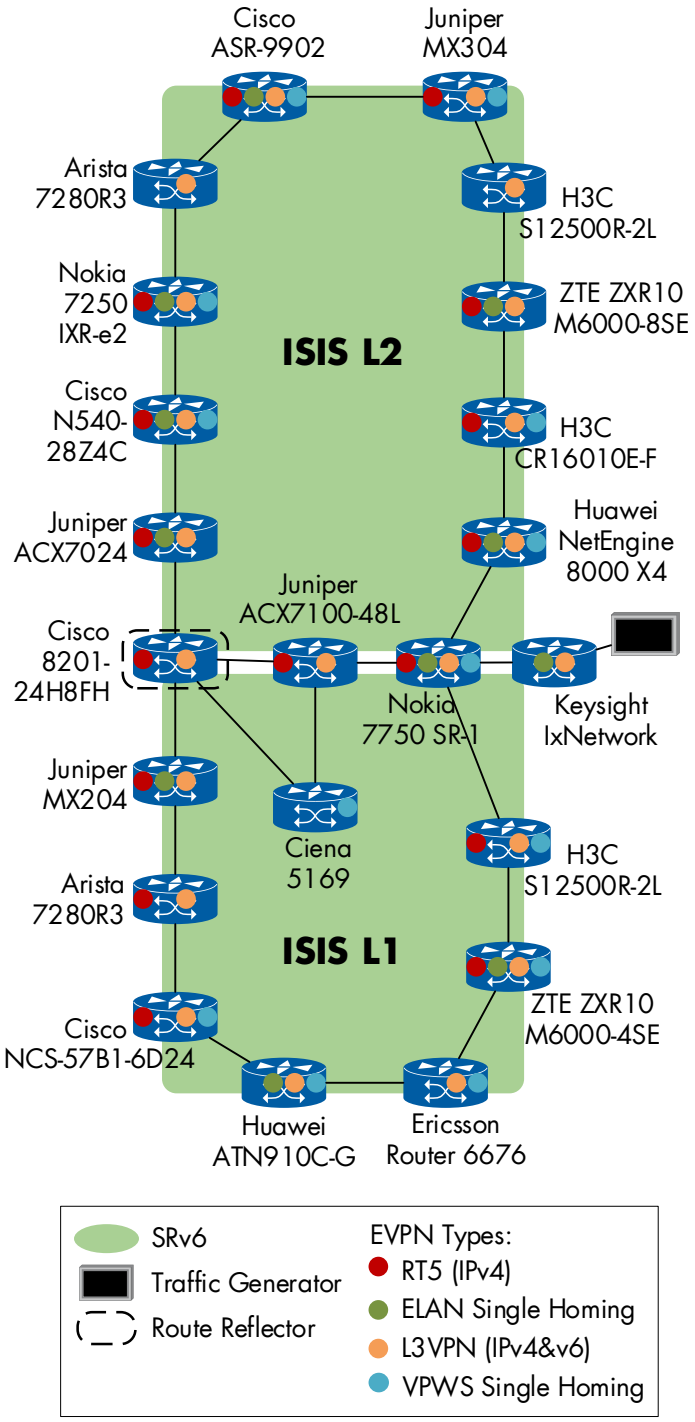


Figure 48: L2/L3 Services over SRv6

Configurations were established for SRv6 locators using 32-bit Locator Block and 16-bit Locator Node μSID format (F3216). Following IETF RFC 9252 (BGP Overlay Services Based on SRv6), vendors built various Layer 2 and Layer 3 VPN services. Some participants implemented the Transposition Scheme defined in RFC 9252 to increase the efficient packing of service routes. This method involves shifting the FUNCTION portion into the label field within a route's Network Layer Reachability Information (NLRI). Given that the remain-

ing segment of the SID remains unchanged for all routes belonging to the same service category, this transposition technique aids in compacting routes into a unified BGP update message, thereby enhancing update efficiency. All participants could correctly decode the SRv6 service SID out of the received route, irrespective of transposition, as long as the lengths were correctly encoded.

The following devices tested successfully for the transposition feature:

- PE node for VPN service: Ciena 5169
- PE node for EVPN service: Ericsson R6676
- PE node for VPN and EVPN services:
  Cisco 8201-24H8FH, Cisco ASR-9902,
  Cisco N540-28Z4C, Cisco NCS-57B1-6D24, H3C CR16010E-F, H3C S12500R-2L,
  Juniper ACX7024, Juniper ACX7100-48L,
  Juniper MX204, Juniper MX304,
  Keysight IxNetwork, Nokia 7250 IXR-e2,
  Nokia 7750 SR-1
- P nodes: Arista 7280R3, H3C CR16010E-F,
  Huawei ATN 910C-G,
  Huawei NetEngine 8000 X4,
  ZTE ZXR10 M6000-4SE, ZTE ZXR10 M6000-8SE.

L3VPN over SRv6

We configured a dual-ring topology for this test case, interlinking two IS-IS levels. On each PE, VRFs were established to support VPNv4 and VPNv6 address families. We validated the control plane by reviewing the routing tables for correct routes, SRv6 service SIDs, and their respective next hops. Then the traffic generation across all PEs was used to verify the data planes.

During the testing, a compatibility issue re-surfaced where one device failed to recognize the END.DT46 with NEXT-CSID message sent by another.

Additionally, to address a specific vendor's dependency on extended-IS TLVs for path validation—a requirement not met by four nodes that did not advertise these TLVs—an SR-Policy was implemented to detour around those nodes.

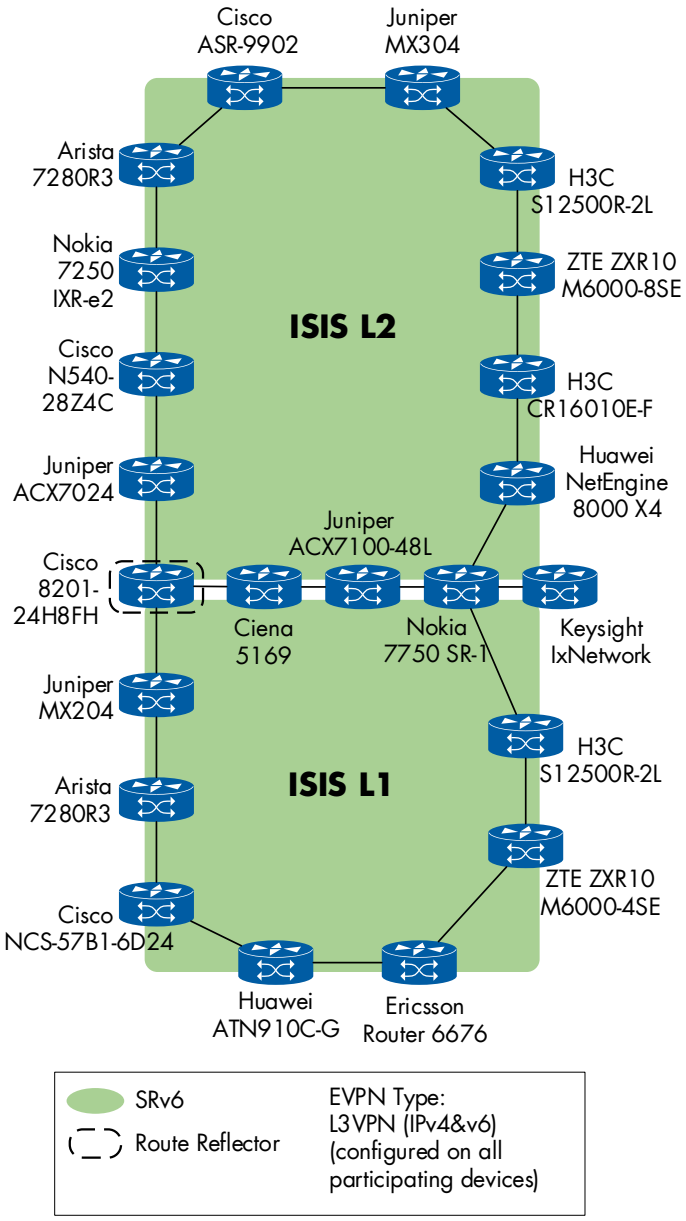
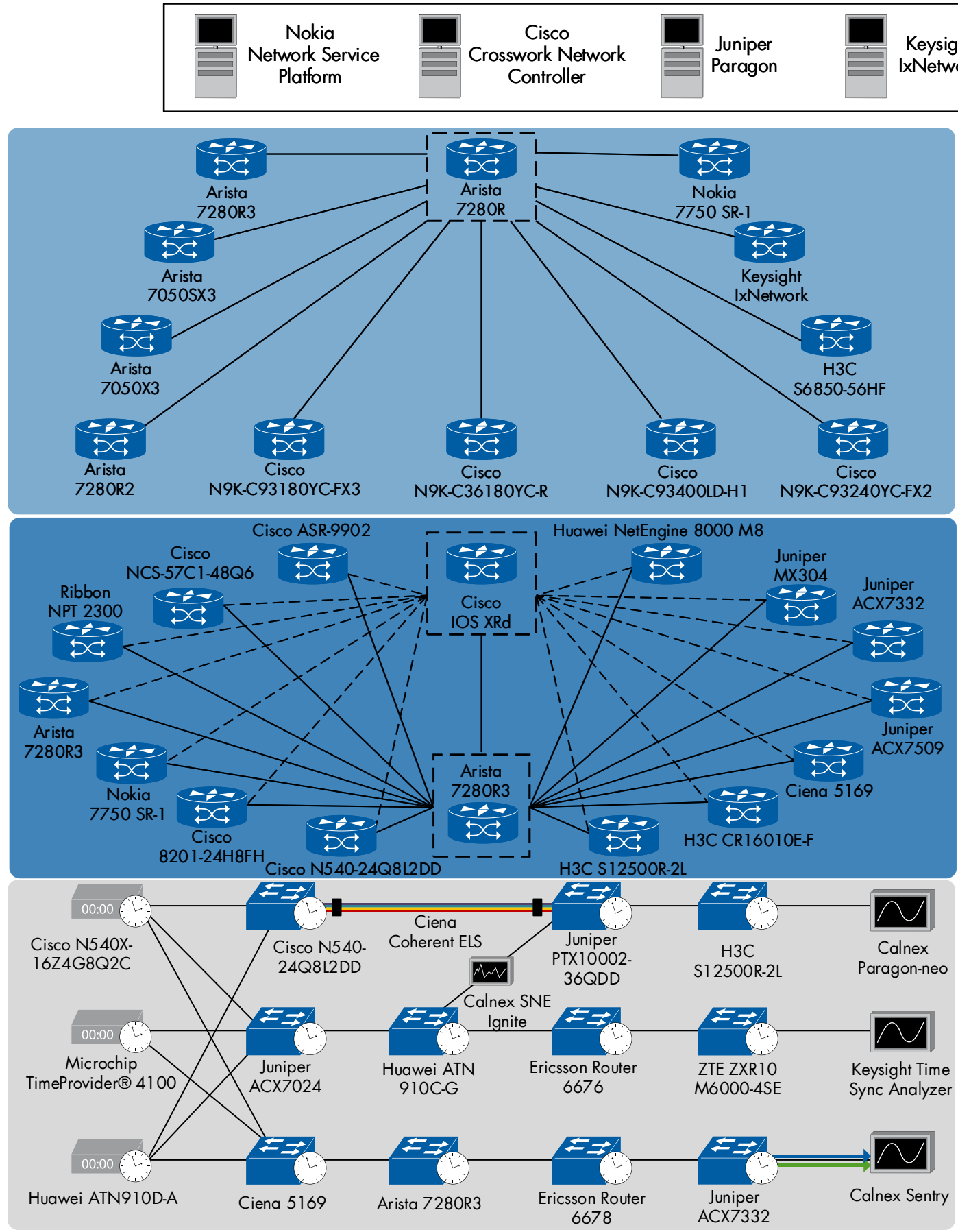


Figure 49: L3VPN over SRv6

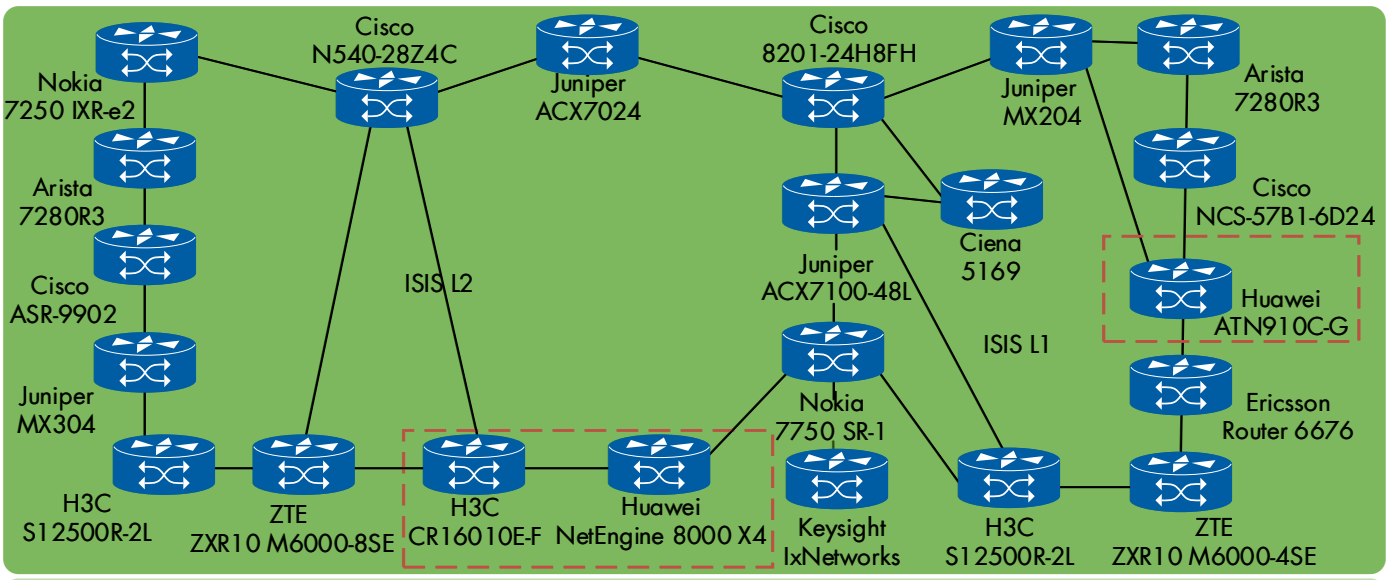
EVPN ELAN Single Homing over SRv6

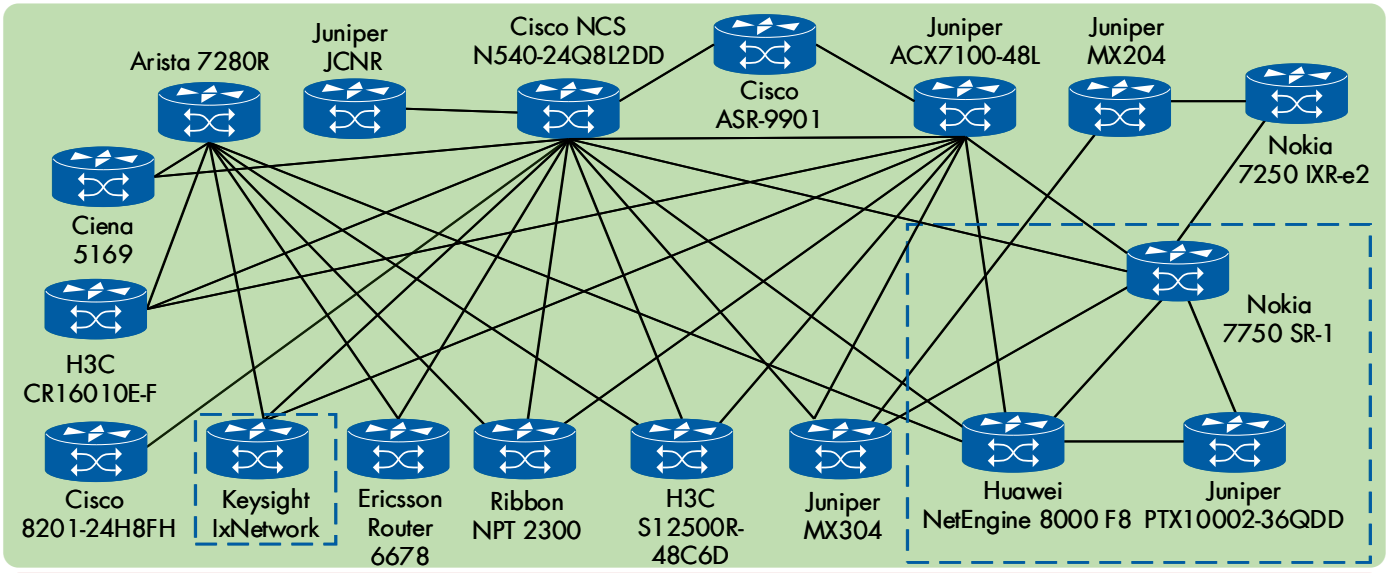
We successfully verified the implementation of EVPN E-LAN services over an SRv6 infrastructure. We checked that the bridge domain on all nodes was operational and in an "Up" state. Additionally, we observed that remote MAC addresses were successfully learned from peer PEs, indicating efficient Layer 2 learning and the successful reception and interpretation of EVPN MAC/IP Advertisement route (Route Type 2).

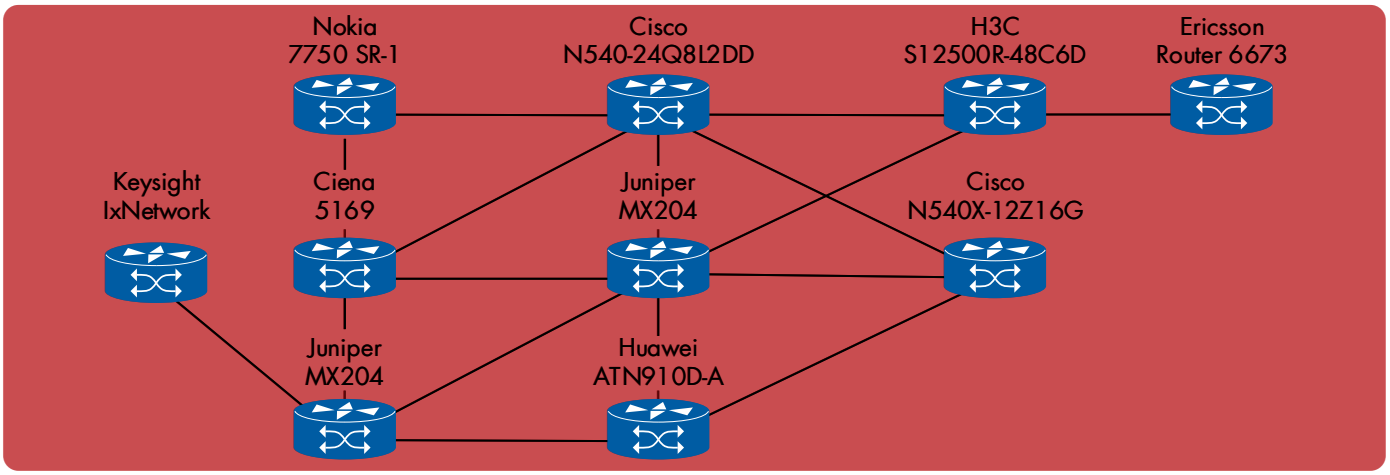
The following devices participated successfully as PE nodes: Cisco ASR-9902, Cisco N540-28Z4C, Huawei NetEngine 8000 X4, Juniper ACX7024, Juniper MX204, Keysight IxNetwork, Nokia 7250 IXR-e2, Nokia 7750 SR-1, ZTE ZXR10 M6000-4SE, ZTE ZXR10 M6000-8SE.


Nokia
Network Service
Platform
Cisco
Crosswork Network
Controller
Juniper
Paragon
Keysig
IxNetw
Arista
7280R3
Arista
7280R
Nokia
7750 SR-1
Arista
7050SX3
Keysight
IxNetwork
Arista
7050X3
H3C
S6850-56HF
Arista
7280R2
Cisco
N9K-C93180YC-FX3
Cisco
N9K-C36180YC-R
Cisco
N9K-C93400LD-H1
Cisco
N9K-C93240YC-FX2
Cisco ASR-9902
Huawei NetEngine 8000 M8
Cisco
NCS-57C1-48Q6
Juniper
MX304
Juniper
ACX7332
Ribbon
NPT 2300
Cisco
IOS XRd
Arista
7280R3
Juniper
ACX7509
Nokia
7750 SR-1
Arista
7280R3
Ciena 5169
Cisco
8201-24H8FH
H3C CR16010E-F
Cisco N540-24Q8L2DD
H3C S12500R-2L
00:00
Cisco N540X-
16Z4G8Q2C
Cisco N540-
24Q8L2DD
Ciena
Coherent ELS
Juniper
PTX10002-
36QDD
H3C
S12500R-2L
Calnex
Paragon-neo
Calnex SNE
Ignite
00:00
Microchip
TimeProvider® 4100
Juniper
ACX7024
Huawei ATN
910C-G
Ericsson Router
6676
ZTE ZXR10
M6000-4SE
Keysight Time
Sync Analyzer
00:00
Huawei ATN910D-A
Ciena 5169
Arista 7280R3
Ericsson Router
6678
Juniper
ACX7332
Calnex Sentry
Core Router
Grandmaster
Orchestrator/
Controller/PCE
Emulator
Boundary/
Slave Clock
Time Error
Analyzer


ht
ork
Huawei
iMaster NCE-IP
highstreet technologies
ht.Connect
Ciena
Navigator NCS


Nokia
7250 IXR-e2
Cisco
N540-28Z4C
Juniper
ACX7024
Cisco
8201-24H8FH
Juniper
MX204
Arista
7280R3
Arista
7280R3
Cisco
NCS-57B1-6D24
Cisco
ASR-9902
ISIS L2
Juniper
ACX7100-48L
Ciena
5169
Huawei
ATN910C-G
ISIS L1
Juniper
MX304
Nokia
7750 SR-1
Ericsson
Router 6676
H3C
S12500R-2L
ZTE
ZXR10 M6000-8SE
H3C
CR16010E-F
Huawei
NetEngine 8000 X4
Keysight
IxNetworks
H3C
S12500R-2L
ZTE
ZXR10 M6000-4SE


Arista 7280R
Juniper
JCNR
Cisco NCS
N540-24Q8L2DD
Cisco
ASR-9901
Juniper
ACX7100-48L
Juniper
MX204
Nokia
7250 IXR-e2
Ciena
5169
Nokia
7750 SR-1
H3C
CR16010E-F
Cisco
8201-24H8FH
Keysight
IxNetwork
Ericsson
Router
6678
Ribbon
NPT 2300
H3C
S12500R-
48C6D
Juniper
MX304
Huawei
NetEngine 8000 F8
Juniper
PTX10002-36QDD


Nokia
7750 SR-1
Cisco
N540-24Q8L2DD
H3C
S12500R-48C6D
Ericsson
Router 6673
Keysight
IxNetwork
Ciena
5169
Juniper
MX204
Cisco
N540X-12Z16G
Juniper
MX204
Huawei
ATN910D-A


BIER
Msr6
Route Reflector
Clock Link – Freq.
1PPS Link
Physical Link
Logicall Link
EVPN-VXLAN
EVPN-MPLS
SDN/NETCONF
SR-MPLS
SRv6
Clocking

EVPN ELAN Multi-Homing over SRv6

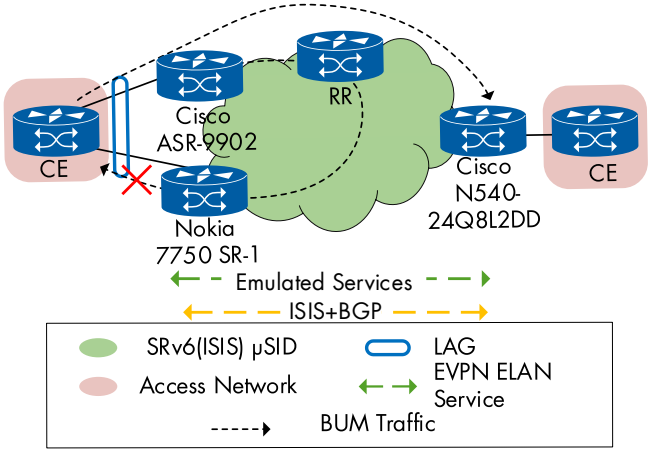


Figure 50: ELAN-Multihoming over SRv6 using Argument Signaling

In our testing this year, we verified a pioneering technique that utilizes the arguments within the SID for SRv6 to manage BUM traffic in an active/active E-LAN multi-homing setup. This approach was designed for ESI split-horizon filtering to prevent broadcast, unknown unicast, or multicast (BUM) packets from being forwarded back to the same multi-homed Ethernet segment where they originated and potentially causing network redundancy and service disruption.

A unique identifier was crafted for each node within an Ethernet Segment. This identifier was then communicated to every node in the same EVPN instance. In the context of SRv6, this is achieved by using the End.DT2M behavior defined in RFC 9252 and utilizes the SRv6 Argument field within the SID to manage BUM traffic forwarding in an E-LAN active/active multi-homing setup. This argument provides a localized mapping to the ESI, enabling the SRv6 network to perform split-horizon filtering effectively.

In our test, we verified that when a multihomed PE received BUM traffic that the other PE had relayed from CE, it identified its unique identifier within the traffic. It withheld the traffic from being re-broadcast into the access layer.

EVPN VPWS Single-Homing over SRv6

We conducted verification of point-to-point EVPN VPWS (Virtual Private Wire Service) over an SRv6 infrastructure and involved a significant number of participants.

This required a comprehensive mesh verification to ensure every possible point-to-point connection was tested. Our method included validating the correct route installation between nodes by configuring loop-back interfaces on the terminal devices and initiating ping tests between their IP addresses. These pings traversed the L2EVPN service, confirming connectivity and route accuracy.

However, one specific setup did not establish an EVPN VPWS session. The issue was in the EVPN Route Type 1 message sent by this device, which used MPLS encapsulation. This encapsulation type is incorrect for our SRv6 context, leading to the failure of the session establishment.

Successful test combinations are listed in Table 9.

	Ciena 5169	Cisco ASR-9902	Cisco N540-28Z4C	Cisco NCS-57B1-6D24	Ericsson R6676	H3C CR16010E-F	H3C S12500R-2L	Huawei ATN 910C-G	Huawei NetEntine 8000 X4	Juniper MX304	Nokia 7250 IXR-e2	Nokia 7750 SR-1	ZTE ZXR10 M600-4SE
Ciena 5169													
Cisco ASR-9902	✓												
Cisco N540-28Z4C	✓	✓											
Cisco NCS-57B1-6D24	✓	✓	✓										
Ericsson R6676	✓	✓	✓	✓									
H3C CR16010E-F		✓	✓	✓	✓								
H3C S12500R-2L		✓	✓	✓	✓	✓							
Huawei ATN 910C-G	✓	✓	✓	✓	✓	✓	✓						
Huawei NetEntine 8000 X4	✓	✓	✓	✓	✓	✓	✓	✓					
Juniper MX304	✓	✓	✓	✓	✓	✓		✓	✓				
Nokia 7250 IXR -e2		✓	✓	✓	✓	✓	✓	✓	✓	✓			
Nokia 7750 SR-1		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
ZTE ZXR10 M6000-4SE	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	

Table 9: EVPN VPWS Single-Homing Test Pairs

RT5 over SRv6

EVPN IP Prefix Route (Route Type 5), defined in RFC 9136, specifies the use of EVPN for Layer 3 or inter-subnet connectivity services.

We confirmed that Layer 3 information was carried correctly within EVPN RT5 by examining the route attributes and the IPv4/IPv6 routing tables on the PEs, which had the correct entries corresponding to the EVPN RT5 routes.

The participating devices were: Arista 7280R3, Cisco 8201-24H8FH, Cisco ASR-9902, Cisco N540-28Z4C, Cisco NCS-57B1-6D24, H3C CR16010E-F, H3C S12500R-2L, Huawei NetEngine 8000 X4, Juniper ACX7024, Juniper ACX7100-48L, Juniper MX204, Juniper MX304, Nokia 7250 IXR, Nokia 7750 SR-1, ZTE ZXR10 M6000-4SE, and ZTE ZXR10 M6000-8SE.

Global IPv4/IPv6 over SRv6

We confirmed that the egress PE can advertise an SRv6 Service SID for IPv4 and IPv6 prefixes in the BGP global routing table. The ingress PE then encapsulated the IPv4/IPv6 payload in IPv6, and copied the SRv6 Service SID as specified by the egress PE to the outer IPv6 header destination address.

According to RFC 9252, the SRv6 Endpoint Behavior was set to one of the options: End.DT4/6 with NEXT-CSID (uDT4/6) or End.DT46 with NEXT-CSID (uDT46). Utilizing SRv6's µDT46 PEs removed the outer IPv6 header and then looked up the IPv4 or IPv6 destination address in the global routing table to process the inner packet.

The same logical and physical test topology was used as shown in Figure 48. In this topology, the following routers were tested successfully for Global Routing Table support (IPv4 and IPv6) over SRv6:

Cisco 8201-24H8FH, Cisco ASR-9902, Cisco N540-28Z4C, Cisco NCS-57B1-6D24, H3C CR16010E-F, H3C S12500R-2L, Huawei NetEngine 8000 X4, Juniper ACX7024, Juniper ACX7100-48L, Juniper MX204, Juniper MX304, Keysight IxNetwork, Nokia 7250 IXR-e2, Nokia 7750 SR-1, ZTE ZXR10 M6000-4SE, and ZTE ZXR10 M6000-8SE.

Prefix Summarization over SRv6

Prefix summarization is one of the key advantages of SRv6. Unlike SR-MPLS, SRv6 maximizes network scale by enabling simpler routing designs and IP route summarization between areas/domains.

Summarizing involves condensing SRv6 locator blocks at the boundaries of each domain. These summarized locators are then distributed to adjacent domains. Summarizing and redistributing/leaking allows any two nodes within the network to establish reachability with thousands less IGP routes. Summarization ensures that the network remains scalable and manageable even as it grows.

In our tests, SRv6 Locator and IPv6 loopback prefix summarization was executed in both directions, and different ABRs were utilized in each test run. Some ABRs were set up to advertise Locator summary advertisements via both the IP Prefix Reachability TLV and the SRv6 Locator TLV. and others only distributed summaries through the IP Prefix Reachability TLV. This didn't raise any issues of resolving service routes from PEs.

One finding from the test revealed that each ABR implemented the SRv6 locator summary injected differently (Internal Down, External Up, External Down, Internal Up), potentially leading to challenges such as load balancing across two ABRs (load balancing is only possible between internal routes, or external routes, but not between internal and external route). When setting up TI-LFA backup (where one ABR is supported through TI-LFA by another ABR), there might be issues with TI-LFA functionality when the primary internal route needs to be backed up by an external route. To avoid these issues, a possible solution would be to define the route types during the summary advertisement; this was not investigated in the testing.

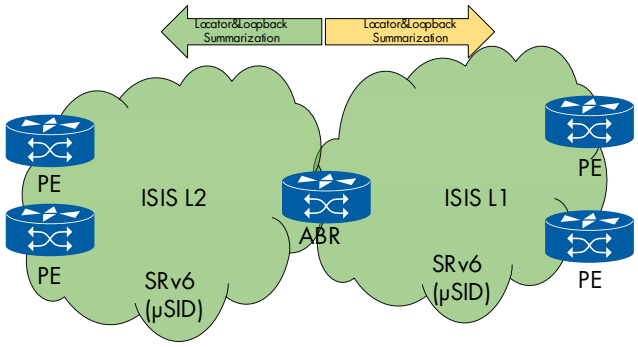


Figure 51: SRv6 Locator & Loopback Summarization

The following devices completed the test as ABR node: Ciena 5169, Cisco 8201-24H8FH, Ericsson R6676, H3C CR16010E-F, H3C S12500R-2L, Huawei NetEngine 8000 X4, Juniper ACX7100-48L, Nokia 7750 SR-1, and ZTE ZXR10 M6000-4SE.

The following devices completed the test as PE Node: Arista 7280R3, Ericsson R6676, H3C CR16010E-F, H3C S12500R-2L, Huawei ATN 910C-G, Huawei NetEngine 8000 X4, Keysight IxNetwork, Nokia 7250 IXR-e2, Nokia 7750 SR-1, ZTE ZXR10 M6000-4SE, and ZTE ZXR10 M6000-8SE.

Unreachable Prefix Announcement (UPA)

Leveraging Summarization in SRv6 maximizes network scale but on the other hand it suppresses individual prefix state that is useful for triggering fast-convergence mechanisms outside of the IGPs - e.g., BGP PIC Edge. The IGP Unreachable Prefix Announcement (UPA) solution defined in the IETF draft raft-ietf-lsr-igp-ureach-prefix-announce describes how to use existing IGP protocol mechanisms to advertise the "loss" of prefix reachability to an individual prefix covered by a summary route. This enables fast convergence away from paths to the node that owns the prefix which is no longer reachable.

In the test, we validated the process for signaling the loss of prefix reachability with an UPA. The network setup consisted of an ABR handling the summary, an Ingress PE, and two Egress PEs.

When the ABR could not connect to a node in the second domain, it recognized that this node's locator was included in the summary prefix. A UPA was created for this locator and sent within the first domain.

After triggering the failure of an egress PE, and upon receipt of the corresponding UPA at the ingress PE node via IS-IS, the BGP Prefix Independent Convergence (PIC) backup path for routes learned from the failed egress PE was activated, and traffic switched to the remaining egress P.

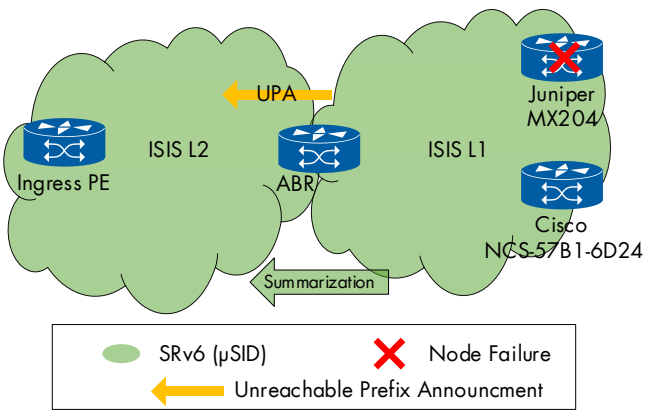


Figure 52: Unreachable Prefix Announcement over SRv6

ABR	Ingress PE	Egress PE
Nokia 7750 SR-1	Cisco N540-28Z4C	Cisco NCS-57B1-6D24, Juniper MX204
Cisco 8201-24H8FH	Nokia 7250 IXR-e2	
Nokia 7750 SR-1	H3C CR16010E-F	
Nokia 7750 SR-1	H3C S12500R-2L	

Table 10: Unreachable Prefix Announcement (UPA), Test Pairs

The successfully tested router combinations for Unreachable Prefix Announcements are shown in Table 10.

One participant could not promptly advertise the UPA, causing the ingress node to depend on BGP convergence for detecting the prefix loss; therefore, their results were not included in the report.

SRv6 TE Policies with Explicit Paths

In SRv6, an SR policy (defined in IETF RFC 9256) can steer traffic over a desired path through the network. It instructs the routers in the network to follow the specified path instead of following the shortest path calculated by the IGP. If a packet is steered into an SRv6-TE policy, the head-end pushes the SID list on the packet. The rest of the network executes the instructions embedded in the list.

During our evaluation, we conducted tests on SRv6 TE policies that utilized explicit paths. Concurrently, we explored automated steering on a per-destination basis. This was aligned with the color and next-hop defined by the service route advertised by the egress node, allowing for color-based Automated Steering into an SRv6 Policy.

The SRv6 Policy implemented on the PE router was configured to use an explicit segment list containing x unique node μSIDs, designed to dictate the precise path across x SIDs before reaching the destination at the egress PE.

The test topology was identical with the "L2/L3 Services over SRv6" topology shown in Figure 48. The following routers participated in the test as SRv6 Policy Headends:

- Ciena 5169
- Cisco 8201-24H8FH
- H3C CR16010E-F, H3C S12500R-2L
- Huawei NetEngine 8000 X4, ATN910C-G
- Juniper ACX 7024, Juniper MX 204
- Nokia 7750 SR-1
- ZTE ZXR10 M6000-4SE

All devices in the topology diagram (Figure 48) participated as transit nodes, except the Arista 7280R3.

The architecture of the packet was such that it had an outer IPv6 header, the destination address which served as a container for six μSIDs (F3216 μSID format was used), detailing the sequence of segments to be traversed.

When more than six segments were defined, the packet also had an SR Routing Header (SRH) housing the remaining segments in the Segment List. Each of these segments acted as µSID containers that can encapsulate six more µSIDs.

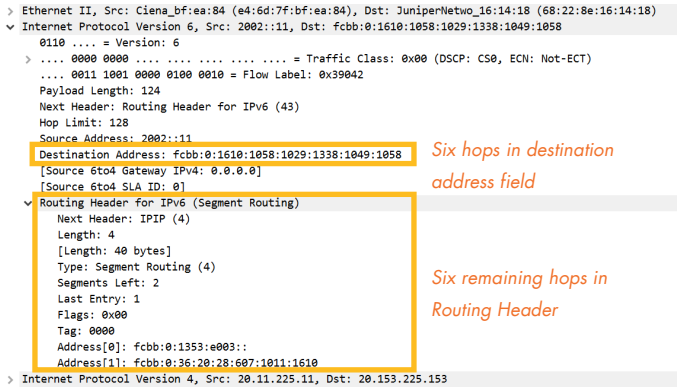


Figure 53: SR Policy Headends test: Packet capture highlighting the placement of the µSID list

Every node along the route was required to execute a shift and forward operation on the µSID list, and upon reaching the final µSID in the destination IP address, the routers had to transfer a new set of µSIDs from the SRH to the Destination Address field of the IPv6 header. Because not all vendors could process the SRH, the segment list was constructed specifically to avoid these nodes as the sixth µSID.

Flexible Algorithm over SRv6

IGP protocols historically compute the best paths over the network based on the IGP metric assigned to the links. On the other hand, IGP Flexible Algorithm (FA) (defined by IETF RFC 9350) enables network operators to tailor the calculation of the IGP's shortest path to suit their particular requirements and preferences. This is done by allowing users to choose a metric (IGP, TE, Delay) and constraints (e.g. Administrative Group). FA uses Prefix-SIDs (SR-MPLS) and SRv6 locators (SRv6) to steer packets along the constraint-based paths.

Our test verified Flex Algorithm over SRv6 by focusing on FA instances with Delay metric (FA 128) and TE metric (FA 129). Certain nodes in the network were equipped with TWAMP for dynamic link delay measurement while others used static delay configuration over the links. The advertisement of the measured link delay was based on IS-IS Traffic Engineering (TE) Metric Extensions (RFC8570).

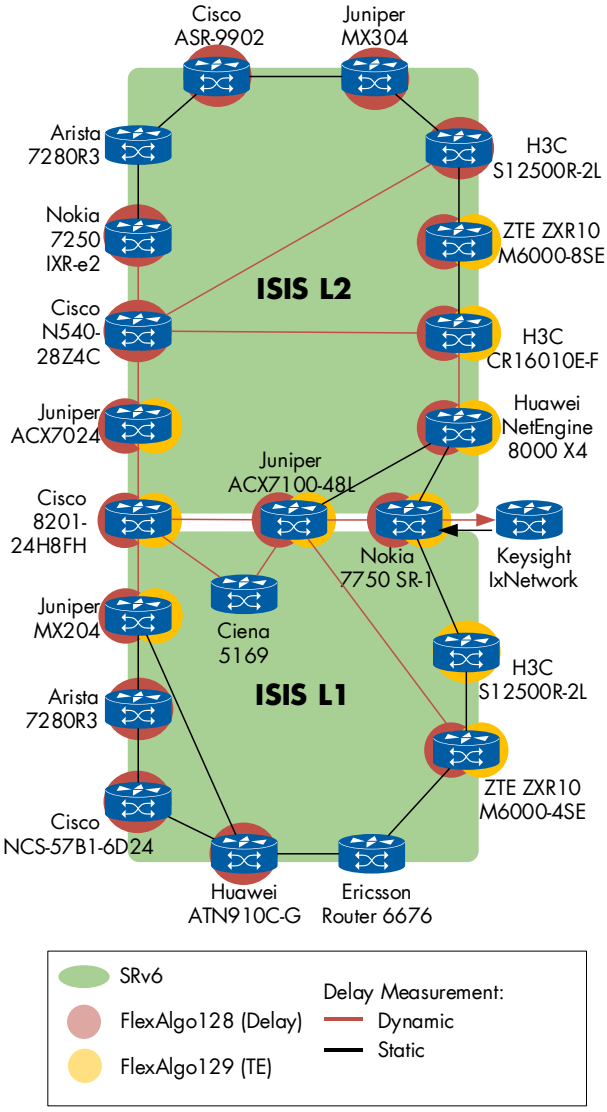


Figure 54: Flexible Algorithms

FlexAlgo Using New Constraints

In this test, we verified new Flex-Algo constraints defined at the IETF: IETF draft draft-ietf-lsr-flex-algo-bw-con specifies additional FA constraints that allow the network administrator to exclude the use of low-bandwidth or high-delay links, so-called min-BW and max-Delay FA constraints. IETF draft draft-ietf-lsr-igp-flex-algo-reverse-affinity specifies an additional FA constraint that allows the inclusion/exclusion of interfaces based on the link admin group value in the reverse direction of the traffic flow, so-called Reverse Affinity FA constraint. Following to the methodology from our SR MPLS test, we successfully validated the newly introduced constraints for Flex Algorithms in SRv6 as follows:

- **FA 132 min-BW constraint:** The operator specifies a minimum bandwidth value for the interfaces part of the FA. IGP excludes interfaces with an interface BW below this value.

- **FA 133 max-Delay constraint:** The operator specifies a maximum link delay value for the interfaces part of the FA. IGP excludes interfaces with a link delay above this value.
- **FA 134 exclude Reverse Affinity constraint:** The operator specifies a reverse link admin group value of interfaces to be excluded from the FA. IGP excludes interfaces matching the specified affinity in the reverse direction of the SPF computation.

These Flex-Algos were verified by generating traffic and verifying that forwarding paths aligned with the constraints.

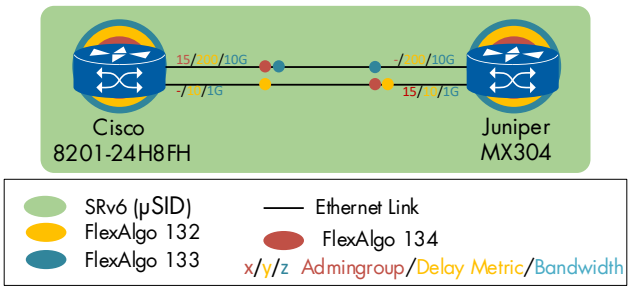


Figure 55: SRv6 FlexAlgo With New Constraints

TI LFA over SRv6

During our evaluation of redundancy failover of the SRv6 topology using TI-LFA, we tested its effectiveness in scenarios that involved µSID local link protection. We noted the failover time and observed acceptable downtimes ranging from 3 to 9 ms.

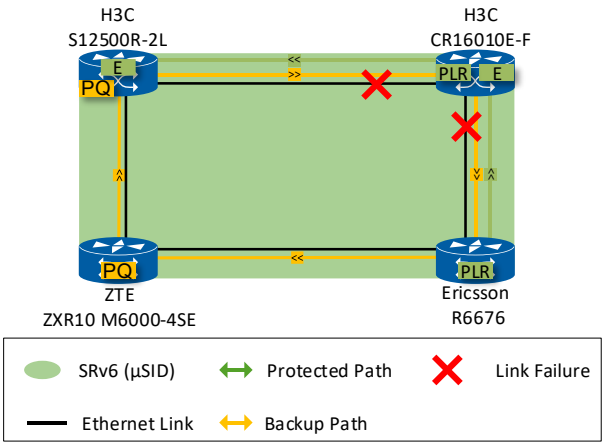


Figure 56: TI-LFA over SRv6 µSID (Local Protection)

SRv6 and SR-MPLS Service Interworking

Our testing included the verification of both L3VPN SRv6/SR-MPLS and L3 EVPN/SRv6 and L3VPN/MPLS Interworking Gateway functionalities.

For L3VPN SRv6/SR-MPLS interworking, we observed the gateway's capability to produce SRv6 VPN SIDs and MPLS VPN labels for all VRF-configured prefixes. The gateway facilitated traffic movement from the MPLS domain to the SRv6 domain by stripping the MPLS VPN label, conducting a destination prefix lookup, and then applying the relevant SRv6 encapsulation. In the opposite direction, it transitioned traffic from the SRv6 to the MPLS domain by removing the IPv6 header, performing a prefix lookup, and affixing the corresponding MPLS VPN and next-hop labels.

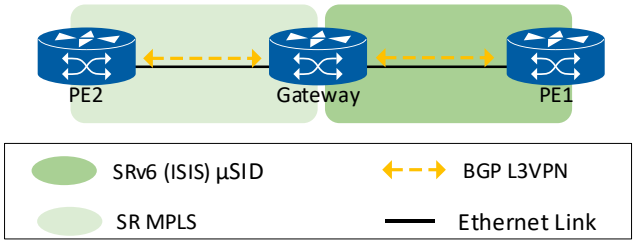


Figure 57: L3VPN SRv6 to SR-MPLS Interworking

The following gateways completed the testing successfully:

- Ericsson Router 6676
- ZTE ZXR10 M6000-4SE, ZXR10 M6000-8SE

The SRv6 PE was implemented by the ZTE ZXR10 M6000-4SE and -8SE, respectively. The SR-MPLS PE was implemented by the Ericsson Router 6676 and the ZXR10 M6000-4SE.

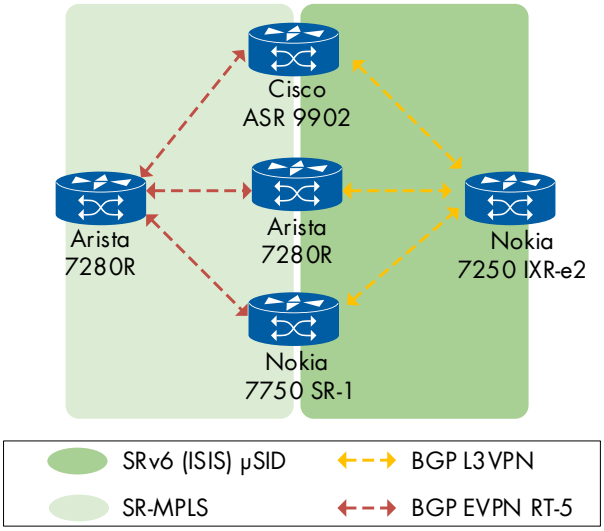


Figure 58: EVPN-RT5 SR-MPLS and L3VPN SRv6 Interworking

In the second part of this test shown in Figure 58 above, we evaluated the interworking between L3 EVPNs over SRv6 on one side, and L3VPN over MPLS on the other side. The gateway's operations were as follows in terms of control plane actions:

- From MPLS to SRv6, the gateway took in routes from the MPLS environment (through EVPN RT5), re-originated them within the L3 EVPN VRF, and associated them with a per-VRF SRv6 SID.
- From SRv6 to MPLS, the gateway imported routes from the SRv6 side (also through EVPN RT5) and re-originated them within the L3VPN VRF, assigning a per-VRF MPLS label.

Path Tracing

Equal-cost multi-path (ECMP) enhances efficiency and resilience in IP networks. On the other hand, it is paramount to manage and troubleshoot them.

The Path Tracing solution, defined in draft-filsfils-ippm-path-tracing-00, unveils detailed insights into the network, providing a record of end-to-end delay, per-hop delay, and load on each egress interface along the packet delivery path. This allows operators to identify current and historical paths, verify packet adherence to these paths, and detect deviations or irregularities.

In this test, Keysight served a dual role: it was the source initiating path-tracing probes and the destination point where these probes ended. The path these probes followed included three Cisco devices, with each interface along the network path uniquely configured for Path Tracing, featuring a distinctive interface ID that remained consistent network-wide.

As the probes made their way through the network, each Cisco router in the topology (see diagram below), upon receiving a probe, attached an IPv6 Hop-by-Hop Option for Path Tracing (HbH-PT) to the packet. This tag, known as Midpoint Compressed Data (MCD), included critical information such as the interface's unique ID, a timestamp, and the interface's current load.

To validate the functionality, we intercepted the probes and scrutinized the tagged data within the packets. This inspection confirmed that the packets indeed carried the MCD information of all three intermediary points, as intended by the test parameters.

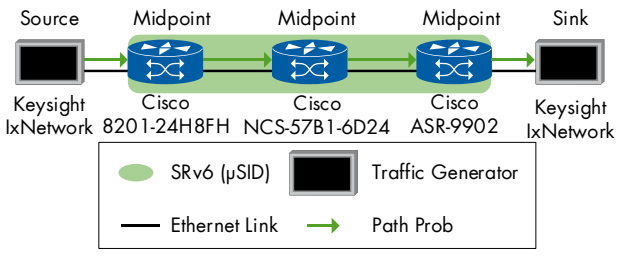


Figure 59: Path Tracing over SRv6

Link Resource Slicing

Link slicing is a technique that helps to distribute the physical bandwidth on links among multiple tenants. It ensures that each tenant gets a minimum bandwidth when there is congestion and also allows the imposition of maximum transmission rates per tenant. This can be achieved using data plane or control plane identifiers.

In this scenario, nodes within the network were configured to support two distinct network slices, each identified by unique IDs. The ingress node assigned incoming traffic to these slices based using SRv6-TE policy. This was achieved by tagging the traffic with a slice ID, which was then embedded into the IPv6 packet header of each packet.

As the traffic progressed through the network, intermediate (middle) and egress nodes were configured to recognize these slice IDs within the packet headers. This enabled them to associate the incoming packets with the corresponding network slices, each predefined with its bandwidth constraints.

To validate this setup, traffic was generated, and monitoring was conducted at an intermediate point within the network to ensure that the slice ID was present in the packet headers, confirming that the slicing mechanism was functioning as intended.

Furthermore, by increasing the traffic volume directed towards one specific slice, it was observed that packet loss occurred indicating that the network was effectively enforcing the bandwidth limitations associated with each slice. This proved that network slicing on network link resources can effectively isolate and guarantee service traffic. The traffic carried on a specific slice is not affected by other traffic. Even if other traffic is discarded due to insufficient bandwidth.

The initial network setup included an extra participant, which shows good interests on the interoperability of link resource slicing. However, it also revealed a divergence in the approach of how the Slice IDs are encapsulated within the IPv6 packet. Despite the guidance provided by IETF documents on Slice ID encapsulation, variations in implementations emerged.

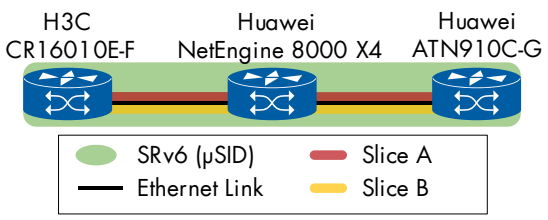


Figure 60: Link Resource Slicing over MSR6

Multicast Source Routing over IPv6 (MSR6)

MSR6 introduces a strategic approach to multicast routing in IPv6 networks, enhancing efficiency and simplifying network design by eliminating the need for routers to maintain per-flow state. This methodology aligns with the evolving demands of modern network infrastructure, aiming to streamline multicast traffic management and improve network performance.

MSR6 OAM

The ping tool could be adapted to test the connectivity and reachability of multicast endpoints, providing insights into the health and efficiency of the multicast routing paths. Traceroute would similarly map the route of multicast packets, helping identify the path and pinpoint any potential routing issues or delays.

We verified the connectivity within the MSR6 network by confirming the proper establishment of a MSR6 tunnel. Then, we conducted ping and traceroute commands within a specific subdomain, successfully testing communication between the leaf nodes and the root.

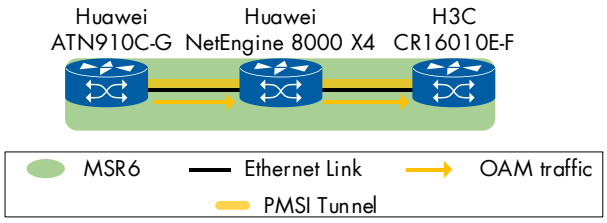


Figure 61: MSR6 Ping and Traceroute

MVPN over MSR6

We examined how MVPN operates over an MSR6 architecture to transport VPN IP multicast traffic. We set up Layer 3 VPNv4 over an SRv6 infrastructure.

Keysight IxNetwork served as both the source and destination for the multicast traffic while performing IGMP joins to mimic real multicast receivers, while Huawei and H3C acted as rooted and leaves.

The test proceeded with the establishment of BGP MVPN peer relationships between the PEs in the same MVPN. This enabled the exchange of MVPN A-D and C-multicast routes.

An Inclusive Provider Multicast Service Interface (I-PMSI) tunnel was created to link all PEs within the MVPN, with the provision to switch to a Selective Provider Multicast Service Interface (S-PMSI) tunnel based on specific configured criteria (data rate).

The multicast traffic initiated by IxNetwork flowed seamlessly to the receiving CEs (emulated by Keysight), affirming that the I-PMSI tunnel was correctly established and operational. Furthermore, the S-PMSI tunnel was successfully engaged when the conditions were met, demonstrating the network's effective handling of multicast group dynamics.

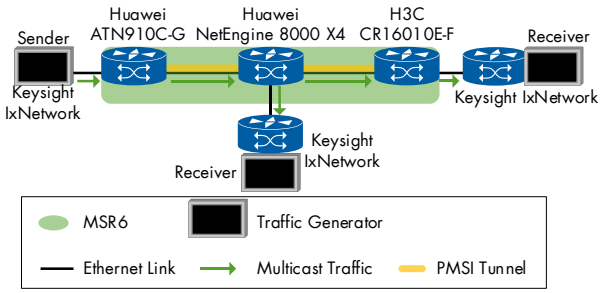


Figure 62: MVPN over SRv6

MSR6 dual root protection

As part of our MSR6 tests, we evaluated the protocol's link protection capabilities. We initiated multicast traffic streams from the sender node towards two roots.

The multicast traffic was simultaneously transmitted through two PSMI tunnels. The receiving leaf node was configured to accept the multicast stream from the primary tunnel linked to the master root while discarding the stream from the secondary tunnel linked to the backup root. We confirmed that the leaf node was receiving traffic from both root nodes, ensuring it delivered only a single stream to the receiver.

In the event of a failure at the master root, the leaf node was designed to swiftly identify the tunnel disruption through flow-based dual-root protection and switch to accepting the multicast flow from the secondary tunnel. We verified that the leaf node executed a switchover from the primary to the backup tunnel and the multicast traffic continued to reach the receiver.

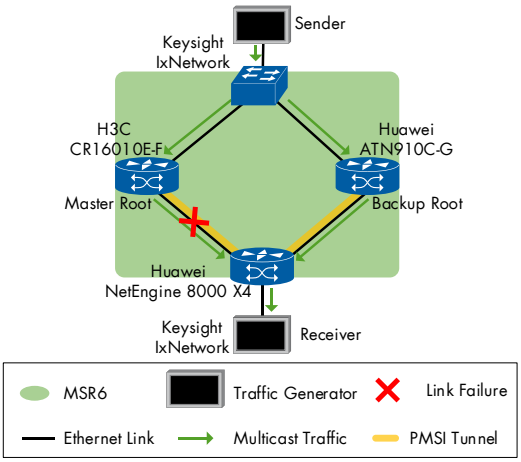


Figure 63: Dual Root Protection over MSR6

SDN Interoperability Test Results

Software-defined networking (SDN) evolves dynamically, driven by its reliability and efficiency in network management.

This year, we focused on key protocols such as PCEP, NETCONF, and BGP-TE/BGP-LS to enhance the SR-MPLS and SRv6-based data and control planes. We observed increased support for IPv6 in PCEP sessions and advancements in flexible algorithm discovery and uSID topology visualization. Vendors noticeably increased support for segment routing policy computation and signaling, including color implementation.

We conducted new test cases, including latency-based optimization and PCEP association groups, and revisited past test cases like path computation and L2/L3 VPN provisioning via NETCONF, integrating new features or vendors.

We've noticed broader support for the PCEP extensions to signal SR Policy identifiers (such as Color) as well as SR Policy Candidate Paths and their attributes (e.g., Preference). In addition, we observed more participating vendors that supported both PCC-initiated/PCE-delegated and PCE-initiated instantiation models.

An SDN controller is a centralized entity in software-defined networking that manages the network's traffic paths and resources. It implements network policies and can dynamically adjust to changing network conditions, thereby optimizing network performance and reliability. An additional significant benefit of the SDN controller is its ability to provide a comprehensive visual representation of the network topology. This feature greatly assists network administrators in understanding the overall structure and dynamics of the network, enabling them to detect and address any potential issues swiftly.

Screenshots taken during this year's testing showcase the network topology visualization on the participating SDN controllers (Figures 64-67).

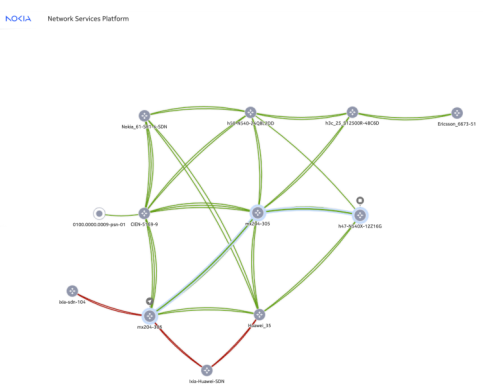


Figure 64: Nokia Network Service Platform (NSP)

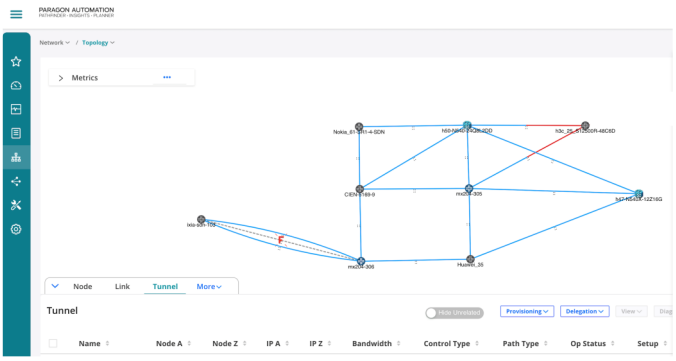


Figure 64b: Juniper Paragon

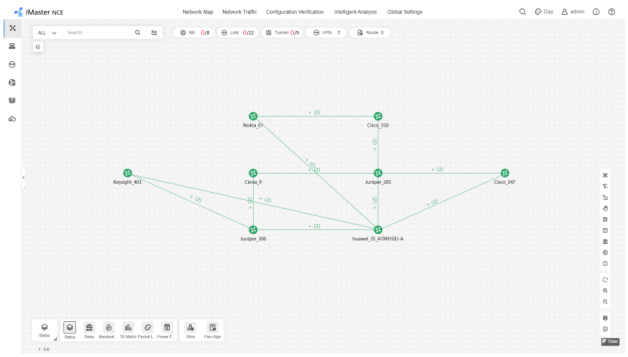


Figure 65: Huawei iMaster NCE-IP

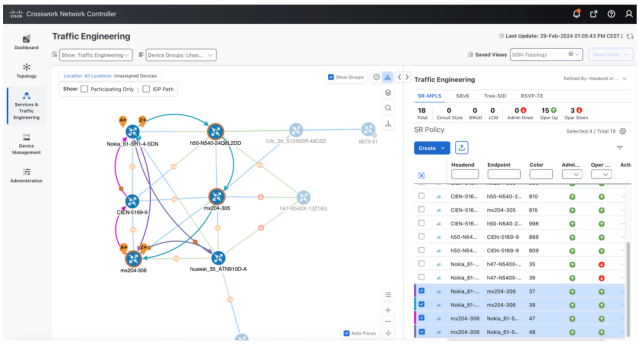


Figure 66: Cisco Crosswork Network Controller

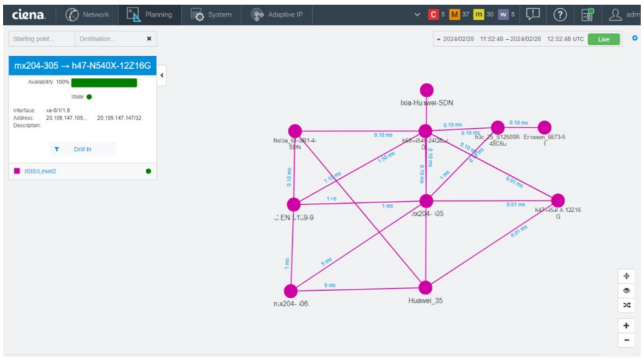


Figure 67: Ciena Navigator (NCS)

Colored Path Computing and Signaling

The primary role of the path computation element (PCE) is to compute and signal a network path to the path computation client, particularly as networks increasingly transition from traditional IP/MPLS to segment routing, encompassing both SR-MPLS and SRv6. The urgency to establish a standardized approach for path computation and signaling in segment-routed networks has grown. An SR Policy, outlined in RFC 9256, consists of various SR Candidate Paths with the same identifying tuple. The IETF draft titled "PCEP extension to support Segment Routing Policy Candidate Paths" (draft-ietf-pce-segment-routing-policy-cp) expands on RFC 8664 to fully accommodate the SR Policy framework. Our evaluation focused on the interoperability between the PCE and the PCC in terms of requesting and signaling an SR policy across different vendors. The tests covered a range of scenarios, including both PCC and PCE-initiated paths, SR-MPLS and SRv6 data planes, and SR policies with and without color.

- The PCCs established IGP (IS-IS L2) adjacencies
- We confirmed the status of the PCEP session between the PCE and the PCC and the Traffic Engineering Database (TED) Synchronization
- The PCE retrieved and visualized the topology
- A path initiation is triggered either from the PCC or the PCE
- We verified the computed path using an L3VPN-steered traffic over the computed path.

Tables 11 and 12 present the vendor combinations that interoperated using PCEPv4 to signal Segment Routing Traffic Engineering (SR-TE) policies to a headend within an SR-MPLS network. The test setup used is shown in Figure 68. These tests incorporate both the PCC-initiated/PCE-delegated and the PCE-initiated instantiation models.

Additionally, we conducted a test using BGP to signal PCE-initiated SR-TE policy to a headend on an SR-MPLS network, where Huawei iMaster NCE-IP served as PCE, and Cisco N540X-12Z16G served as PCC.

PCE	PCC	Colored
Ciena Navigator NCS	Juniper MX204	Yes
Cisco Crosswork Network Controller	Juniper MX204	Yes
	Huawei ATN910D-A	Yes
	Nokia 7750 SR-1	No
Huawei iMaster NCE-IP	Cisco N540X-12Z16G	Yes
	Juniper MX204	Yes
Juniper Paragon Applications	Cisco N540X-12Z16G	Yes
	Huawei ATN910D-A	Yes
	Nokia 7750 SR-1	No
Keysight IxNetwork	Juniper MX204	Yes
Nokia Network Service Platform (NSP)	Cisco N540X-12Z16G	No
	Juniper MX204	No

Table 11: PCE-Initiated Segment Routing Policy Signaling with SR-MPLS

PCE	PCC	Colored
Ciena Navigator NCS	Cisco N540-24Q8L2DD	Yes
Cisco Crosswork Network Controller	Ciena 5169	Yes
	Juniper MX204	Yes
	Huawei ATN910D-A	Yes
	Nokia 7750 SR-1	No
Huawei iMaster NCE-IP	Cisco N540X-12Z16G	Yes
	Juniper MX204	Yes
Juniper Paragon Applications	Cisco N540X-12Z16G	Yes
	Nokia 7750 SR-1	No
Keysight IxNetwork	Juniper MX204	Yes
Nokia Network Service Platform (NSP)	Cisco N540X-12Z16G	No
	Juniper MX204	No

Table 12: PCC-Initiated Segment Routing Policy Signaling with SR-MPLS

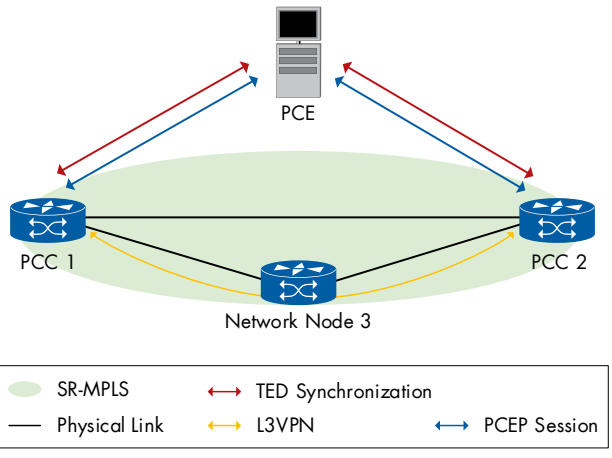


Figure 68: SR Policy Signaling with SR-MPLS

Our testing also included two PCEP tests for signaling

SR-TE policies to a headend on an SRv6 network, following the PCC-initiated/PCE-delegated model. In a particular scenario, PCEPv6 was used at the headend on an SRv6 uSID network, while PCEPv4 was employed at the headend on an SRv6 full SID network. The test setup, as well as the participating devices, are shown in Figures 69 and 70 below.

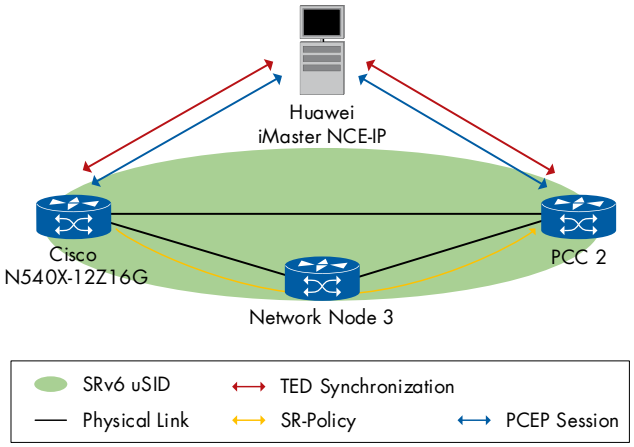


Figure 69: SR Policy Signaling with SRv6 µSID

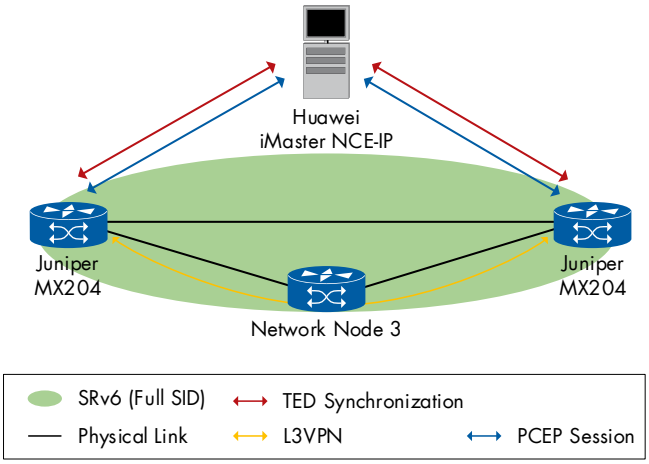


Figure 70: SR Policy Signaling with SRv6 Full SID

The verification process involved validating the SR policy status on both the PCC and PCE nodes and confirming the correct steering of the traffic of an L3 VPN service over the signaled SR policies. In the tests involving BGP and SRv6-uSID, the policy was signaled only to the headend PCC, which restricted our verification methods to checking the policy status on both the PCC and the PCE rather than using L3VPN or LSP ping.

We encountered several challenges. For one particular combination, the L3VPN could not be configured correctly; thus, LSP Ping was utilized to verify the correct instantiation of the SR policy. In another unlisted test run, no PCEP packets were transmitted from the PCE to the PCC, preventing the completion of the test.

PCC—Dynamic Paths Instantiation

The ability of the path computation client (PCC) to adapt to real-time network conditions is a significant advantage of SDN networks. In this test, we evaluated the PCC's ability to request a path from the path computation element (PCE) upon receiving a new L3VPN route from a PE router, marking them with a specific color extended community. This process activated on-demand segment routing policies calculated at the headend PCC to reduce latency towards the BGP next hop. After establishing the routing policy, the PCC reported the outcome to the PCE.

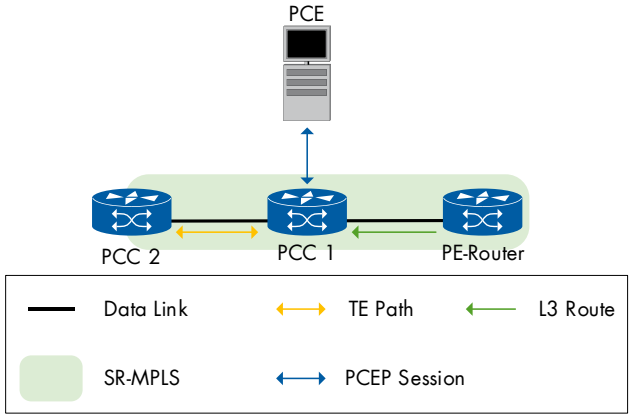


Figure 71: PCC Dynamic Paths Instantiation

There were two successful test combinations:

- PCC: Ciena 5169
  PCE: Cisco Crosswork Network Controller
  PE: Cisco N540X-12Z16G
- PCC: Cisco N540X-12Z16G
  PCE: Ciena Navigator NCS
  PE: Ciena 5169

Advertisement of SR Policies using BGP-LS

This test evaluated the support of BGP-LS extensions defined in the draft "Advertisement of Segment Routing Policies using BGP Link-State" (draft-ietf-idr-bgp-ls-sr-policy) by both the segment routing traffic engineering (SRTE) head-end node and the PCE. It aimed to assess the effectiveness of BGP-LS in reporting traffic engineering information, such as SR policies, to the PCE. The test steps included verifying BGP-LS connectivity and monitoring the policy status reported from the PCC to the PCE, demonstrating the system's capability to maintain clear and accurate TE data communication. The test topology is detailed in Figure 72 on the next page.

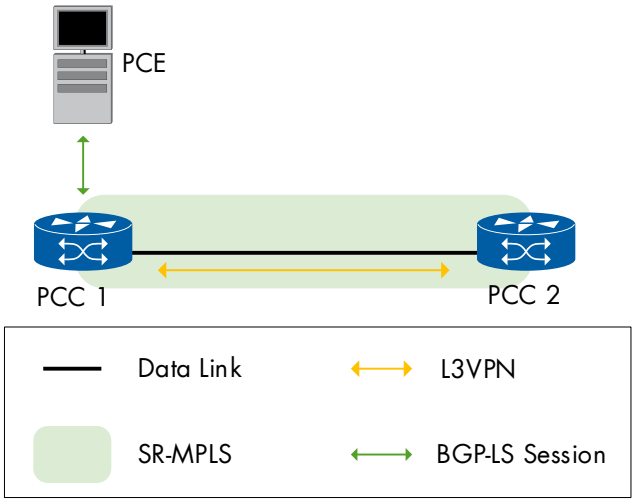


Figure 72: Advertisement of SR Policies using BGP-LS with SR-MPLS

This test succeeded with three PCE/PCC combinations:

- PCE: Cisco IOS XRd
  PCC: Huawei ATN910D-A
- PCE: Huawei iMaster NCE-IP
  PCC: Cisco N540X-12Z16G
- PCE: Keysight ixNetwork
  PCC: Cisco N540X-12Z16G

PCEPv6

PCEP operates fundamentally over TCP. Once the TCP session is established, a PCEP session between the PCE and PCC is constructed on this connection. With the evolving network infrastructure shifting towards IPv6 for its extended addressing capabilities and enhanced security features, ensuring that these sessions are fully operable in an IPv6 environment is crucial. In this test, we confirmed that the PCEP session is fully operable in an IPv6 environment.

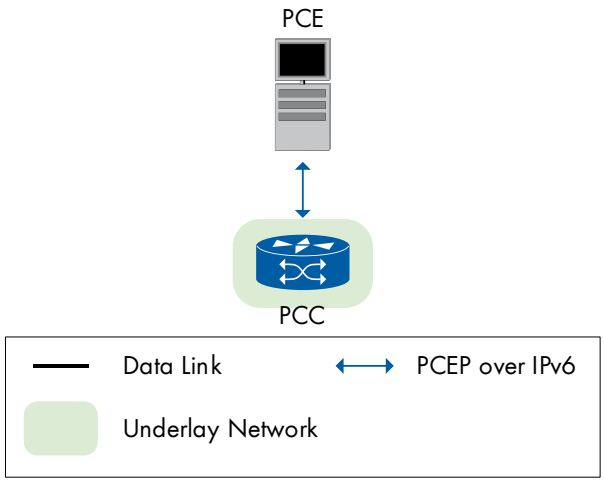


Figure 73: PCEPv6

Once the PCEP session was established over IPv6, we monitored the PCE and PCC session statuses to validate that the sessions were set up correctly. The PCE and PCC pairs that interoperated in this test are:

PCE	PCC
Cisco IOS XRd	Huawei ATN910D-A
Cisco IOS XRd	Keysight IxNetwork
Huawei iMaster NCE-IP	Cisco N540X-12Z16G
Huawei iMaster NCE-IP	Keysight IxNetwork
Keysight IxNetwork	Huawei ATN910D-A
Keysight IxNetwork	Cisco N540X-12Z16G

Table 13: PCEPv6 Test Combinations

Latency-Based Optimization

In today's networks, application-specific performance is becoming more and more of a critical metric and network optimization goal. Performance is not limited to the bandwidth but more towards the sensibility of a user action in real-time. This is significant in financial networks, online conferencing, and many other applications. PCEP provides a mechanism to compute end-to-end paths based on multiple metrics. In this test, we verified the path computation element's (PCE) ability to initiate and optimize the paths with consideration of the latency on the different possible paths. The devices in the topology advertised the IGP latency metric. The PCE calculated a path prioritizing the lowest cumulative latency. To test latency-based optimization, we increased the IGP metric on the chosen links of this path. Following this adjustment, we monitored how the PCE recalculated and signaled a new optimized route that again adhered to the principle of minimal cumulative latency.

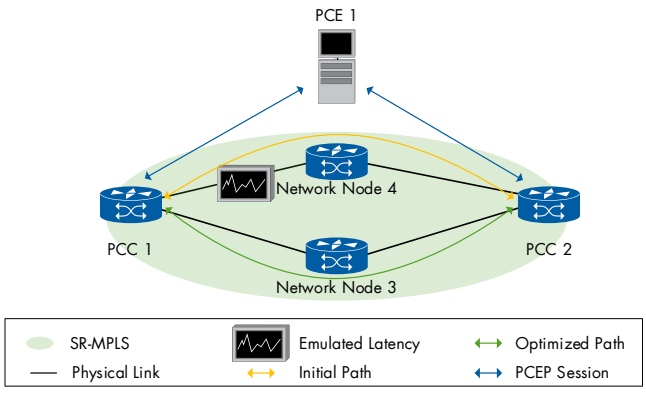


Figure 74: Latency-Based Optimization

Three combinations completed the test successfully:

- PCE: Ciena Navigator NCS
  PCC: Cisco N540X-12Z16G
- PCE: Cisco Crosswork Network Controller
  PCC: Ciena 5169
- PCE: Nokia Network Service Platform (NCS)
  PCC: Juniper MX204

In one test run, we encountered an issue where the PCE did not adjust the path dynamically following the change in latency. However, it was able to compute a path with the lowest latency during the initial path computation.

Flexible Algorithm Discovery and Visualization

This test verified the BGP-LS extensions for Flexible Algorithm Advertisement (RFC 9351). These extensions enable the advertisement of IGP Flexible Algorithm Definition (FAD) and Application-Specific Link Attributes (ASLA) as part of the network's topology information.

PCE	PCC	Data
Ciena Navigator NCS	Cisco N540X-12Z16G	SR-MPLS
	Nokia 7750 SR-1	SR-MPLS
	Keysight IxNetwork	SR-MPLS
Cisco Crosswork Network Controller	Huawei ATN910D-A	SRv6
	Nokia 7750 SR-1	SRv6
	Keysight IxNetwork	SRv6
	Huawei ATN910D-A	SR-MPLS
	Nokia 7750 SR-1	SR-MPLS
	Keysight IxNetwork	SR-MPLS
	Ciena 5169	SR-MPLS
Huawei iMaster NCE-IP	Cisco N540X-12Z16G	SRv6
	Nokia 7750 SR-1	SRv6
	Keysight IxNetwork	SRv6
Keysight IxNetwork	Cisco N540X-12Z16G	SRv6
	Nokia 7750 SR-1	SRv6
	Huawei ATN910D-A	SRv6
	Cisco N540X-12Z16G	SR-MPLS
	Nokia 7750 SR-1	SR-MPLS
	Ciena 5169	SR-MPLS

Table 14: FlexAlgo Discovery Test Combinations

We confirmed the PCE's capability to accurately discover a Flexible Algorithm (FA) instance, including its definition, the participating nodes, and their attributes, such as prefix-SIDs and link attributes. We also visualized the FA's topology.

In the test, one PCE was connected to one PCC, including a BGP-LS setup between PCE and PCE. SR-MPLS and SRv6 scenarios were tested likewise. The test combinations are shown in Table 14.

SRv6 µSID Topology Discovery and Visualization

The BGP-LS address family extensions play a crucial role in advertising SRv6 information in a multivendor environment, distributing SRv6 segments, their behaviors, and other related data across all SRv6-capable nodes.

Building on this foundation, this test evaluated the PCE/Controller's capability to retrieve this propagated SRv6 information. Specifically, the SRv6 uSID segments, SRv6 prefixes (locators), and the attributes of nodes and links.

The test topology was straightforward, as before: In each combination, one PCE was connected to one PCC using BGP-LS. The PCC used SRv6 with µSIDs. Successful test combinations were:

PCE	PCC
Cisco Crosswork Network Controller	Keysight IxNetwork
	Huawei ATN910D-A
	Nokia 7750 SR-1
Huawei iMaster NCE-IP	Cisco N540X-12Z16G
	Nokia 7750 SR-1
	Keysight IxNetwork
Keysight IxNetwork	Cisco N540X-12Z16G
	Nokia 7750 SR-1
	Huawei ATN910D-A

Table 15: SRv6 µSID Topology Discovery

PCEP Association Group: Diversity

When redundancy is required, it is necessary to establish two disjoint (diverse) paths between two nodes without any common links, nodes, or SRLG groups. In this test, the path computation client (PCC) requested disjoint paths from the path computation element (PCE), which then computed the paths and signaled them to the PCC.

This test confirmed that the Path Computation Element with PCEP can effectively compute and signal diverse paths to Path Computation clients in a multi-vendor environment.

Executing the test, we followed this procedure:

- The DUTs initiated the IGP adjacencies
- We validated the PCEP session, PCE path instantiation, and LSP state synchronization
- The PCC requested disjointed paths from the PCE

We checked the PCCs' configuration and the PCE interface to verify the correct computation and signaling of the disjointed paths.

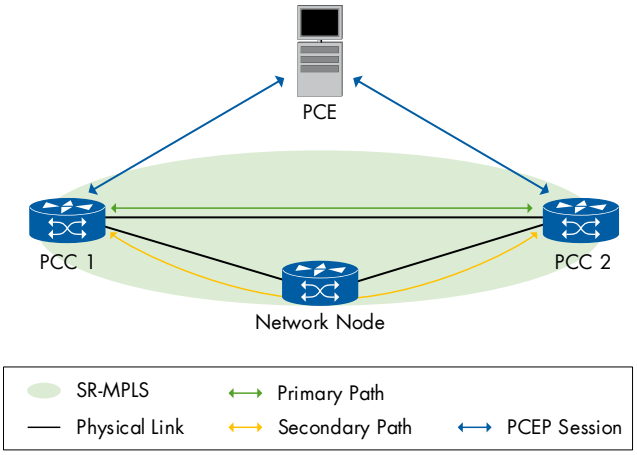


Figure 75: Disjoint Paths Computation

The following combinations of devices participated:

PCE	PCC
Cisco Crosswork Network Controller	Nokia 7750 SR-1
	Juniper MX204
	Cisco N540X-12Z16G
Nokia Network Service Platform (NSP)	Juniper MX204
	Cisco N540X-12Z16G
	Nokia 7750 SR-1

Table 16: Results of PCEP Association Group: Diversity

PCEP Association Group: Policy

In software-defined networks, the Path Computation Element (PCE) computing a path is a fundamental operation ensuring the separation between the data and control planes. Path computation typically relies on various metrics, such as Interior Gateway Protocol (IGP) or latency metrics. PCEP Association Groups enable the PCE to predefine multiple policies that can be selected and applied while computing new paths.

In this context, a policy might be defined as a group of configuration parameters that either a Path Computation Client (PCC) or PCE (PCEP speakers) can apply to an LSP or a group of LSPs. Specifically, in the test scenario we conducted, the policy associated with an LSP or group of LSPs included configuration parameters like the optimization objective (latency), maximum latency (SLA), and threshold latency (to trigger an optimization task). The PCC requested a path and specified the desired policy from the PCE. Subsequently, the PCE computed a path that adhered to the policy constraints and returned it to the PCC, demonstrating the effectiveness of applying these policy parameters in real-world network management.

The test combination shown in Figure 76 successfully completed this test scenario.

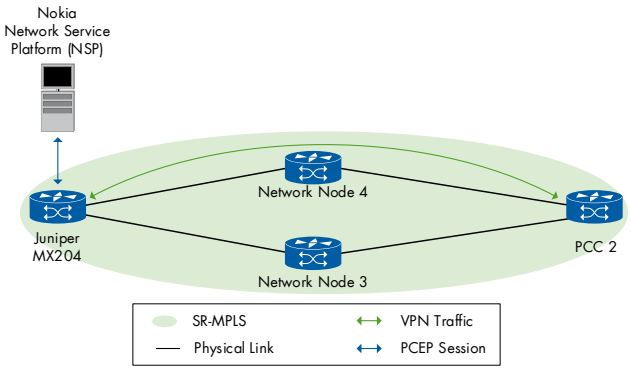


Figure 76: Setup for PCEP Association Group: Policy

PCEP Binding SID

The draft "Carrying Binding Label/Segment Identifier (SID) in PCE-based Networks" proposes a significant improvement for managing binding labels or SIDs in networks using segment routing. It introduces a method whereby the binding label/SID can minimize the stack depth of SIDs required at nodes, especially access nodes with constrained forwarding capacities. In this scheme, a Midpoint/Gateway PCC node located in an intermediate network position modifies the SID list it receives from the headend PCC by appending necessary SIDs for the path to the tailend.

The PCE's role is to provide the Binding SID (BSID) and the identity of the Midpoint/Gateway node to the headend PCC. This information is needed by the headend to direct the traffic toward the tailend, ensuring efficient path utilization and reducing the overhead on nodes with limited SID processing capabilities.

In this test, we signaled the binding SID to the Midpoint/Gateway PCC but didn't proceed with the path instantiation on the headend PCC.

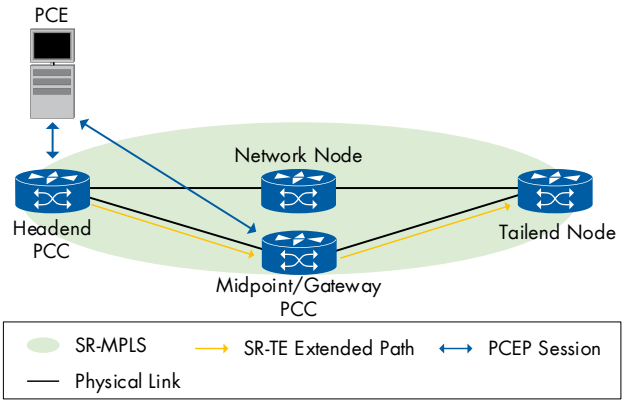


Figure 77: PCEP Binding SID

We verified the test by checking the configuration on the Midpoint/Gateway PCC, showing the BSID and the LSP towards the tailend.

The Nokia Network Service Platform (NSP) functioned as PCE in this test, and successfully completed tests with midpoint/gateway PCCs Nokia 7750 SR-1 and Juniper MX204.

NETCONF Transport Slicing Controller

In this test, the network controller, acting as the Transport Slice Manager and following the guidelines of the draft "IETF Network Slice Service YANG Model" (draft-ietf-teas-ietf-network-slice-nbi-yang-02), used NETCONF to establish a transport slice. It integrated IETF-defined YANG models with SR and SR-TE policies communicated via PCEP. The setup featured IETF-L3NM Route Policy and SR-TE Policy with QoS configurations for L3VPN, ensuring the slice met targeted performance and routing standards.

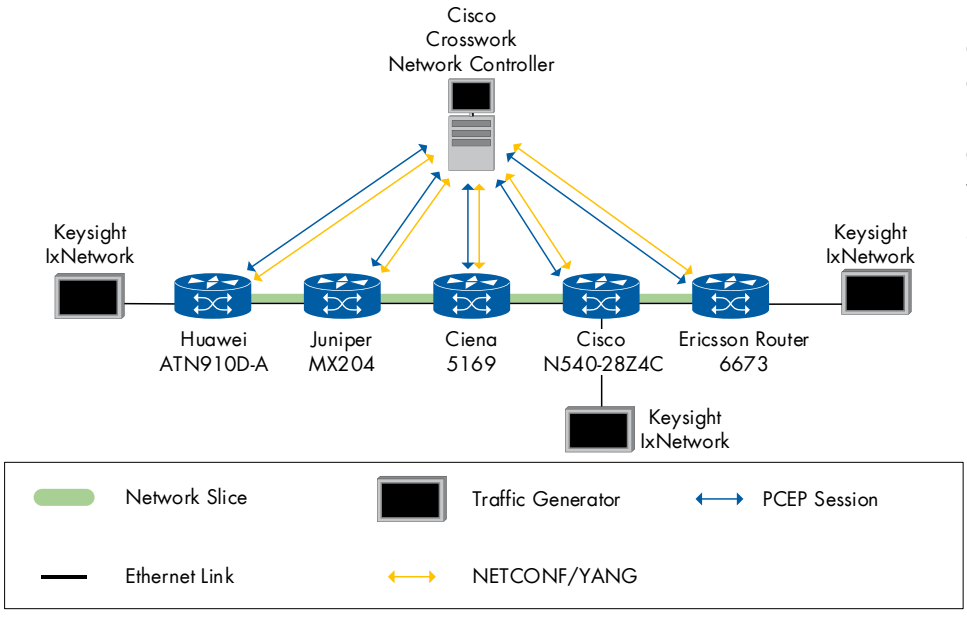


Figure 78: NETCONF Transport Slicing

The slice's configuration and functionality were validated by generating and monitoring bidirectional traffic, as depicted in Figure 78. However, we encountered an issue where the SR policy could not be signaled to one router via PCEP, though its NETCONF-based slice configuration functioned correctly.

L3VPN/L2VPN Provisioning

Integrating Layer 2 (L2) VPNs into modern networks, alongside Layer 3 (L3) VPNs, offers a comprehensive connectivity solution that enhances network flexibility, security, and control. L2 VPNs extend broadcast domains across dispersed sites, complementing L3 VPNs' ability to segment network traffic efficiently.

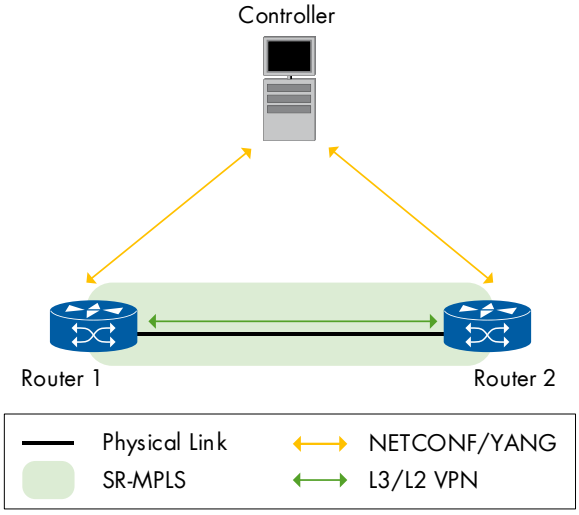


Figure 79: L2/L3 Service Provisioning

NETCONF standardization is crucial in this integrated VPN environment, streamlining the configuration and management of diverse network devices. It enables automated L2 and L3 VPN provisioning, reducingcomplexity and enhancing operational reliability.

We established NETCONF sessions between the network controllers and routers during the VPN provisioning process. For the L3VPN, we deployed VRF configurations and confirmed their connectivity via ping tests. We applied EVPN VPWS configurations for the L2VPN and verified its connectivity using ping tests, ensuring both VPN types were provisioned correctly and functional.

The following devices successfully completed the L3VPN provisioning interoperability test:

Controller	Router 1	Router 2
Cisco Crosswork Network Controller	Juniper MX204	Cisco N540X-12Z16G
	Ericsson R6673	
	Huawei ATN910D-A	
	Ciena 5169	Cisco N540-28Z4C
Huawei iMaster NCE-IP	Juniper MX204	Huawei ATN910D-A
	Cisco N540-28Z4C	
Keysight IxNetwork	H3C S12500R-48C6D	Juniper MX204

Table 17: L3 Service Provisioning Results

The following devices successfully completed the L2VPN provisioning interoperability test:

Controller	Router 1	Router 2
Cisco Crosswork Network Controller	Juniper MX204	Cisco N540-28Z4C
	Huawei ATN910D-A	
Huawei iMaster NCE-IP	Cisco N540-28Z4C	Huawei ATN910D-A
	Juniper MX204	

Table 18: L1 Service Provisioning Results

In the test scenarios involving the Cisco Crosswork Network Controller, we utilized IETF L2NM RFC 9291 Service YANG for L2 services and IETF L3NM RFC 9182 Service YANG for L3 services.

Additionally, in the particular test setup involving the Cisco Crosswork Network Controller, Ericsson Router 6673, and Cisco N540X-12Z16G, the provisioned L3VPN was configured to support dual-stack, enabling compatibility with both IPv4 and IPv6 protocols.

Multipoint L2 VPN Provisioning

This year, our L2VPN testing expanded to include Multipoint L2VPN provisioning, where the controller established a unified L2VPN across three devices, enhancing direct connectivity and network integration. The Cisco Crosswork network controller leveraged the IETF L2NM—RFC 9291 YANG model, which was then translated into specific device configurations implemented via the NETCONF protocol.

Meanwhile, the Huawei iMaster NCE-IP utilized a different configuration and service provisioning approach, which also achieved the desired outcomes.

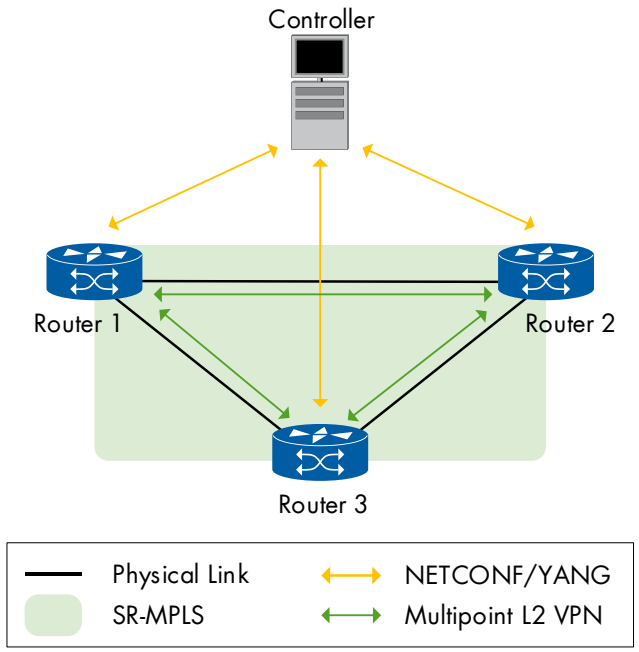


Figure 80: Multipoint L2 Service Provisioning (Router 1 was Cisco N540-28Z4C in all cases)

Controller	Router 2	Router 3
Cisco Crosswork Network Controller	Juniper MX204	Juniper MX204
Huawei iMaster NCE-IP	Cisco N540-28Z4C	Huawei ATN910D-A

Table 19: Multipoint L2 Service Provisioning Results

Telemetry—gNMI

gNMI (gRPC-based Network Management Interface) telemetry, as defined in the OpenConfig framework, has advanced to enable real-time data streaming and model-driven insights, enhancing network visibility and interoperability. For our test procedure, we established gNMI subscriptions to routers using the OpenConfig-interface YANG telemetry model, following the NETCONF session status confirmation.

We retrieved PCEP session details, BGP neighbor statuses, and CPU and memory utilization metrics in one test combination. We observed inconsistencies in the telemetry stream's message structures associated with additional prefixes in the telemetry stream. This issue was solved during testing by applying a new patch to the controller.

In the second test combination, we generated a service assurance graph based on RFC 9417, "Service Assurance for Intent-Based Networking Architecture," which depicted the status of an L3VPN and obtained the BGP neighbors' status. However, for retrieving PCEP peer status, some routers lacked the gNMI path, driving the controller to resort to CLI scripts to acquire those parameters.

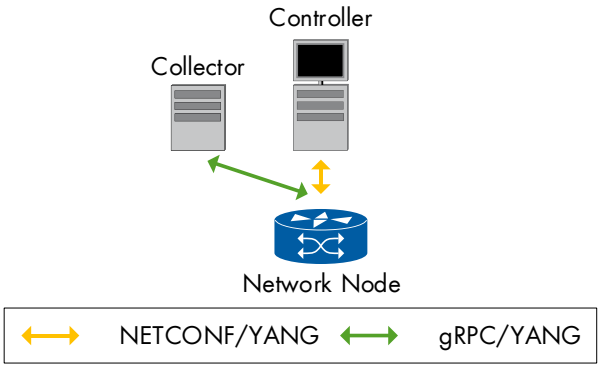


Figure 81: gNMI Telemetry

Controller / Collector	Network Node
Cisco Crosswork Network Controller	Juniper MX204
Juniper Paragon Applications	Cisco N540-28Z4C

Table 20: gNMI Telemetry Results

System Inventory

In this test, an SDN controller used NETCONF to retrieve system inventory details from network devices, leveraging open interfaces like openconfig-inventory, openconfig-platform, and ietf-alarms. These interfaces facilitate the collection of hardware and software information—such as serial numbers, model numbers, and firmware versions—as well as dynamic operational details like NTP service status, alarms, fan status, and operational temperatures.

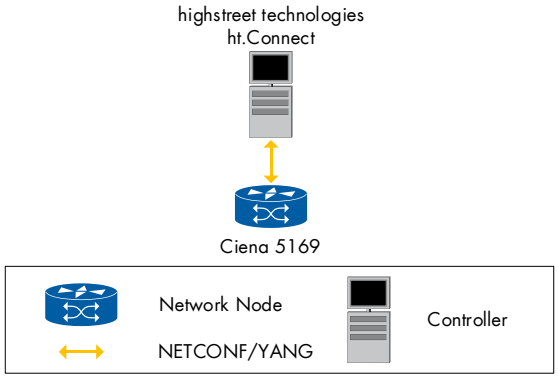


Figure 82: System Inventory Test Setup

Time Synchronization Test Results

Time Synchronization – or more specifically, frequency, phase, and time synchronization – is an essential facet of modern networking, integral to the uninterrupted functionality of networks spanning diverse sectors such as enterprise environments, data centers, and service provider networks.

The intricacy and significance of deploying a robust Time Synchronization network necessitates detailed planning and precise execution.

Working with participating vendors, the EANTC team designed and carried out comprehensive tests specifically aimed at addressing the evolving needs of the industry.

Our 2024 event focused on addressing the challenges and leveraging the opportunities introduced by recent advancements in network technologies, notably the growing significance of 5G networks and the adoption of Open RAN architectures.

Our testing standards have been rigorously updated to align with the demanding synchronization needs of 5G networks, ensuring comprehensive coverage of their sophisticated requirements like time error requirements, which are defined in the ITU-T G.8271.1.

For this year, we abandoned some of the old tests we have repeatedly performed during the last couple of years (i.e., PTP profiles support and simple failover scenarios) and replaced them with more advanced tests:

- Interworking Gateway Profile: This function becomes vital when the existing network must support both Partial and Full Timing Profiles simultaneously, requiring boundary clocks to translate between profiles while maintaining peak performance.
- PTP over DWDM: This test used the Ciena Coherent ELS System as a transport layer for PTP packets, using 400ZR+ pluggable.
- Multi-Boundary Clock Holdover: designed to verify the holdover performance of each Boundary Clock in one topology.
- O-RAN scenarios: Passive port monitoring tests.

Some planned test cases could not be executed due to functional limitations of the participating devices: none of the available devices allowed vPRTC testing due to the lack of support for the same profile in two domains. Different participating vendors' MACsec implementations were not interoperable: MACsec was implemented in a proprietary way by some vendors. Additionally, there are unresolved systematic PTP challenges when sending PTP frames encrypted using MACsec: For tight network synchronization, PTP requires accurate timestamping of the packet. However, MACsec requires insertion and removal of the 24-to-32-byte long MACsec header on all or some of the frames on the link, causing large delay variations between the egress timestamping point and the link connector (and similarly on the ingress). The PTP protocol assumes that the delay on a link is constant. With MACsec, however, this is not the case. Anyway, PTP over MACsec was tested with two devices from the same vendor in previous years.

Boundary Clock Class D Conformance

Class C/D Boundary Clocks are engineered to meet the stringent requirements for Time Synchronization in modern 5G networks. Specifically, applications indicated by the recommendation ITU-T G.8271/Y.1366 require more precise clocks, such as LTE intra-band contiguous carrier aggregation and NR MIMO or TX diversity.

Despite the increasing general presence of Class D Boundary Clocks in the industry, we haven't observed new requirements being defined for Class D specifications recently. This year, we witnessed a significant enhancement in the time error results of participating Class D Boundary Clocks, as in the past years, it was difficult to pass the ±5ns absolute time error low pass filtered for a single boundary class node. This year, most of the devices that participated passed this criterion; Additionally, these Class D devices also easily met the network limits when used in a chain of devices.

This indicates that, even without new standards or benchmarks being set, the technology and implementation of Class D Boundary Clocks are evolving, improving their functionality.

It is essential to clarify that this first test case evaluates the time error performance of an individual device rather than its interoperability capabilities. This evaluation was a prerequisite to including Class D clocks in chain testing procedures. We utilized the Calnex Paragon-neo and Keysight Time Sync Analyzer to simulate the Grandmaster and Slave Clocks, positioning the device under test as an intermediary Boundary Clock. In compliance with ITU G.8273.2, we evaluated the device's performance by measuring the low-pass filtered two-way time error $\max|TE_L|$, applying a threshold of 5 nanoseconds (ns).

The Boundary Clocks were configured to enable

Precision Time Protocol (PTP) and Synchronous Ethernet (SyncE) towards the Slave Clock, employing the PTP G.8275.1 hybrid profile.

We configured the enhanced TLV in ESMC from G.8264 in all these test runs, but in some cases, the boundary clock could understand Enhanced SyncE but not propagate it downstream. Also, we observed in one case that the boundary clock didn't use the physical SyncE signal from the master node and instead recovered the frequency sync from the received PTP. The following snapshot configurations snippet illustrates the status of one device that participated in the test, detailing the parameters we used.

```
Clk qualified Priority: 1
Configured QL: PRC
ESMC QL : ePRC
Clock source type: ifd
Clock Event: Clock locked
Wait-to-restore: 0
```

Calnex Paragon-neo and Keysight Time Sync Analyzer devices were used to verify the results of this test.

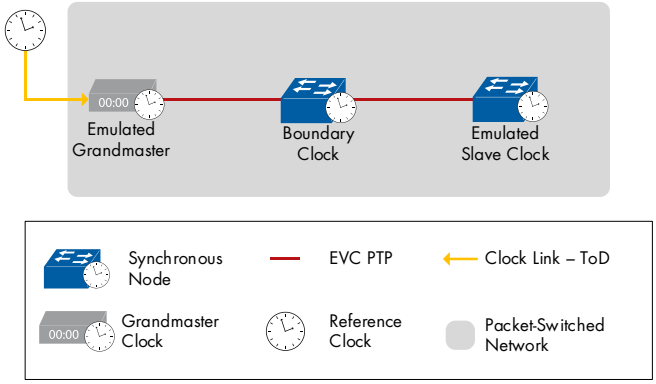


Figure 83: Conformance Test Boundary Clock Class D

The following devices successfully executed the Class D conformance test with different port speeds as shown:

- Arista 7280R3 (100GbE)
- Ciena 5169 (10/25/100GbE)
- Cisco N540-24Q8L2DD (10/100GbE)
- Cisco N540X-16Z4G8Q2C (10/100GbE)
- Ericsson Router 6676 (10/100GbE)
- Ericsson Router 6678 (100GbE)
- H3C S12500R-2L (10GbE)
- Huawei ATN910D-A (10/100GbE)
- Juniper ACX7100-48L, ACX7024 (10GbE)
- Juniper ACX7332, ACX7509 (10GbE)
- Juniper MX304 (10GbE)
- Juniper PTX10001-36MR (100GbE)
- Juniper PTX10002-36QDD (100GbE)
- Microchip TimeProvider 4100 (1GbE)
- ZTE ZXR10 M6000-8SE (10GbE)

Boundary Clock Class C Conformance

We conducted a conformance assessment for Class C Boundary Clocks in line with recommendation ITU-T G.8273.2 Clause 7.1.4.

The clause includes two different requirements:

- Relative constant time error, which should be within ±12ns range.
- Relative dynamic time error low-pass filtered noise generation (MTIE) should be in the range of 14ns.This involved measuring the constant time error across two ports of a Boundary Clock.

Accordingly, some devices were tested at various Ethernet port speeds to ensure a comprehensive evaluation. Both Calnex Paragon-neo and Keysight Tine Sync Analyzer devices were used to verify the results of this test.

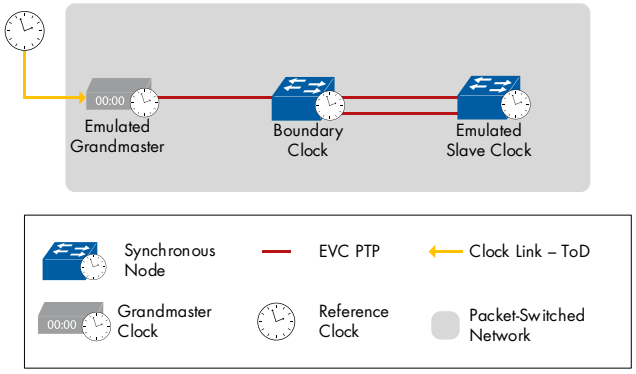


Figure 84: Conformance Test Boundary Clock Class C

The following devices successfully executed the Class C conformance test with different port speeds as shown:

- Arista 7280R3 (100GbE)
- Ciena 5169 (25/100GbE)
- Cisco N540-24Q8L2DD (100GbE)
- Ericsson Router 6676 (100GbE)
- Ericsson Router 6678 (100GbE)
- H3C S12500R-2L (10/100GbE)
- Huawei ATN910C-G (10GbE)
- ZTE ZXR10 M6000-4SE (10GbE)

Time Synchronization Source Failover

This test was a part of resiliency tests of Time Synchronization while having two boundary clocks in a chain, which makes the topology a more real-world scenario. The test topology consisted of the following elements:

- **Grandmaster A (GM-A)** connected to the Global Navigation Satellite System (GNSS) as a reference, used as the main reference for the topology.
- **Grandmaster B (GM-B)** connected to GNSS and used as a backup.
- **Boundary Clock-1 (BC-1),** connected to both GMs and configured to prefer the GM-A when locked to the GNSS, using the links' local priorities.
- **Boundary Clock-2 (BC-2)**, connected to BC-1 and the Time Error Analyzer device, providing both PTP and SyncE output to the measurement device.

The test started when GM-A and GM-B were locked on the GNSS reference. The Boundary Clock-1 was locked with both PTP and SyncE from the GM-A. The first phase of measurement was started for 1000 seconds to be able to calculate the Constant Time Error cTE; then, the GNSS was disconnected from the GM-A, causing the BC-1 to switch reference to GM-B as the source.

GNSS was then also disconnected from GM-B, so the configured local priorities caused the reference of the BC-1 to switch back to GM-A in holdover as their reference.

GM-B was then reconnected to GNSS, followed by reconnection of the GM-A GNSS, while measurement of 1PPS and PTP 2way TE from BC-2 continued. The limits were set to G.8271.1 level 6A (260 ns).

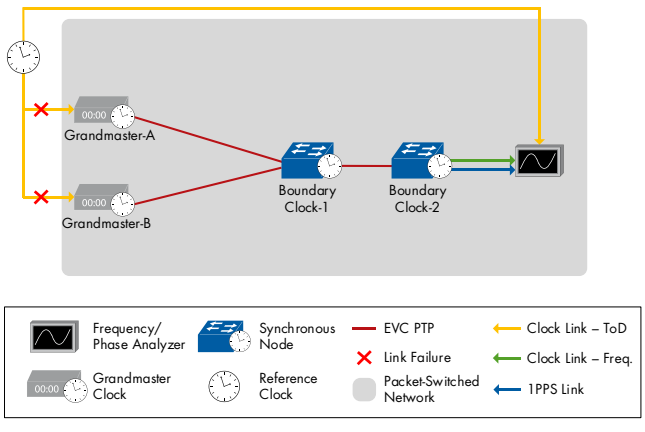


Figure 85: Time Synchronization Source Failover

One test run used a 400GbE optical link between the two BCs, with Juniper ACX7332 as BC-1 and BC-2 being Ericsson R6678.

Calnex Sentry was used to verify the output of the Boundary Clock during this test.

GM A	GM B	BC-1	BC-2
Huawei ATN910D-A	Ericsson 6676		Huawei ATN910C-G
Microchip TimeProvider 4100	Cisco N540X-16Z4G8Q2C	Ciena 5169	Cisco N540-24Q8L2DD
	Microchip TimeProvider 4100	Juniper ACX7332	Ericsson 6678
Cisco N540X-16Z4G8Q2C	Ciena 5169	Arista 7280R3	

Table 21: Time Synchronization Source Failover Results

Holdover with Enhanced Sync-E Support

Enhanced Synchronous Ethernet (eSyncE) provides physical layer frequency support to PTP-aware devices in full-timing support networks, improving performance to enable the stringent synchronization requirements of modern telecommunication networks.

This test verified the ability of a chain of boundary clocks, configured to use eSyncE, to maintain acceptable values of time error during the loss of the Grandmaster Global Navigation Satellite System (GNSS) reference and later emulated the case of PTP loss from GM and through the chain. eSyncE ESMC messages were also captured and analyzed to verify that the eSyncE TLV was being processed as required by each device in the chain.

The testing process began by designating GM-A as the primary Grandmaster (GM), which served as the main source for both Precision Time Protocol (PTP) and Enhanced Synchronous Ethernet (eSyncE) throughout the chain. We conducted a prolonged measurement of the time error over 1000 seconds, enabling us to determine the constant time error cTE under stable conditions accurately.

Subsequently, we proceeded to the next phase by isolating GM-A from the GNSS, triggering a failover mechanism within the network. This action caused the first boundary clock in the sequence, followed by the entire chain, to switch to GM-B as the new time source for PTP and eSyncE functionalities.

To simulate a realistic operational scenario, we initiated a controlled disruption of the PTP connection between GM-B and the chain's first boundary clock (BC). This allowed us to evaluate the performance of the enhanced SyncE during a holdover period, assessing the network's resilience in the absence of PTP synchronization.

To further increase the test's realism, we systematically turned off the PTP links between successive boundary clocks along the chain at intervals of 350 seconds. This systematic approach enabled us to observe the impact of each disconnection on the time error at the end of the chain, providing valuable insights into the network's Time Synchronization under progressively degrading conditions.

During one specific test scenario, an anomaly was observed where a device continued to broadcast a ClockClass of 248 despite being in a holdover state. This irregularity adversely influenced the downstream boundary clock in the chain, prompting it to enter an out-of-specification holdover state. However, upon disabling the PTP downlink from the affected node—consistent with our 350-second phased approach—the subsequent boundary clocks reverted to normal holdover behavior, correctly transmitting ClockClass values of 135 or 165.

This comprehensive testing procedure highlighted the resilience and failover capabilities of the network's synchronization architecture. It underscored the critical importance of precise control and monitoring of Time Synchronization mechanisms, especially in complex, layered network infrastructures.

EANTC verified the results of this test using the Calnex and Keysight measurement devices. For all test runs, the Maximum relative time was below 260ns which complies with the time error limit recommendation of ITU.T T G.8271/Y.1366 (03/2020) with a Maximum relative time level of 6A, 260ns.

The topologies in Figures 86 and 87 show the setups used in this test.

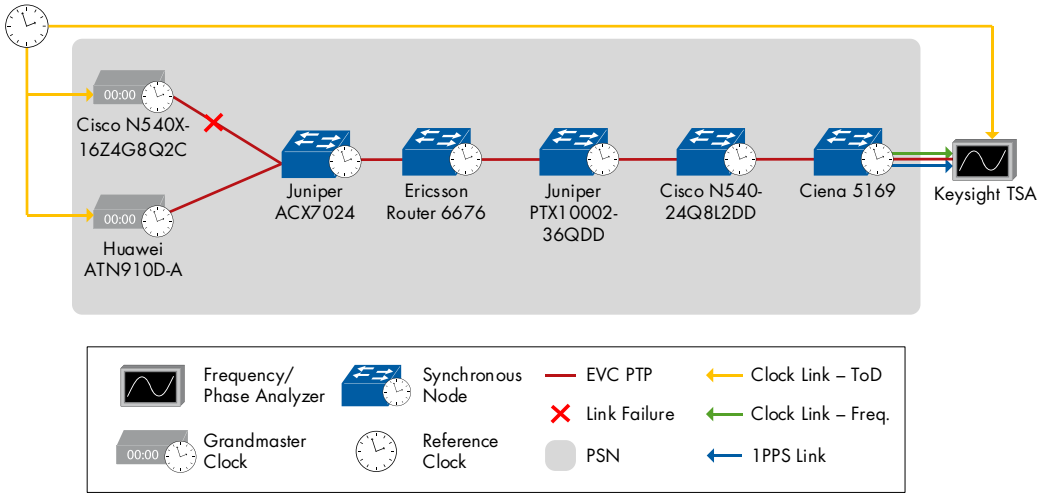


Figure 86: Holdover with Enhanced SyncE Support, Run 1

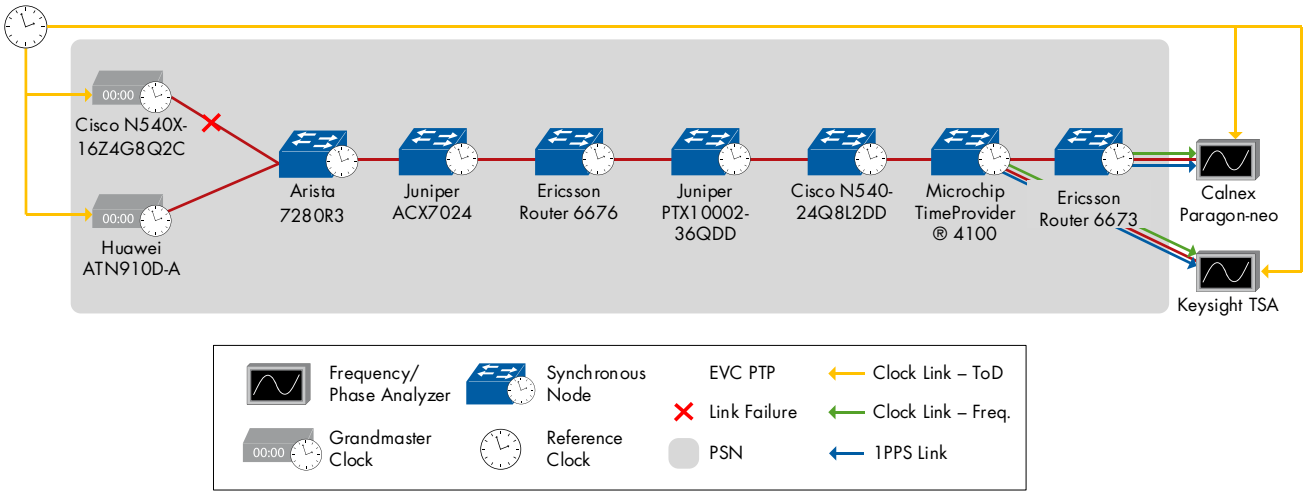


Figure 87: Holdover with Enhanced SyncE Support, Run 2

Calculating Time Error Limits for Boundary Clocks

Deployed synchronization networks are commonly chains of boundary clocks, T-BCs, with well-defined devices and network time error performance specifications. This test case validated the performance of a chain based on the number and class of T-BCs in it.

Two different scenarios were evaluated:

- Chain of Class C Boundary Clocks: G.8273.2 defined Performance estimation for chain of Boundary Clocks class A/B/C and specifies details for calculating limits for chains of Boundary Clocks. For cTE, since the accumulation is additive, the value for the chain is the cTE for a single device multiplied by N. For five class C T-BCs, this is 5 · 10 = 50 ns. For quantities such as dTEL MTIE and dTEL TDEV, which have a square root of the sum of squares accumulation, if the respective value for a single T-BC is x, the calculation is $\sqrt{N} \cdot x^2$.
- Chain of Class D Boundary Clocks: Class D T-BCs have only a single metric, max|TEL|, defined, for which the accumulation is not specified. max|TEL| may be considered as all cTE, all dTE, or a combination. Considering all as cTE gives a best-case limit of N· max|TEL|, i.e., N·5, used for these test evaluations. It was also observed that if max|TEL| was considered as only dTE, i.e., the limit could be calculated as $\sqrt{N} \cdot x^2$, all test cases also passed this tighter limit.

We also had to consider the PRTC budget, which is ±100ns for the Grandmaster we used for the GNSS time error test, except when we used the Calnex NEO or the Keysight Time Synchronization Analyzer as GMs which supported a PRTC budget was set to zero.

Figure 88 shows all the successful test runs and devices, which all remained within the Class D limits.

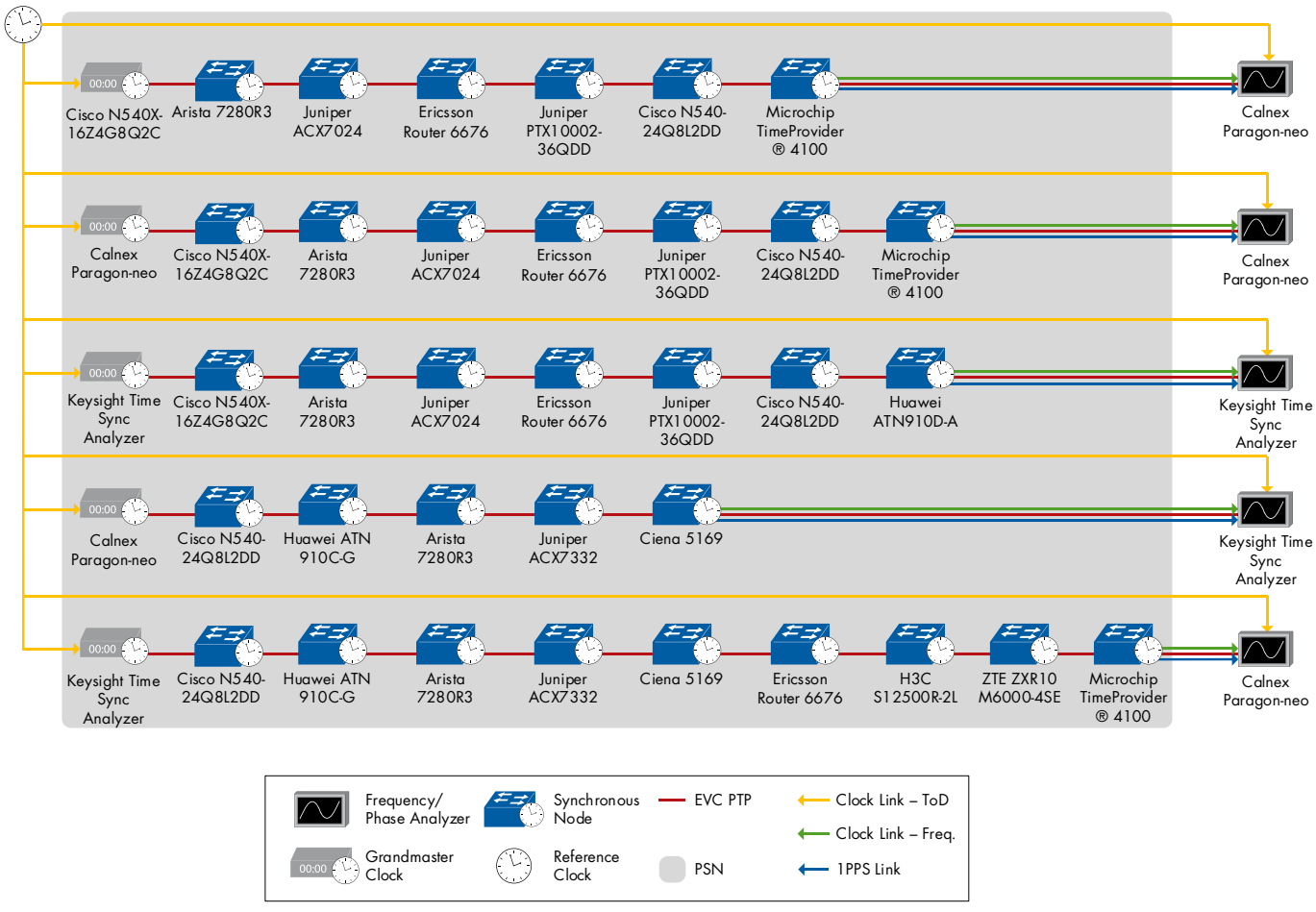


Figure 88: Calculating Time Error Limits for Boundary Clocks

Delay Asymmetry Detection/ Measurement

Controlling asymmetric delay on links carrying PTP messages is critical and one of the most challenging issues for network Time Synchronization.

At this year's event, we tested one scenario to showcase the participating devices' abilities to detect the applied asymmetry, as we had already performed this test multiple times in previous years.

In this scenario, with one Grandmaster and two Boundary Clocks, the Grandmaster (Keysight Time Sync Analyzer) and Boundary Clock-1 (Ericsson Router 6676) were referenced to GNSS (via a splitter) through an antenna on the roof of EANTC's lab. PTP profile G.8275.2 was used across the whole chain.

When the Grandmaster and Boundary clock-1 were locked to GNSS. Boundary Clock-2 (Huawei ATN910D-A) used Boundary Clock-1 as the timing source. When the GNSS connection to Boundary Clock -1 was disconnected, Boundary Clock 1 reverted to using PTP from the Grandmaster as its timing source, with the Boundary Clock-2 1PPS absolute time error being measured across this transition. We restarted the measurements, introduced an asymmetric delay of 300 ns using the Keysight Time Sync Analyzer, and waited for the Boundary Clock-1 to detect the asymmetry.

The Boundary Clock-1 detected the 300ns delay and, at the same time, didn't allow the asymmetry to be transported to the second Boundary Clock, which helped keep the Time Error within the respected limit. We measured the Maximum Absolute Time Error, and the following diagram's devices complied with G.8271 accuracy level 4.

The following diagram contains the test topology with the devices that participated in this test.

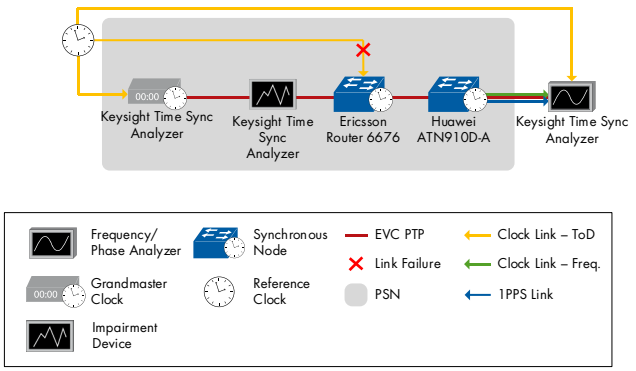


Figure 89: Delay Asymmetry Detection/Measurement

Boundary Clock Interworking Function Performance

Brownfield networks may present challenging configuration scenarios, where using only the Full Timing Support through the whole network is not an option. On the other hand, using the Partial Timing Support alone is also not suitable for all synchronization scenarios. In such cases, both PTP profiles should be configured. This situation requires the Boundary Clocks to be able to translate between the two profiles, complying with the specifications for the respected profile.

In this test, we used a grandmaster, a boundary clock that implemented the interworking function gateway role, and a slave clock. In addition to analyzing the time error from the output of the Slave Clock, we connected the Boundary Clock to the analyzer to capture the PTP packets and check the PTP flags.

We performed this test by configuring the Boundary Clock to translate the Full Timing Support Profile G.8275.1 to the Partial Timing Support G.8275.2 and vice versa, checking the packets of the PTP for the flags, ClockClass, and Clock Accuracy advertised by the boundary clock while measuring the time error at output of the slave clock.

The following diagram shows the topologies used during this test.

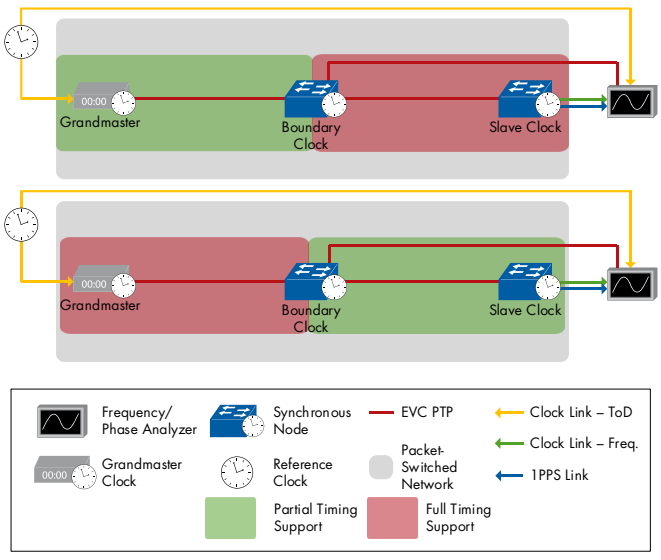


Figure 90: Boundary Clock Interworking Function Performance

Both Calnex Paragon-neo and Keysight Tine Sync Analyzer devices were used to verify the results of this test.

The following table shows the successfully tested device

combinations for interworking between G.8275.2 and G.8275.1:

Grandmaster	Boundary	Slave Clock
Interworking G.8275.2-to-G.8275.1		
Keysight Time Sync Analyzer	Ericsson 6676	Ciena 5169
	Ciena 5169	Ericsson6676
	Huawei ATN910D-A	Arista 7280R3
Interworking G.8275.1-to-G.8275.2		
Calnex Paragon-neo	Huawei ATN910D-A	Arista 7280R3
Keysight Time Sync Analyzer	Ericsson R6676	Ciena 5169

Table 22: Test Results of Boundary Clock Interworking Function Performance

Passive Port Monitoring

Annex G of ITU-T G.8275.1 defines the optional feature of Passive Port Monitoring (PPM), which has proven to be a valuable feature in the operation of synchronization networks.

The PPM feature can monitor the PTP (Precision Time Protocol) phase/time difference between the passive port and the slave port for T-BC/T-TSC nodes.

PPM can be utilized in the following use cases:

- Provisioning: Measure and compensate for asymmetry in network nodes.
- Monitoring: Use the PPM feature to monitor and compare the PTP phase/time difference between different clock sources throughout the network.
- Analysis: Continuously measure the PTP phase/time difference between the ports on a given T-BC/T-TSC node from different upstream time sources. This helps identify devices that impact clock quality.

PPM had been tested at EANTC interop events in previous years, but that testing occurred before its standardization by the ITU-T. Therefore, this year marks the first time at EANTC that we are testing PPM in accordance with Annex G of ITU-T G.8275.1.

In this test, we verified two scenarios involving PPM, where an Impairment Device introduced asymmetry, enabling us to measure an offset with PPM as follows:

1. BC1 with PPM configured, measuring the offset from PPM towards the slave clock locked to the GM, as the following diagram shows.

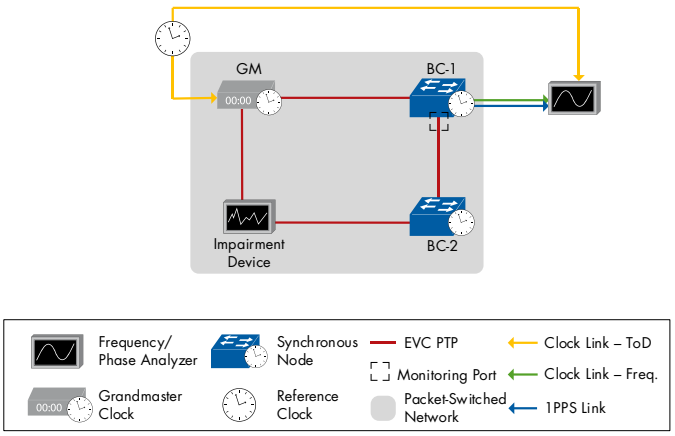


Figure 91: Passive Port Monitoring Setup 1

The following devices participated successfully in this setup (in all cases, the impairment function was provided by Calnex SNE Ignite):

GM	BC1	BC-2
Calnex Paragon-neo	Ericsson Router 6676	Huawei ATN910C-G
	Huawei ATN910D-A	Ericsson Router 6676
	Juniper PTX10002-36QDD	Ericsson Router 6676
Cisco N540X-16Z4G8Q2C	Cisco N540 -24Q8L2DD	Juniper PTX10002-36QDD

Table 23: Passive Port Monitoring Results, Setup 1

The following diagram shows that BC1 with PPM is configured and connected to GNSS to compare the results.

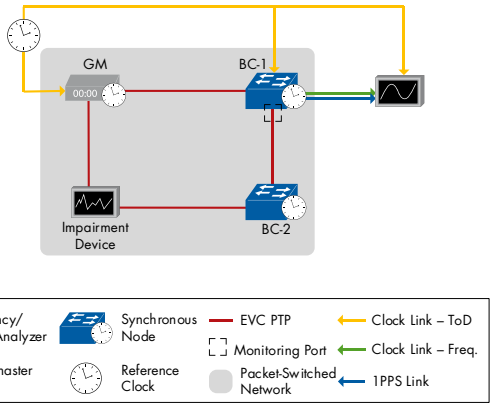


Figure 92: Passive Port Monitoring Setup 2

The first two combinations of table 23 participated in the "setup 2" test configuration successfully as well.

Open RAN Implementations

The Open Radio Access Network (Open RAN) represents a pivotal advancement in the telecommunications industry, attracting widespread interest for its potential to revolutionize network operations.

As professionals across the networking world delve into exploring, testing, and implementing O-RAN solutions, the significance of its components, particularly the fronthaul network, cannot be overstated.

The fronthaul network is essential for the efficient and effective operation of the O-RAN architecture, necessitating precise and reliable Time Synchronization to ensure system integrity, functionality, and performance.

This interest has led us to systematically investigate various Time Synchronization scenarios within the fronthaul network to assess and rigorously ensure its performance and reliability.

O-RAN Fronthaul LLS-C2 (Option-A)

In our tests, we emulated the O-RAN Fronthaul LLS-C2 (Option-A) configuration, with a notable modification: the substituting the Distributed Unit (O-DU) with a Boundary Clock.

Our setup involved a Grandmaster, a Boundary Clock, and two distinct timing paths originating from the Boundary Clock, each incorporating one Hub-Site Router (HSR) and one Cell-Site Router (CSR).

Both CSRs were connected to a time error analyzer to measure the relative Time Error and the 1 Pulse Per Second (1PPS) absolute time error. The test run successfully met all measurement criteria outlined by the O-RAN Alliance in the document O-RAN.WG9.XTRP-TST.0-R003-v03.00 for FR2, demonstrating compliance with established standards.

It is crucial to highlight that while integrating an O-DU typically enlarges the time error budget, the exemplary results obtained in our tests suggest that including an O-DU is unlikely to affect the overall performance outcomes adversely. This conclusion is supported by the accuracy and reliability of the Time Synchronization achieved in our scenario without the O-DU.

Figure 93 shows the topology used for this test, and Table 24 lists the test combinations successfully evaluated in this setup.

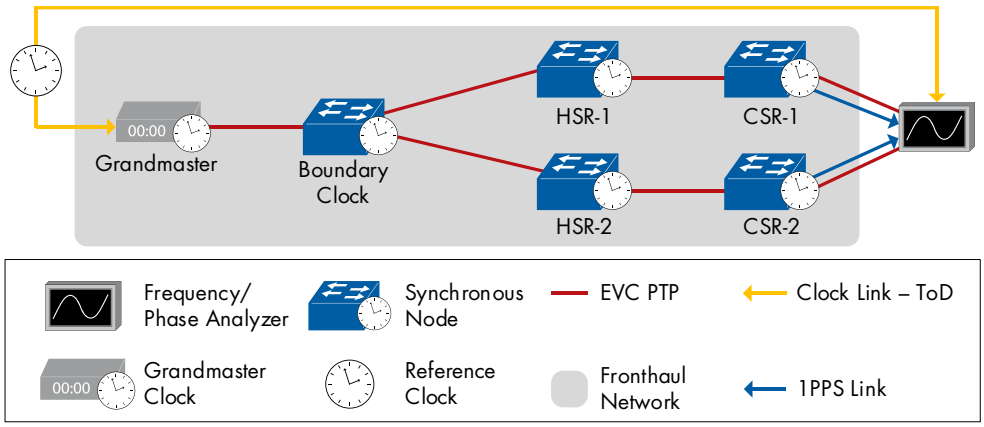


Figure 93: O-RAN Fronthaul LLS-C2 (Option A)

GM	BC	HSR1	CSR1	HSR2	CSR2	Measure-
Keysight Time Sync Analyzer	Ciena 5169	Juniper ACX7332	Arista 7280R3	Ericsson Router 6676	Huawei ATN910D-A	Keysight Time Sync Analyzer
Microchip TimeProvider 4100	Ciena 5169	Arista 7280R3	Juniper ACX7332		Huawei ATN910D-A	
Cisco N540X-16Z4G8Q2C	Huawei ATN910D-A	Arista 7280R3	Juniper ACX7332		Ciena 5169	Calnex Paragon-neo

Table 24: O-RAN Fronthaul LLS-C2 (Option A) Test Results

O-RAN Fronthaul LLS-C3 Configuration with GM from Midhaul

Furthermore, we performed a test to emulate an LLS-C3 scenario per the O-RAN.WG9.XTRP-SYN-v03.00 document, where the Grandmaster is positioned at the Midhaul.

Figure 94 shows the test topology used to validate this configuration. In all test runs, Keysight Time Sync Analyzer implemented the Emulated O-CU, O-DU, and O-RU functionality, and served as the traffic generator. Additionally, in all test runs, the Juniper ACX7100-48L functioned as the HSR (Boundary Clock).

The following table lists the results.

Grandmaster	Cell Site Router (Boundary Block)
Keysight Time Sync Analyzer	Juniper ACX7024
	Ciena 5169
	Arista 7280R3
	Cisco N540X-16Z4G8Q2C
	Ericsson Router 6676
	Huawei ATN910D-A
Microchip Time-Provider 4100	Huawei ATN910D-A

Table 25: O-RAN Fronthaul LLS-C3 with GM at Midhaul Results

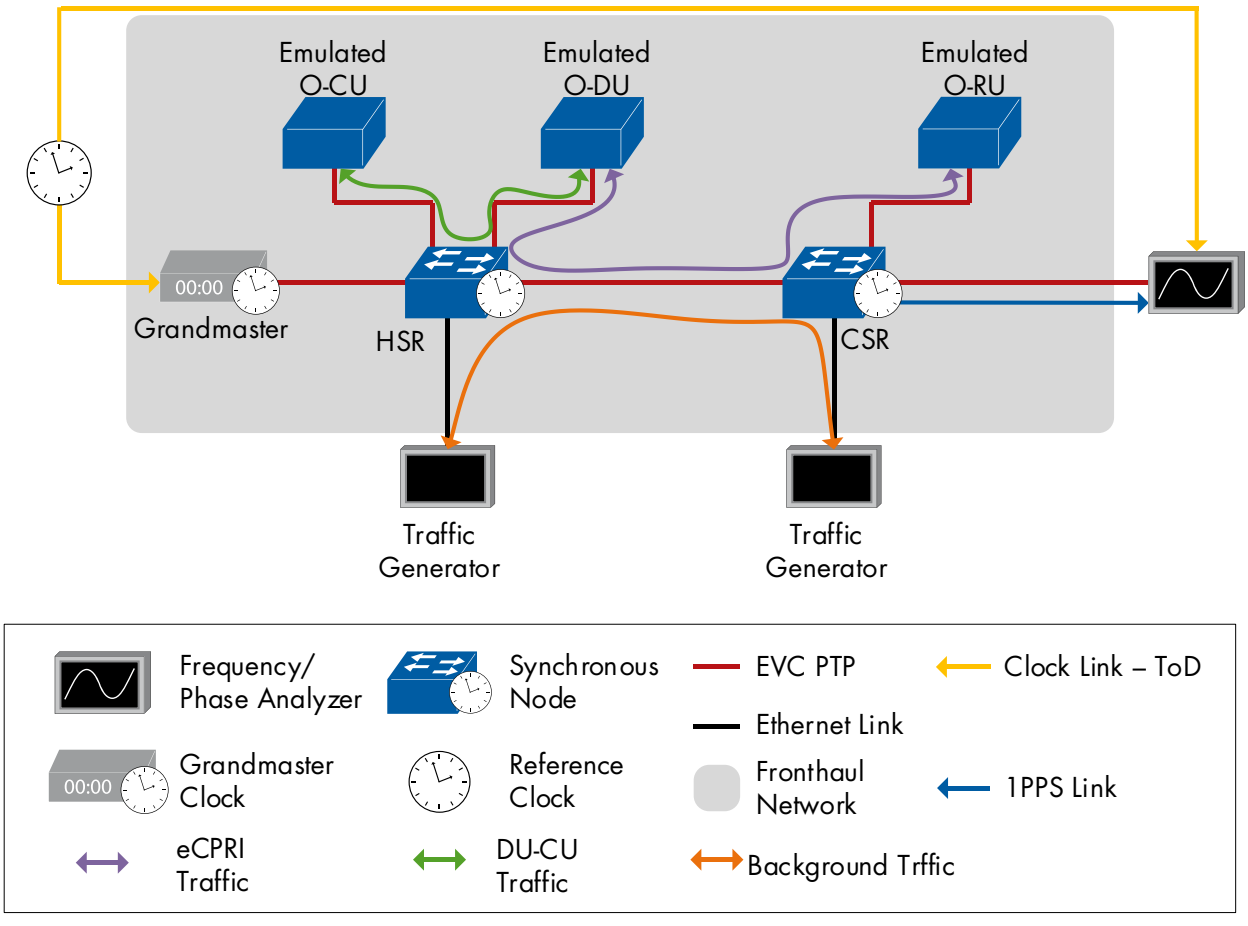


Figure 94: O-RAN Fronthaul LLS-C3 Configuration with Grandmaster at Midhaul

PTP over DWDM Line System

Dense wavelength-division multiplexing (DWDM) is an optical transmission technology that uses multiple wavelengths of light to combine several data streams onto a single optical fiber. DWDM could be found in backbone networks, metro areas, and access networks, providing a scalable way to meet growing bandwidth demands.

We used Ciena Coherent Edge Line System (ELS) for this test to create the DWDM connection between a Boundary Clock and a Slave Clock. The DWDM line rate was 400Gbps and the configured wavelength was 1564.68 nm. Ciena Navigator NCS SDN controller was used to configure the Ciena ELS for the needed power, frequency, and speed.

Test Run-1:

- The Cisco N540-24Q8L2DD (Router B in Figure 95) locked towards Juniper PTX10002-36QDD (Router A in Figure 95) with PTP and SyncE over the 400GE interfaces via the Ciena Coherent ELS.
- Additionally, Cisco N540-24Q8L2DD was connected to Juniper PTX10002-36QDD via a 400GE ZR+ interface with a direct fiber cable, and to Cisco N540X-16Z4G8Q2C via a 10GE interface (dotted red lines in Figure 95)
- We configured Passive Port Monitoring (PPM) to measure the offset between all the different ports. We also recognized the different offset with 400GE ZR+ interfaces via Ciena Coherent OLS and fiber cable.

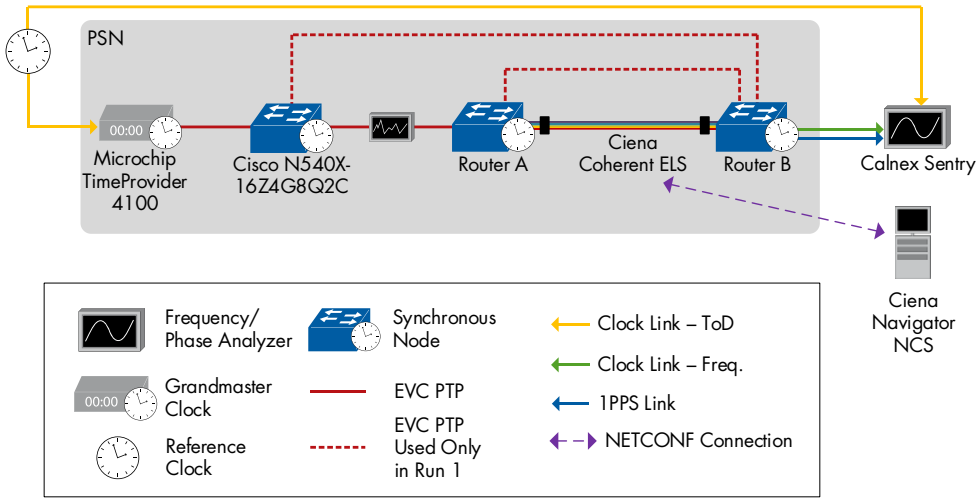


Figure 95: PTP over DWDM, Test Runs 1 and 2

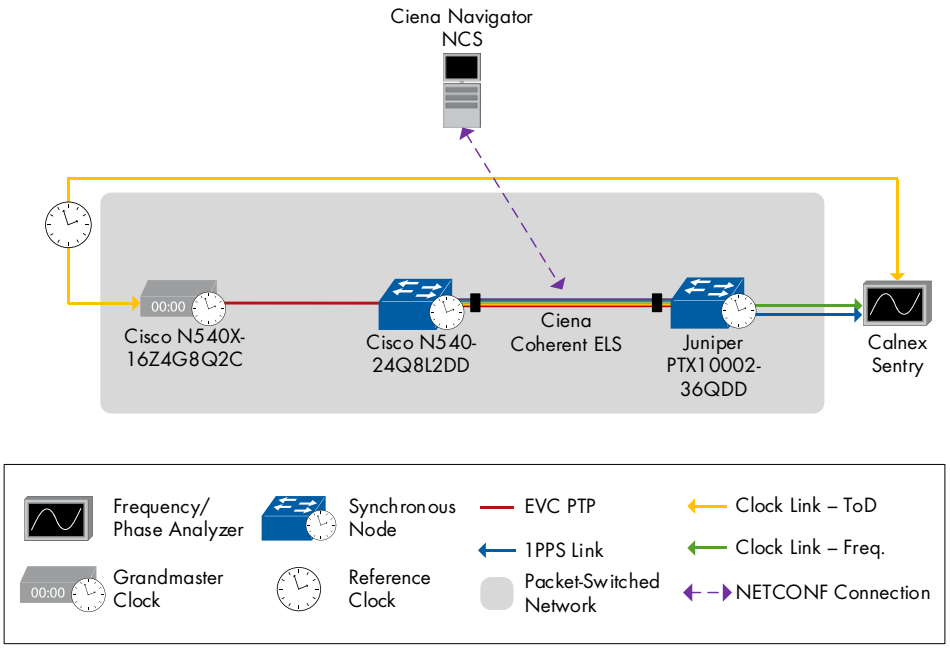


Figure 96: PTP over DWDM, Test Run 3

For test runs 2 and 3:

- The direction of timing was reversed whereby Cisco 540-24Q8L2DD took the role of Router A; it was PTP and SyncE locked to Cisco N540X-16Z4G8Q2C. The Juniper PTX10002-36QDD took the role of Router B in this run; it was PTP & SyncE locked to Cisco 540-24Q8L2DD across the ELS link. The Cisco 540-24Q8L2DD used the second 400GE direct connection to monitor any offset coming back from the PTX10002-36QDD. The red dotted optional links were *not* used in these test runs.

In all cases, we used ±25 ns as the limit of the accepted time error on the 1PPS output of the last clock of the chain.

During this test, two other cases with other vendors failed because the Slave Clock couldn't lock on the Boundary Clock through the DWDM line system. Although the connection was up and running, the PTP failed to lock.

Figures 95 and 96 show the successful test runs for this test.

Overnight Multi-Boundary Clock Holdover Test

Resilience scenarios have always been an essential part of our annual testing event, aiming to emulate failures in the network to identify the best configurations and solutions to mitigate the impact of PTP (Precision Time Protocol) source loss in the network.

This test verifies the ability of boundary clocks to maintain acceptable time error values during the loss of a PTP reference. This results in the boundary clocks entering a holdover state while maintaining frequency lock with the GM or main BC.

As illustrated in the topology, we initiated the test with all boundary clocks connected to a Cisco N540X-16Z4G8Q2C. All clocks had the grandmaster clock, which is the Microchip 4100, as a traceable PTP and SyncE source.

In the next step, we disrupted the PTP connection between the Cisco N540X-16Z4G8Q2C and the downstream boundary clocks (BCs) and measured the time error over 12 hours.

The results of this test were remarkable, as all measured time error values at the output of the downstream BCs were within 7ns of the GNSS reference from the GM throughout the 12-hour measurement period.

Figure 97 shows the topology and the devices that participated in this test..

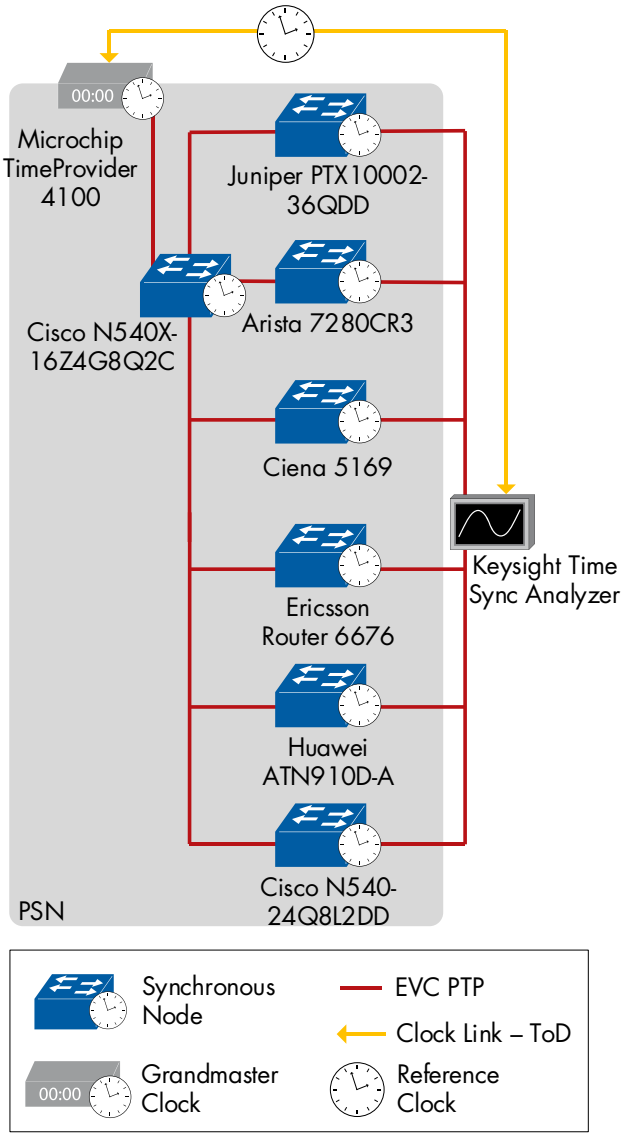


Figure 97: Overnight Multi Boundary Clock Holdover Test

Clocking Show Cases For Paris

At the end of our clocking tests, we prepared two showcase scenarios for demonstrration at the MPLS-SDN-AI World Congress in Paris.

For the first showcase, we packed some of the most important clocking test cases into a single complex test scenario including the following tests:

- A structure of O-RAN LLS-C2 topology with an additional Boundary Clock
- Passive Port Monitoring on two devices
- PTP over DWDM
- Interworking Gateway Function between PTP Profile 8275.1 to 8275.2 and vice versa
- Asymmetry Detection on one of the devices.

Figure 98 depicts the showcase topology.

The second live clocking test demonstration shows a chain of nine (9) Class-D boundary clocks, which is documented previously in this document under the section "Calculating Time Error Limits for Boundary Clocks".

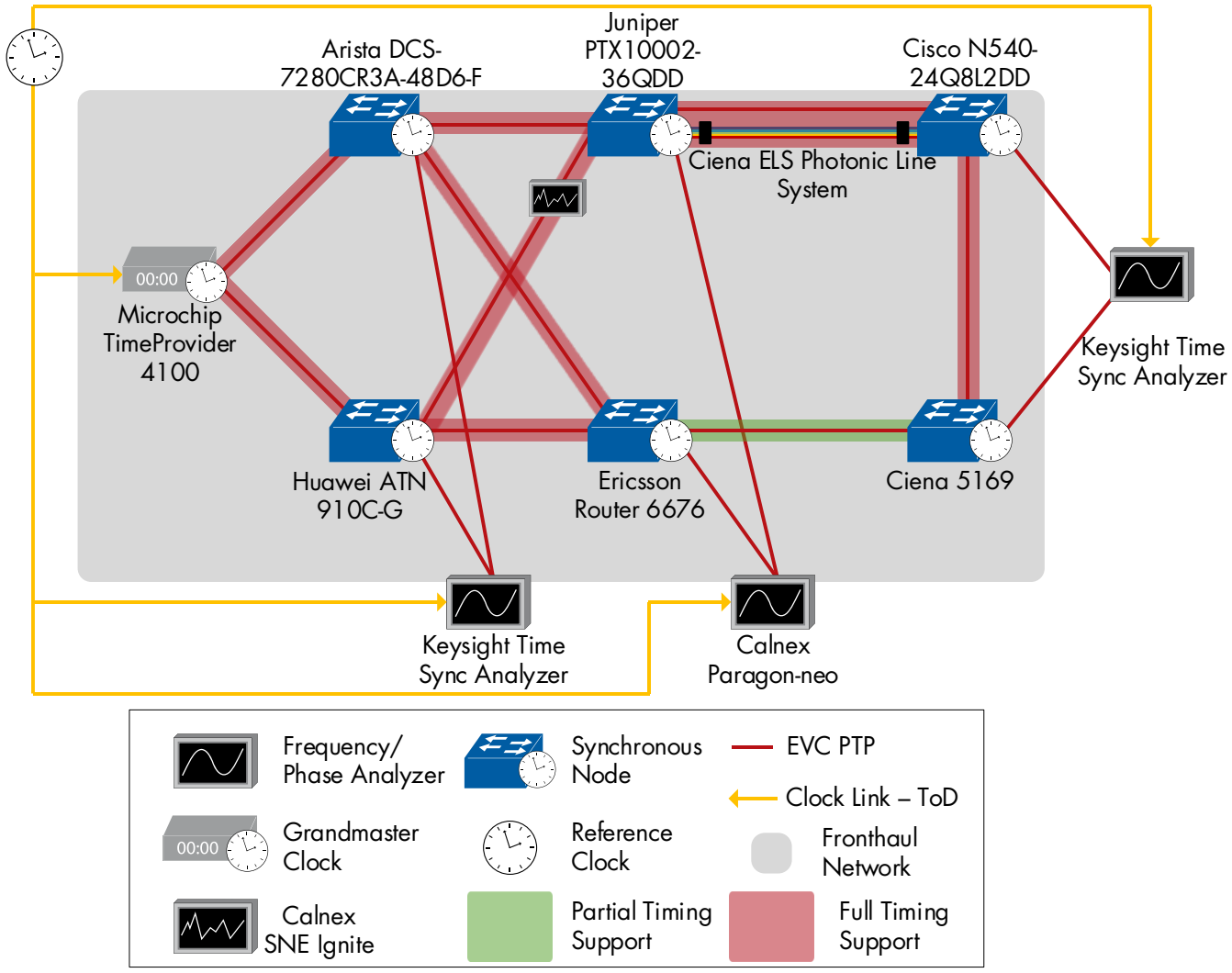


Figure 98: Mixed Clocking Showcase

Conclusion

Reflecting on over 20 years of conducting multi-vendor testing at the Upperside World Congress, the EANTC team considers ourselves fortunate to have played a role in this dynamic industry. The diverse range of solutions from various manufacturers, coupled with the persistent demand from network operators to ensure the modernization of their networks by evaluating and implementing next-generation features and services, has led to a scenario where most service provider networks operate with multiple vendors concurrently within the same network and domain.

This blend of solutions and advanced features drives the industry towards maintaining a standards-oriented ecosystem, fostering continuous innovation, and sustaining a relatively healthy level of competition. This was evident across all technology domains covered during our latest EANTC MPLS SDN interoperability test event again.

Looking ahead, EANTC remains committed to supporting the innovation and production readiness of multi-vendor SDN/Segment Routing transport networks and applications in the future.

About EANTC

EANTC (European Advanced Networking Test Center) is a leading independent test lab for telecommunication technologies. Based in Berlin, Germany, the company offers vendor-neutral, realistic, and high-quality testing and consultancy services for vendors, service providers, and enterprises.

EANTC's performance and scalability, interoperability, proof of concept, acceptance tests, and network audits cover established and next-generation fixed and mobile network technologies.

Our technical expertise focuses on network technologies like 5G, Open RAN, SD-WAN, and security testing.

We organize a unique style of interoperability and performance test events covering advanced technologies such as SDN transport and Open RAN.

https://www.eantc.de

info@eantc.de

EANTC AG, Salzufer 14, 10587 Berlin, Germany
info@eantc.de, www.eantc.de
[v1.1 20240404]