EANTC

White Paper

Multi-Vendor MPLS SDN Interoperability Test

Table of Contents

Editor's Note

For a second time, EANTC and participating vendors are presenting the results of our annual MPLS SDN interoperability test event to a virtual World Congress. For one more year, we cannot present an impressive lineup of the most innovative multi-vendor SDN network in Paris to you.

Together with ten outstandingly committed manufacturers, we ran a hybrid hot staging event for two weeks in July. All hardware was shipped to EANTC, installed in our lab, and supported locally as well as remotely. Like last year, the remote collaboration worked really well. 80 people from 10 companies across three continents and 15 time zones worked together very efficiently and trustfully to build some of the most complex and advanced multi-vendor transport use case scenarios yet.

Two main use cases dominated again: 5G x-haul transport, and datacenter interconnection (used both in enterprises and cloud deployments). We have been covering these use cases for multiple years now, each time extending the test coverage. This year, we noticed more realistic and sophisticated configurations and more complete coverage of implementations, increasing the likelihood of interoperability in complex deployments.

Specifically, as an industry-first, we achieved multi-vendor interoperability for advanced 5G slicing with Flexible Algorithm (Flex Algo) with five vendors. 5G Standalone transport with differentiated slice transport requirements becomes more realistic with this achievement. The participating Flex Algo implementations have developed a lot since our first test, both regarding the underlay and the Segment Routing (SR-MPLS and SRv6; BGP-LS) integration.

We evaluated many possible combinations of affinities and rules. As a result, we would consider the participating Flex Algo implementations ready for deployments.

One of the most sophisticated and time-consuming tests in this event related to the provisioning of SR policies with colored flows—another building block towards slicing in transport networks. Our event has certainly contributed to resolving a number of implementation issues, and we are happy to report successful results with all participants in the end. As part of the industry move to more automated provisioning, we tested on-demand dynamic tunnel creation by SDN controllers by PCE protocol with great results as well. The SR implementations of all vendors participating in this area were very solid. Segment Routing over IPv6 (SRv6) continues to gain acceptance in the industry (we had four participating vendors this time), while is still more challenging in multi-vendor combinations than SR-MPLS—for example, path computation (PCEP) was not yet supported for SRv6.

Participating EVPN implementations showed robust interoperability across all combinations. We tested EVPN over MPLS and VXLAN transport successfully again—this time focusing on multicast over EVPNs and on the seamless integration with legacy VPLS services. Don't take the level of success and maturity for granted across the industry, however; these statements apply only to vendors regularly participating in our interop event series.

Clock synchronization support is another strong pillar of 5G transport requirements. We have continued our series of synchronization tests. For the first time in the industry, interoperability of "Class D" Boundary Clocks has been evaluated successfully, including a chain of such clocks, adhering to the highest precisions defined by the ITU today (260ns Level 6A). To speed up the initial synchronization, Synchronous Ethernet (SyncE) got enhanced recently. We have tested the interoperability of the new standard with very positive results for the first time as well. Finally, we have validated PTP synchronization over FlexE interfaces which may play an important role in 5G fronthaul environments in the future.

Finally, you may notice that this event does not cover NETCONF/YANG orchestration tests. EANTC has split off these tests to a separate series of interop events. The interoperability of device and service management, performance monitoring, and telemetry deservers more time and focus.

The type of management plane testing is quite different as well and regularly got in conflict with the control plane tests in previous events.

Meanwhile, we have published two white papers with NETCONF/YANG interop results for the transport network and will present the results alongside during the SDN World Congresses now and in the future.

One of the advantages of virtual conferences is the availability of presentations at any place and time. EANTC will contribute a wide range of in-depth topic videos with live demonstrations recorded during the hot staging event. We will publish all videos via our YouTube channel. And of course, with this white paper, it is our pleasure to disclose as many details as it takes to understand and reproduce our results independently. Our team, together with the teams of the participating vendors, is ready to answer any further questions. We hope that this white paper will introduce the current state of the art and guide you well through the complex and advanced multi-vendor SDN deployment options today.

Introduction

The MPLS SDN Interoperability Test 2021 aimed to test and improve the interoperability of the multi-vendor implementations through different technologies in the industry and help the new standards be more mature and better to be widely used. This year's tests were dominated by the 5G and the network slicing and their techniques, which covered a wide range of transport networks areas and technologies as follows:

Segment Routing: With the 5G implementations in small contexts being rolled out, the development and standardization of the used technologies and techniques are continued. SR as source routing technology could be useful for simplifying the network and moving the path information from the transit network nodes to the packet itself. It helps to make the network highly responsive and more reliable, and resilient. Having those benefits in hand, we successfully tested multiple interop features and VPN technologies with SR and SRv6, including EVPN and L3VPN. TI-LFA (Topology-Independent Loop-Free Alternate) was also successfully tested over SR-MPLS, SR-MPLS with remote and local Shared Risk Link Group (SRLG), and over SRv6. The tests showed great reliability this technology adds to the network.

Flexible Algorithm: Flex Algo is the latest addition to Segment Routing traffic engineering and works with both SR-MPLS and SRv6 data planes. We focused on defining various Flex Algos with different constraints or metrics and observed how these values were flooded via ISIS-TE metric extensions. Then, we verified that traffic forwarding was restricted only to nodes or links participating in that algorithm. Using the dynamic delay measurement and Prefix Metric was a great example of the powerful potentials of this field.

EVPN: This area focused on the Data Center EVPN, Integrated Routing and Bridging (IRB). We observed multi-homing and single homing setups for these network services. Additional EVPN capabilities over BGP as the unified control plane for simplifying network services were all on the agenda, Optimized Inter-Subnet Multicast (OISM), IGMP proxy, and EVPN fault management over CCM (Continuity Check Message).

SDN: This section was dedicated to testing the Path Initiation and Computation techniques, besides managing SR policies through BGP. We also observed an interesting On-Demand Next-hop test. This area of testing is still immature regarding the PCEPv6 and SRv6. We hope to test these technologies in the next year's event.

Flexible Ethernet: This year, FlexE focused on Channelization and Physical Isolation, Bonding, and Dynamic Bandwidth Adjustment to optimize the use of fiber capacity. Channelized techniques such as channel isolation, bonding, and sub-rating were all key parts of this test. We created SR-MPLS overlay over FlexE in a multi-vendor environment.

Clock Synchronization: This area has always been a rich part of our event; this year wasn't an exception. Although we had abandoned some traditional, more functionality-oriented tests, we could test more real-life scenarios and new technologies. For the first time at our event, we carried out individual tests for Boundary Clocks Class C/D, which enabled the passed devices to participate in a chain of Class C or D tests. Enhanced SyncE was a new topic this year. We had the opportunity to test it with multiple vendors and got great results. PTP over MACsec was tested, only one vendor supported it. We hope to see more support and more implementations for PTP security next year. PTP over FlexE was also successfully tested, although the PTP packets weren't sent through the FlexE tunnels but used the FlexE enabled interface. We also feel that we must appreciate the Calnex test suite, which allowed us to perform advanced time synchronization tests like the Class C/D tests, and FlexE with a compelling and direct automated testing process.

Interoperability Test Results

This white paper documents only positive results (passed test combinations) individually with vendor and device names. Failed test combinations are not mentioned in diagrams; they are referenced anonymously to describe the state of the industry. Our experience shows that participating vendors quickly 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—beta—solutions and enables a safe environment 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.

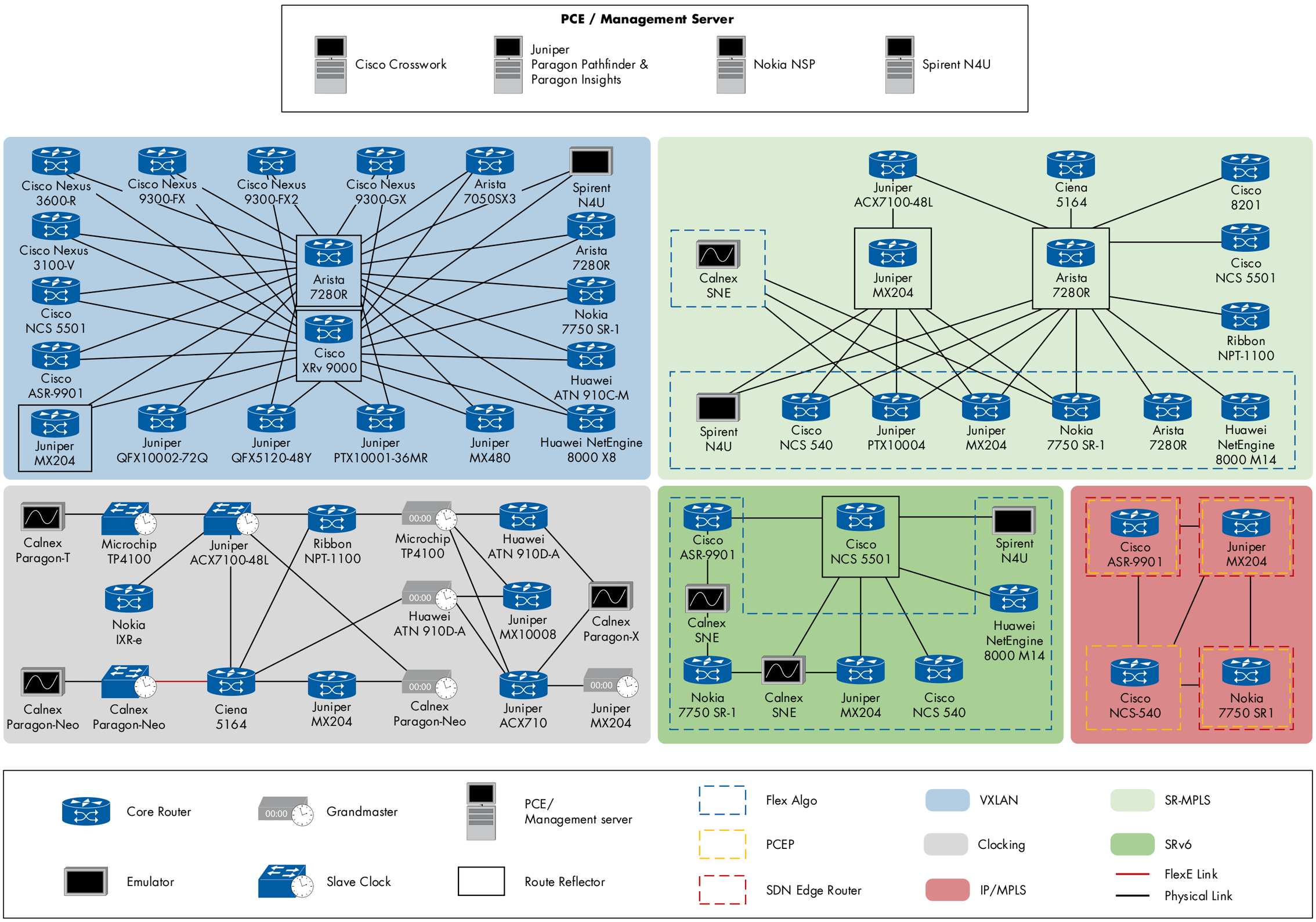
Test Equipment

With the help of participating test equipment vendors, we generated and measured traffic, emulated and analyzed control and management protocols, and performed clock synchronization analysis. We thank Calnex and Spirent for their test equipment and support throughout the testing.

Participating Vendors and Devices

Vendor	Device
Arista Networks	7050SX3 7280R
Calnex Solutions	Paragon-Neo Paragon-T Paragon-X SNE
Ciena	5164
Cisco	8201 ASR 9901 Crosswork XRv 9000 NCS 540 NCS 5501
Cisco Nexus	3100-V 3600-R 9300-FX 9300-FX2 9300-GX
Huawei Technologies	ATN 910C-M ATN 910D-A NetEngine 8000 M14 NetEngine 8000 M8
Juniper Networks	ACX5448-M ACX710 ACX7100-48L MX10008 MX204 MX480-MPC10E Paragon Pathfinder + Paragon Insights PTX10001-36MR PTX10004 QFX10002-72Q QFX5110-48S QFX5120-32C QFX5120-48Y QFX5210-64C
Microchip	TP4100
Nokia	7750 SR1 IXR-e NSP
Ribbon Communications	NPT-110
Spirent Communications	N4U

Table 1: Participating Vendors and Devices


PCE / Management Server
Cisco Crosswork
Juniper Paragon Pathfinder & Paragon Insights
Nokia NSP
Spirent N4U
Cisco Nexus 3600-R
Cisco Nexus 9300-FX
Cisco Nexus 9300-FX2
Cisco Nexus 9300-GX
Arista 7050SX3
Spirent N4U
Cisco Nexus 3100-V
Arista 7280R
Cisco NCS 5501
Arista 7280R
Nokia 7750 SR-1
Cisco ASR-9901
Cisco XRv 9000
Huawei ATN 910C-M
Juniper MX204
Juniper QFX10002-72Q
Juniper QFX5120-48Y
Juniper PTX10001-36MR
Juniper MX480
Huawei NetEngine 8000 X8
Juniper ACX7100-48L
Ciena 5164
Cisco 8201
Calnex SNE
Juniper MX204
Arista 7280R
Cisco NCS 5501
Ribbon NPT-1100
Spirent N4U
Cisco NCS 540
Juniper PTX10004
Juniper MX204
Nokia 7750 SR-1
Arista 7280R
Huawei NetEngine 8000 M14
Calnex Paragon-T
Microchip TP4100
Juniper ACX7100-48L
Ribbon NPT-1100
Microchip TP4100
Huawei ATN 910D-A
Nokia IXR-e
Huawei ATN 910D-A
Juniper MX10008
Calnex Paragon-X
Calnex Paragon-Neo
Calnex Paragon-Neo
Ciena 5164
Juniper MX204
Calnex Paragon-Neo
Juniper ACX710
Juniper MX204
Cisco ASR-9901
Cisco NCS 5501
Spirent N4U
Calnex SNE
Huawei NetEngine 8000 M14
Nokia 7750 SR-1
Calnex SNE
Juniper MX204
Cisco NCS 540
Cisco ASR-9901
Juniper MX204
Cisco NCS-540
Nokia 7750 SR1
Core Router
Grandmaster
PCE/ Management server
Flex Algo
VXLAN
SR-MPLS
PCEP
Clocking
SRv6
Emulator
Slave Clock
Route Reflector
SDN Edge Router
IP/MPLS
FlexE Link
Physical Link

Segment Routing

The launch of 5G is in full gearing up, which opens new opportunities for end-to-end business VPN technologies and infrastructure transition. Segment Routing (SR) is a source routing paradigm, widely discussed with network simplification. The test of this event continues to focus on a core question: Is the transport layer ready for the challenges of 5G preparation?

The test goals cover SR with end-to-end business VPN support to confirm wide network services interoperability support; fast reroute as a key part which provides resiliency through TI-LFA technology with improved path reliability; and more features such as fault isolation carried out in both of its versions SR-MPLS and SRv6.

SR-MPLS with VPN Services

SR-MPLS re-uses MPLS data plane by including labels with global significance. The label distribution works over both ISIS and OSPF. The IETF RFC8667 and RFC8665 standards define these route extensions respectively for SR-MPLS. In the test of both versions ISIS extensions and OSPF extensions, we verified the SR-MPLS creation and transport for end-to-end VPN services.

L3VPN over SR-MPLS

We verified L3VPN interoperability over SR-MPLS data plane. We expected that SR-MPLS transports the L3VPN services transparently based on its label-based data plane. The participating PEs successfully established ISIS/OSPF sessions with each other. The ISIS/OSPF routing table displayed the PEs loopback prefixes alongside their corresponding prefix Segment IDs (SID). In addition, all L3VPNs that were full meshed went up and VRF routes included both IPv4 and IPv6 routes carrying a service label.

In the data plane, packets were forwarded with label stack of two labels (SR transport label—if through P node—and VPN service label), or only with single VPN service label on the last link, and with implicit 0 label used by SR transport. All traffic was received at the endpoints without any loss.

The devices that successfully demonstrated interoperability are:

- Arista 7280R (Spine), Ciena 5164, Cisco NCS 5501, Juniper ACX7100-48L, Juniper MX204, Juniper PTX10004, Nokia 7750 SR1, Ribbon NPT-1100, and Spirent N4U

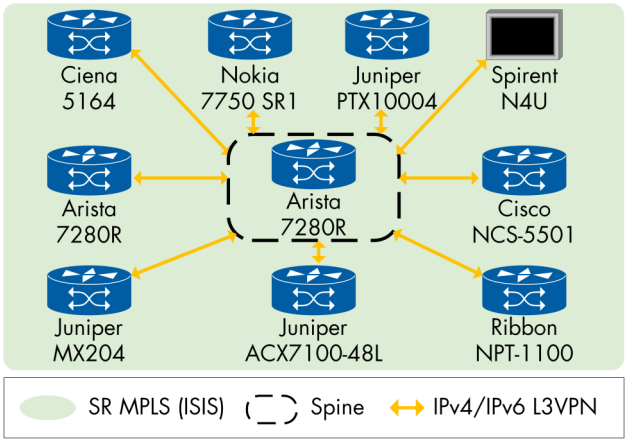


Figure 1: L3VPN Over SR-MPLS (ISIS)

The devices that successfully demonstrated interoperability are:

- Cisco NCS 540, Juniper ACX7100-48L, Juniper MX204 (PE and Spine), Juniper PTX10004, Nokia 7750 SR1, Spirent N4U

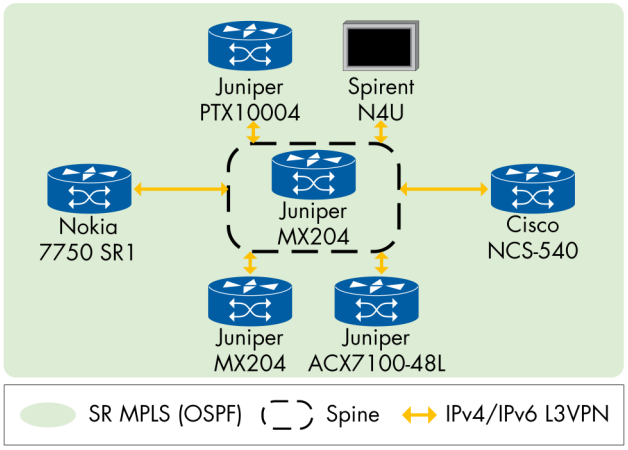


Figure 2: L3VPN MPLS OSPF

IPv6 BGP-LU

The SR domain is known as a set of nodes participating in the source based routing model. Service providers can create different network instances within a SR domain which is administratively maintained as a whole. It is required to include a new protocol for the SID exchanges between interior gateway protocol (IGP) instances. BGP Labeled Unicast (BGP-LU) provides capability to exchange labels across inter-region networks. The egress PE advertises its local loopback via BGP-LU. The gateway router at the area border assigns new BGP-LU labels (Prefix SIDs) in the area for reflected BGP-LU routes, and re-advertises BGP-LU prefixes changing the next-hop to its local loopback (next-hop self). The ingress PE encapsulates packets with this BGP-LU label together with a list of SIDs within the area.

When the packet arrives at the gateway router, it removes the label and encapsulates it with a list of SIDs of the next network. We created two ISIS instances within the same IPv6 IGP domain. All PEs successfully established L3VPN services over the IPv6 BGP-LU. We observed BGP-LU labels exchanged between both ISIS instances, as expected. All traffic was successfully received at the endpoint.

The participating devices were both ingress and bridging nodes:

- Juniper MX204, Nokia 7750 SR1, and Spirent N4U emulated the egress node

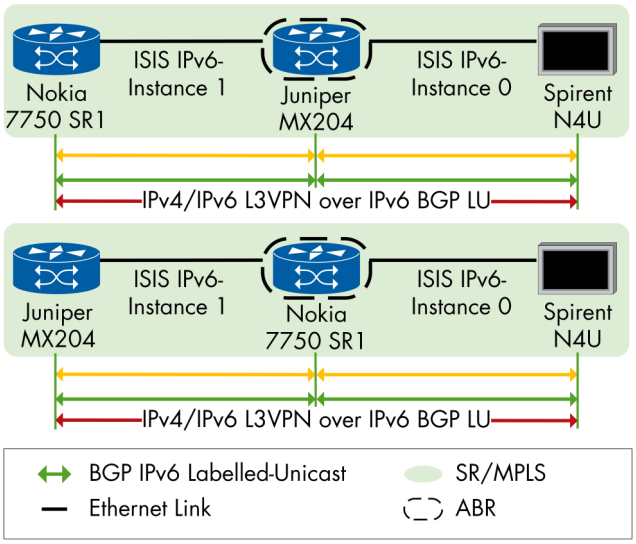


Figure 3: IPv6 BGP-LU

SRv6 with VPN Services

With IPv6 as its data plane, SRv6 becomes a new type of key driver for next generation network. It is the instantiation of SR deployed on the IPv6 data plane. The following tests verified functionality of various VPN services over SRv6.

EVPN VPWS over SRv6

We verified EVPN E-Line over SRv6 in this test. EVPN is a common overlay technology in the data center to create virtual networks on top of Layer 2 and Layer 3 physical networks. EVPN provides powerful control plane and supports Carrier Ethernet services, such as E-Line. The draft-ietf-bess-srv6-services defines EVPN E-Line service over SRv6. Multiprotocol BGP (MP-BGP) EVPN enables communication between hosts in different EVPN segments by distributing Layer 3 reachability information in the form of either a host IP address route or an IP prefix.

We observed single homing PEs that successfully established EVPN Virtual Private Wire Service (VPWS) with each other. For EVPN VPWS single homing, the participating devices were:

- Huawei NetEngine 8000 M14, Nokia 7750 SR1, Spirent N4U as emulated node
- Route Reflector: Cisco NCS 5501

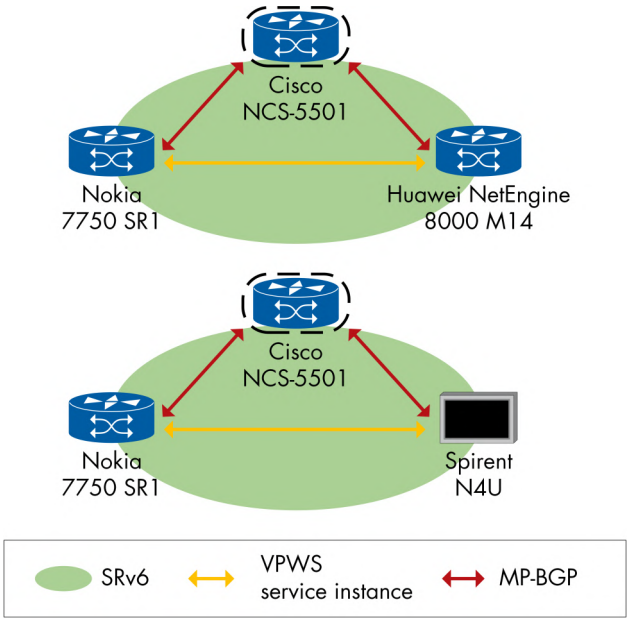


Figure 4: EVPN VPWS Service over SRv6

Initially, we planned to test multi-homing E-Line service over SRv6. However, we observed an interoperability issue during EVPN service establishment. The receiver PE expected to receive Auto-Discovery (AD) per Ethernet Segment Identifier (ESI) route with SRv6 L2 Service TLV, which explains that SRv6 encapsulation is in use. However, the sender only sent the AD per ESI route without any extension, which was dropped by the receiver. Therefore, the EVPN failed to establish, and there were two such cases, so we skipped the multi-homing test.

According to the draft, draft-ietf-bess-srv6-services-07, SRv6 BGP-based Overlay Services (Internet-Draft, 2021), the TLV has the following definition: "A Service SID enclosed in an SRv6 L2 Service TLV within the BGP Prefix-SID attribute is advertised along with the A-D route". The TLV instruction is clear, however, there are different interpretations. One option representing the receiver expects it to be mandatory, as multiple encapsulations might exist in a network, which takes place in a migration phase. In that case, the EVPN shall have a clear choice to identify which encapsulation is required. On the other side, the sender considered this to be optional. We recommend that the standard body takes it as feedback and provides assistance with a clear interpretation.

EVPN Route Type 5 over SRv6

We verified EVPN L3VPN support over SRv6. EVPN extends Layer 2 connectivity to overlay which supports multi-tenancy and is flexible. For tenancy connections across subnets, EVPN provides decentralized gateway function which supports routing based on BGP protocol. Therefore, EVPN also provides L3VPN service model via its unified BGP control plane. The draft-ietf-bess-srv6-services defines EVPN L3VPN over SRv6.

MP-BGP EVPN enables communication between hosts in different EVPN segments by distributing Layer 3 reachability information in the form of either a host IP address route or an IP prefix. BGP Type 5 route operates without an overlay next hop or a Type 2 route for recursive route resolution.

The participating PEs successfully established EVPN L3VPN over the SRv6.

The following DUTs successfully participated in the test, as

- PE: Huawei NetEngine 8000 M14, Nokia 7750 SR1, and Spirent N4U as emulated node
- Route Reflector: Cisco NCS 5501

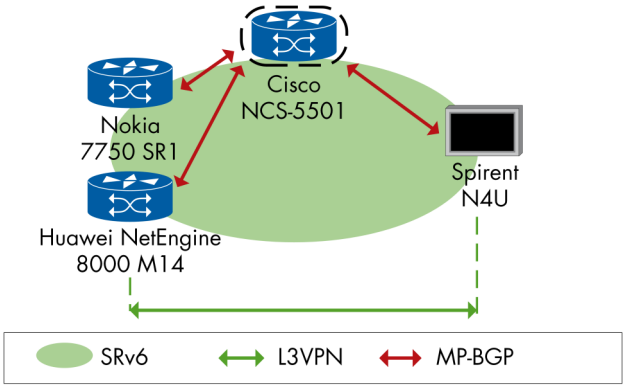


Figure 5: EVPN L3VPN Service over SRv6

IPv4/IPv6 Global Routing Table over SRv6

We verified the correct interoperability between vendors when advertising Internet (Global Routing table) IPv4 and IPv6 prefixes over BGP peerings using IPv6 transport, with a pure IPv6 next-hop in the BGP Next-Hop Attribute and SRv6 information in the Prefix-SID attribute for forwarding over a SRv6 data plane.

The Internet routing table continues to grow, 9.7 million till 2021 per the CIDR report. In addition to the scale perspective, for routers that expect to carry a functional Internet routing table in the transition of a new transport layer, network operators are expected to deploy firstly with the correct functionality.

On IPv6 data plane of SRv6, IPv6-BGP session is capable of carrying both IPv4 and IPv6 prefixes. The SRv6 End.DT6, End.DT4 or End.DT46 SID (identifier of the endpoint with decapsulation and specific IPv6, IPv4 or IP table lookup) carried in the route presents how the route is reachable.

We observed IPv4 and IPv6 prefixes advertised via IPv6 BGP over the SRv6 network. All traffic was received for the advertised routes as expected.

The participating devices were:

- Cisco NCS 5501 (spine), Cisco NCS 540, Cisco ASR-9901, Huawei NetEngine 8000 M14, Juniper MX204, Nokia 7750 SR1, and Spirent N4U as emulated advertising node

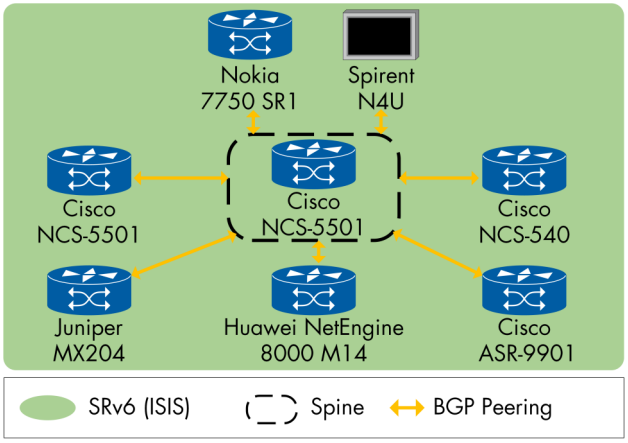


Figure 6: BGP IPv4/IPv6
Global Routing Table over SRv6

L3VPN over SRv6

BGP provides L3VPN signaling capabilities over SRv6. SRv6 Service TLV is an extension of BGP Prefix-SID Attribute to achieve signaling of SRv6 SIDs for L3VPN as defined in draft-ietf-bess-srv6-services. To make the test more interesting, we explored the function bits of SID carried in SRv6 network.

SR leverages the source routing paradigm. An ingress node steers a packet through a ordered list of instructions, called segments, identified by SID. The SRv6 SID is composed of a locator ID plus a function. The function is defined locally on the node where it is executes, and the length is variable, commonly known between 16 and 96 bits.

We verified two options of the SID function bits, 16 bits and 20 bits. All participating PEs successfully established the L3VPN over the SRv6 with its two versions of SID function bits.

The devices that demonstrated interoperability using the 16 bits SID function were:

- Cisco NCS 5501 (spine), Cisco NCS 540, Cisco ASR-9901, Juniper MX204, Nokia 7750 SR1, Spirent N4U as emulated advertising node

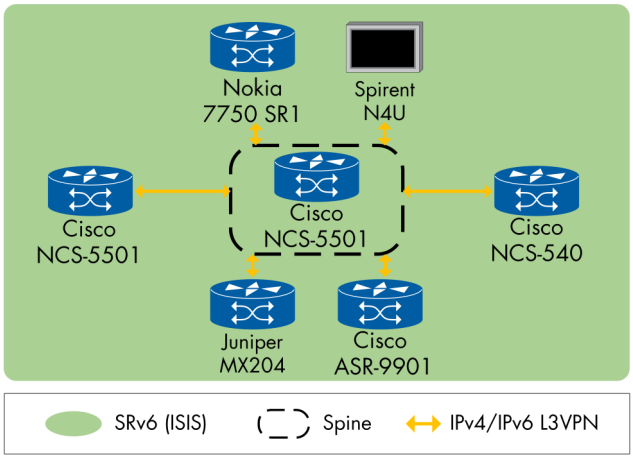


Figure 7: L3VPN over SRv6 with 16 Bits SID Function

The devices involved in the test were:

- Huawei NetEngine 8000 M14, Juniper MX204, Nokia 7750 SR1, Cisco ASR-9901 (spine, route reflector)

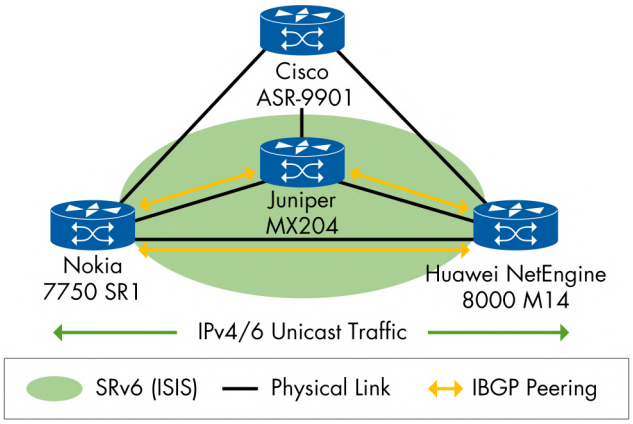


Figure 8: L3VPN over SRv6 with 20 Bits SID Function

When 20 bits function was used some vendors did not accept peer's SID higher than 16 bit even when version upgrade for BGP routes started to accept routes carry high-bit function values, but the data plane did not match.

Segment Routing LSP Ping/Traceroute

The RFC8287 defines the LSP ping and traceroute method for SR with MPLS data plane. Similar to conventional LSP ping/traceroute, the SR fault detection and isolation tools are also based on echo request and echo reply. But SR LSP ping/traceroute include a new TLV type, the Segment ID sub-TLV. On receipt of the sub-TLV carried in an echo request sent by the sender LSR, the LSR responder needs to check the segment ID obtained from the sub-TLV with the local advertised segment ID, to determine if the echo request has been forwarded from the correct path. The LSP ping/traceroute response is carried in an echo reply.

We verified LSP ping and traceroute for SR-MPLS and ping for SRv6 only.

The successfully interoperating devices were:

- Cisco ASR-9901, Huawei NetEngine 8000 M14, Juniper MX204, Nokia 7750 SR1

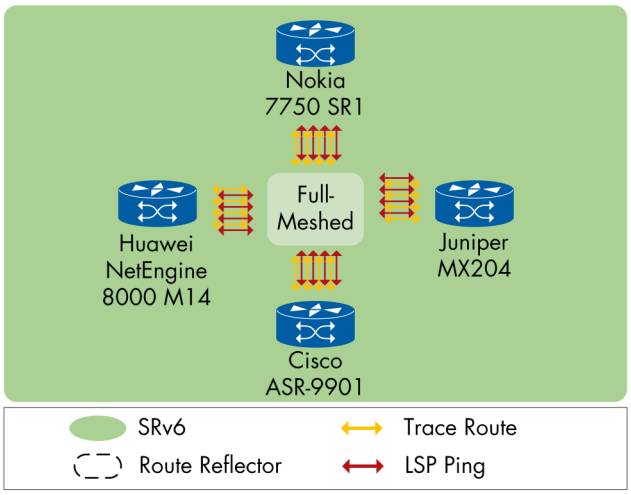


Figure 9: SR LSP Ping/Traceroute with SRv6

The following devices successfully participated in the test:

- Cisco NCS 5501, Huawei NetEngine 8000 M14, Juniper PTX10004, Nokia 7750 SR1

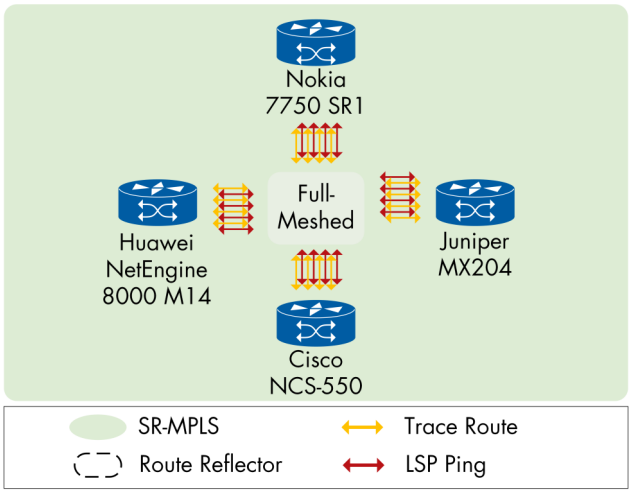


Figure 10: SR LSP Ping/Traceroute with MPLS

In addition to the previous mesh networks, the pairs shown in Table 2 were successfully tested with DUT1 sending ping and traceroute to the respected DUT2.

One test pair showed malformed TLV types, failing to send either an echo request or an echo response. We excluded this pair from the test.

Segment Routing Anycast Advertisement

Anycast describes a set of SR capable routers carrying the same SID. It simplifies the network design through load-balancing in servicer provider networks. We verified that anycast SID is exchanged in the SR network and traffic, in which traffic is load-balanced to the nodes within the anycast network.

We created two Anycast networks to verify the isolation between them. We observed that Anycast SIDs were learned in the segment routing network. By inserting a SID list including the Anycast SIDs, we observe that the traffic entered the expected Anycast networks respectively, and showed load sharing in the Anycast network as expected.

The following DUTs successfully participated in this test:

- Arista 7280R, Ciena 5164, Ribbon NPT-1100, and Spirent N4U as traffic generator

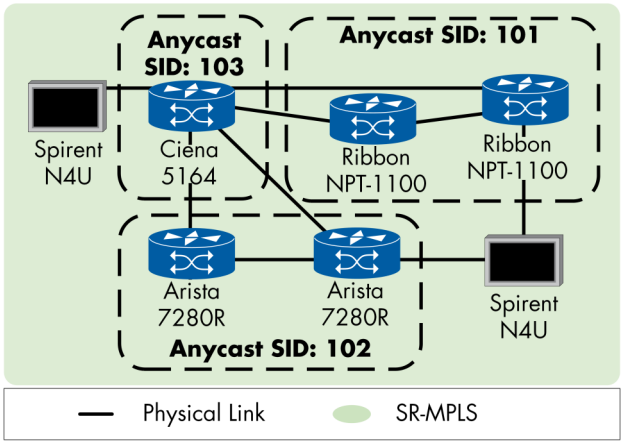


Figure 11: SR Anycast Advertisement

DUT1	DUT2
Ciena 5164	Huawei NetEngine 8000 M14
Huawei NetEngine 8000 M14	Nokia 7750 SR
Arista 7280R	Ciena 5164
Ribbon NPT-1100	Arista 7280R
Huawei NetEngine 8000 M14	Ribbon NPT-1100
Cisco NCS-5501	Arista 7280R, Huawei NetEngine 8000 M14, Ribbon NPT-1100

Table 2: Test Pairs of LSP Ping/Traceroute with SR-MPLS

Topology-Independent Loop-Free Alternate

Segment Routing aims to be a transport technology that support services with tight Service Level Agreement (SLA) guarantees. Therefore, SR must provide a local repair mechanism capable of restoring end-to-end connectivity in case of a link failure, node failure, and local/remote Shared Risk Link Group (SRLG) failure. In case of a link failure, the destination is protected against the failure of a link. In the node failure scenario, the destination is protected against a failure of a node on the primary path. The SRLG protection describes the situation, in which the destination is protected assuming that a configured set of links sharing fate with the primary link has failed. Topology-Independent Loop-Free Alternate (TI-LFA) relies on SR to build a protection mechanism based on proven IP-FRR concepts. TI-LFA does not require any additional signaling between the Point of Local Repair (PLR) and the repair node — typically called PQ node.

TI-LFA over SR-MPLS

We measured convergence time of TI-LFA link protection and verified the local SRLG feature. The local SRLG protection consists of choosing a reroute path that does not include SRLG links because they might share the same risk as the protected path. We also confirmed the support of the local micro-loop prevention. Micro Loop happens when various nodes in the network have different convergence times, and when loop duration is longer than their TTL, it may cause traffic loss. We built a full-mesh topology consisting of four nodes to test link and SRLG TI-LFA over the SR-MPLS network. The participated vendors configured the network nodes with an L3VPN service. Prior to the link failure, the ingress PE (PLR) forwarded the traffic to the directly connected egress PE. To simulate the link failure, we asked the vendor of egress PE to disconnect the link between egress and ingress nodes (the protected link), simultaneously the traffic was still flowing from the traffic generator toward the ingress PE. We observed 34ms out of service time of link protection. The expected time was under 50ms. The out of service time of local SRLG was 4ms-29ms. Expected was under 50ms.

The following DUTs successfully participated in TI-LFA over SR-MPLS while supporting local micro-loop prevention:

- PLR: Ribbon NPT-1100, Ciena 5164
- PQ: Arista 7280R, Huawei NetEngine 8000 M14

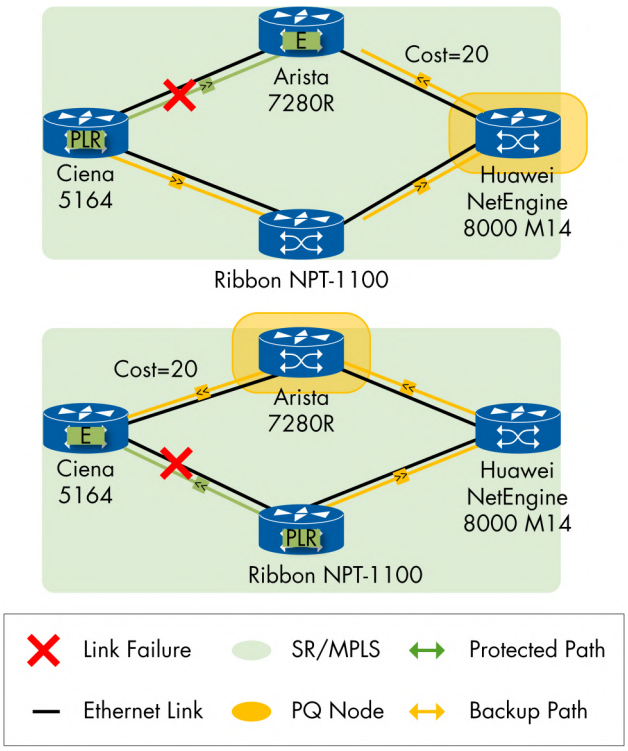


Figure 12: TI-LFA with Link Protection

The following DUTs successfully participated in TI-LFA local SRLG while supporting local micro-loop prevention:

- PLR (with local SRLG): Arista 7280R, Huawei NetEngine 8000 M14
- PQ: Ribbon NPT-1100, Ciena 5164

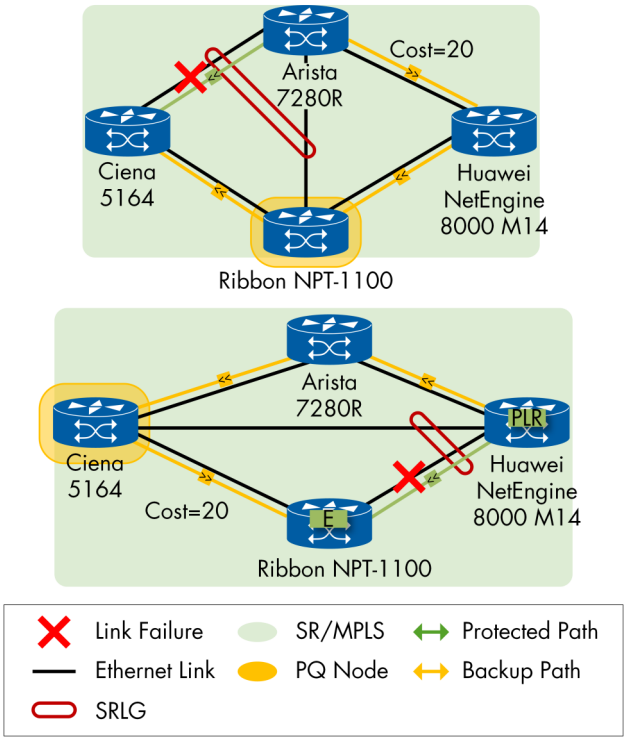


Figure 13: TI-LFA with Local SRLG

TI-LFA with Local and Remote SRLG Protection over SR-MPLS

We verified network convergence of end-to-end VPN services using TI-LFA with remote SRLG protection.

"draft-ietf-rtgwg-segment-routing-ti-lfa-05" defines that in SRLG protecting mode, the destination is protected assuming that a configured set of links sharing fate with the primary link has failed. A local SRLG protecting backup considers only the directly connected links while computing the backup path. On the other hand, a remote SRLG protecting backup also considers links not directly connected to the PLR.

We measured 1ms-10ms out of service time. Expected is a time under 50ms. Table 3 shows the devices that successfully participated in the test.

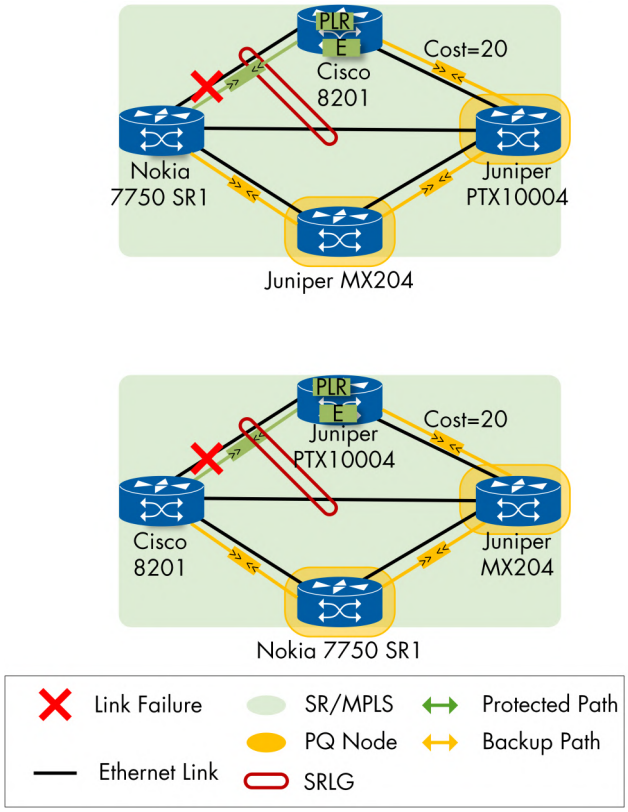


Figure 14: TI-LFA Remote and Local SRLG

TI-LFA with SRv6

The next step is to verify the TI-LFA local SRLG with the SRv6 data plane.

Full mesh topology was implemented and all vendors took turns in participating as PLR nodes while enabling local or remote SRLG as well.

We measured 1ms-38ms convergence time of TI-LFA over SRv6. Expected is a time under 50ms.

The participating nodes were:

- Huawei NetEngine 8000 M14, Juniper MX204, and Nokia 7750 SR1 with local SRLG protection
- Cisco ASR-9901 with remote SRLG protection

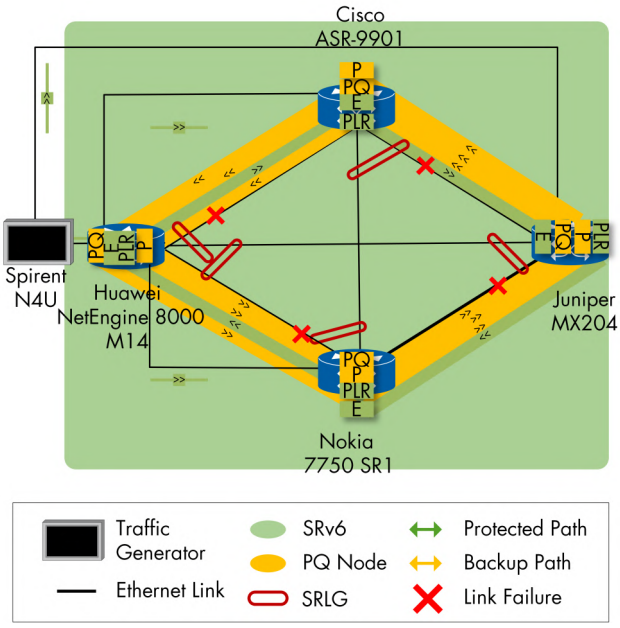


Figure 15: TI-LFA with SRv6

Local SRLG Ingress/ Remote SRLG Egress	Local SRLG PQ	Remote SRLG Ingress/ Local SRLG Egress	Remote SRLG PQ
Cisco 8201	Juniper MX204	Juniper PTX10004	Nokia 7750 SR1
Nokia 7750 SR1	Juniper PTX10004	Cisco 8201	Juniper MX204

Table 3: Test Pairs of TI-LFA Remote and Local SRLG (SR-MPLS)

Flexible Algorithm

Network operators are facing the challenge of 5G slicing, especially to adapt transport to application needs so that the same underlay network can support logical separation and isolation.

The combination of Segment Routing with Flexible Algorithm (Flex Algo) allows network operators to solve arising issues of not only transport for 5G networks but also empowering operators with increasing control over the service quality (delay or bandwidth) guarantee and routing.

The current trends of networking offer with policy intelligence a layered testing of multi-Flex Algo plane networks. The test goals consisted of three parts, namely the creation of underlay IGP, which successfully presented attribute extensions in the network to gain insight into the data required by building Flex Algo. The Flex Algo basic functionality tests included the creation of Flex Algo multi-planes with various policy constrains, both common technologies SR-MPLS and SRv6 were on board. At last, the advanced Flex Algo feature test consisted of prefix metric propagation, including multiple IGP instances.

Underlay ISIS for SR-MPLS with Flex Algo

We verified that ISIS TE metric extensions RFC5305 and RFC8570 were exchanged to flood link attributes such as delay and affinity attributes. Next generation networks define the IGP underlaying network, not only to build the SIDs required by the SR, but more specifically, operators can extract from its extended attributes the information that application requires to separate data planes (low latency, disjoins path).

We created an ISIS-SR topology among participating vendors, together with a defined set of metrics on each DUT. Vendors configured static link delay values, TE metric and Administrative Group. We then verified that these values were distributed via ISIS TE metric extensions.

Metric	Value
IGP metric	102
TE metric	20
Administrative group	blue (marked in Fig 16)
Unidirectional delay metric	11 ms

Table 4: Values of the Used Metrics

The information distributed using IS-IS TE Metric Extensions can then be used to make path-selection decisions based on network performance.

Therefore, we observed the metric value from the network to ensured that it was equal to configured. Once ISIS sessions went up, we collected the routing table via CLI from each DUT. As expected, we observed delay Sub-TLV exchanged over the ISIS underlay. We also observed administrative groups formed in the network. The displayed value was equal to the configured value.

The following DUTs successfully participated as ISIS router with extensions:

- Arista 7280R, Cisco NCS 5501, Huawei NetEngine 8000 M14, Juniper MX204, Juniper PTX10004, and Nokia 7750 SR1

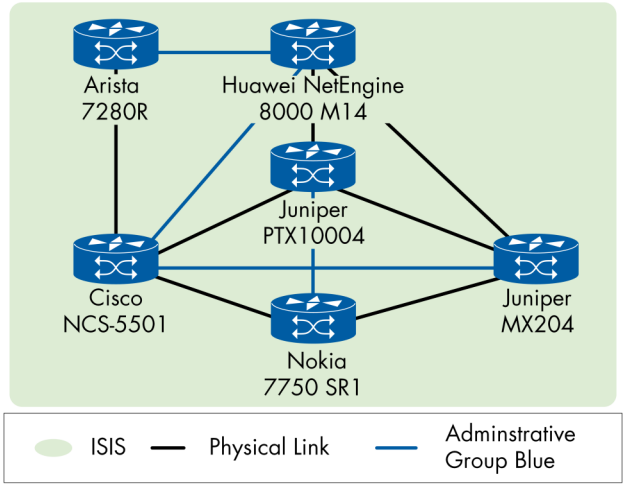


Figure 16: ISIS for SR-MPLS with Flex Algo

Link Delay Measurement in ISIS Underlay

We verified the ability of the underlay network to perform link delay measurement and subsequently link delay propagation in ISIS. One major part of this test included the previously created ISIS topology for SR-MPLS with Flex Algo, at where to extend this functionality. Another significant part of this test included creating a new ISIS topology for SRv6 with Flex Algo and using the delay metric as main focus of the new topology. The Two-Way Active Measurement Protocol (TWAMP) is a protocol defined by the IETF (RFC5357) that was initially designed for IP networks. Now it is commonly used in SR for delay measurement. It performs two-ways delay measurements in a continuous fashion. The TWAMP light version reduces the control protocol overhead, allowing for a more compact implementation.

Flex Algo	Link Metric
138	Low delay metric

Table 5: Flex Algo Definition for SRv6

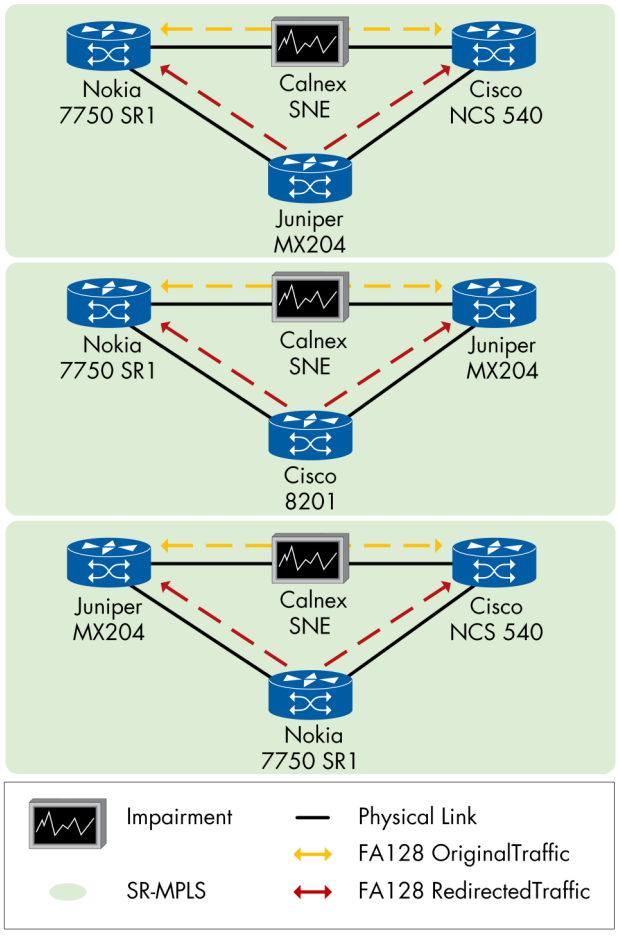


Figure 17: SR-MPLS with TWAMP

The following devices successfully participated in the SR-MPLS with TWAMP test:

- Cisco NCS 540, Juniper MX204, Nokia 7750 SR1, and Cisco 8201 as transit node

The following devices successfully participated in the SRv6 with TWAMP test:

- Cisco ASR-9901, Huawei NetEngine 8000 M14, Juniper MX204, and Nokia 7750 SR1

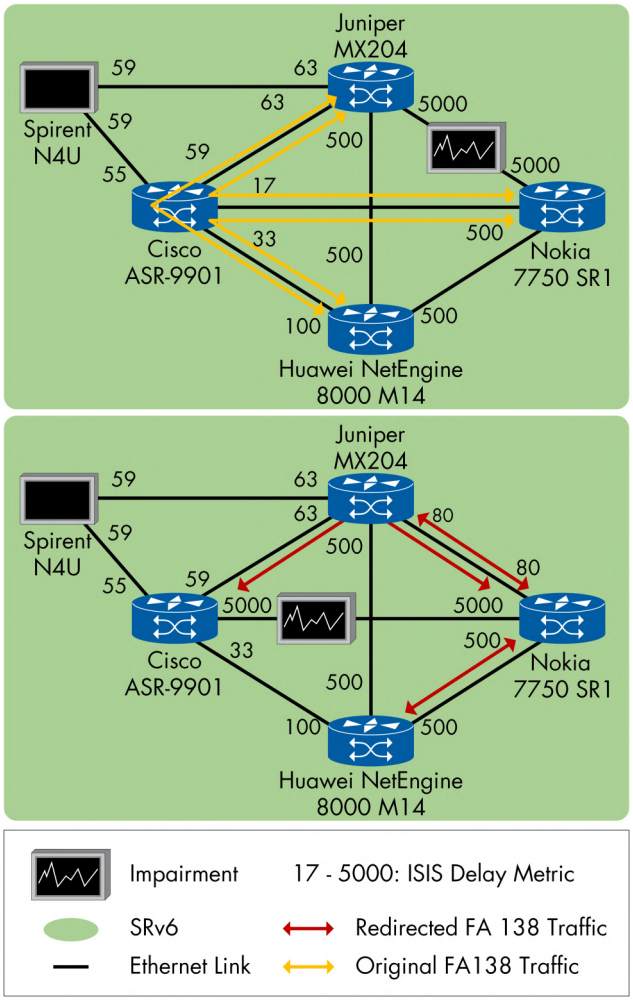


Figure 18: SRv6 with TWAMP

Initially, we observed failed TWAMP session due to bad checksum received. The vendor soon fixed the issue by adding a patch.

SR-MPLS with Flex Algo

Segment Routing with Flex Algo is currently one of the most promising developing fields in the 5G slice architecture. According to 5G context, network slicing is the collection of a set of technologies to create dedicated logical networks as a service. For a transport network, network slicing requires the underlying network to support partitioning of the network resources to provide application-specific networking and computing from a shared pool. The slices may be seen as virtual networks.

Flex Algo is a solution that allows IGPs to compute constraint-based paths over the network, where a router can then associate one or more SR Prefix-SIDs or SRv6 locators with a particular Flex Algorithm. Each such Prefix-SID or SRv6 locator then represents a path that is computed according to the identified Flex Algorithm.

The foundation of Flex Algo is based on a single octet Flex Algorithm identifier with a value between 128 and 255 carried in SR-Algorithm Sub-TLV. Each router interested to take part in a Flex Algo needs to advertise the Flex Algo in its SR-Algorithm sub-TLV under Router Capability. The Flex Algo specification covers both ISIS and OSPF. In this event, we tested the ISIS underlay for Flex Algo. In addition to the default Flex Algo 0, service providers can combine any Flex Algo with network performance metrics, such as low latency, or with traffic engineering methods (like the use of affinity metrics) to create one or multiple Flex Algo(s) meeting such requirements.

This provide a huge degree of flexibility and adaptability to the network, allowing virtual data plane based on traffic needs.

The goal of the test includes Flex Algo multi-planes and isolation, to allow the existing ISIS underlay network to unfold its full potential.

We created three Flex Algo definitions and expect that the ISIS underlay shall calculate a set of nodes and links of each Flex Algo, and based on the collected performance metric information to form three different Flex Algo planes. We created a L3VPNs in each of the Flex Algo definitions. In different parts of the network, we added traffic and expected each VPN to follow its own Flex Algo path.

Flex Algo 128 – Low Delay

We performed two tests for Flex Algo 128 (low latency), depends on how delay value was configured. Based on the TWAMP-light delay measurement, a group of vendors created Flex Algo 128 using measured delay value, as shown in Figure 17 (SR-MPLS with TWAMP). The delay value collected by the ISIS came from the delay measurement value measured by the DUT via TWAMP-light. This method by default resulted shortest path with less delay to be included in the Flex Algo 128 (as shown in the Table below).

In addition, the delay measurement feature added one more test step, to verify that when the delay value changed dynamically, the creation of Flex Algo 128 was based on the dynamic value.

L3VPN Traffic	Pass-Through (Traffic Path)		
	Flex Algo 128	Flex Algo 129	Flex Algo 130
Arista 7280R –> Juniper MX204	Huawei NetEngine 8000 M14	Cisco NCS 5501 - Nokia 7750 SR1	Huawei NetEngine 8000 M14
Juniper MX204 –> Arista 7280R	Huawei NetEngine 8000 M14	Juniper MX204 - Cisco NCS 5501	Huawei NetEngine 8000 M14
Juniper MX204 –> Cisco NCS 5501	-	Nokia 7750 SR1	-
Cisco NCS 5501 –> Juniper MX204	-	Nokia 7750 SR1	-
Huawei NetEngine 8000 M14 –> Cisco NCS 5501	-	Juniper MX204	-
Cisco NCS 5501 –> Huawei NetEngine 8000 M14	-	Nokia 7750 SR1 - Juniper MX204	-

Table 6: Flex Algo SR-MPLS Policy and Traffic Path

Therefore, we increased with the impairment device 100ms delay over the short path. When the max. delay value increased, although another path was farther but the total delay value was less than the max. delay. We expected the latter link to be selected in Flex Algo 128. As shown in Figure 17 (SR-MPLS with TWAMP), these DUTs participated in this test.

There were also vendors who configured network links with the static delay value, so that ISIS link state collection consisted of static delay value. The Flex Algo 128 therefore included the links with smaller delay value configured.

Flex Algo 129 – Exclude Blue Links

We verified Flex Algo 129 creation based on Affinity link attribute. We expected that none of the blue links shall be included in this Flex Algo.

Flex Algo 130 – Low TE Metric

We verified TE metric distribution and calculation of Flex Algo 130 based on TE metric. We used the same TE value per link as shown in section (Underlay ISIS for SR-MPLS with Flex-Algo). The Flex Algo 130 included the links from the shorted path.

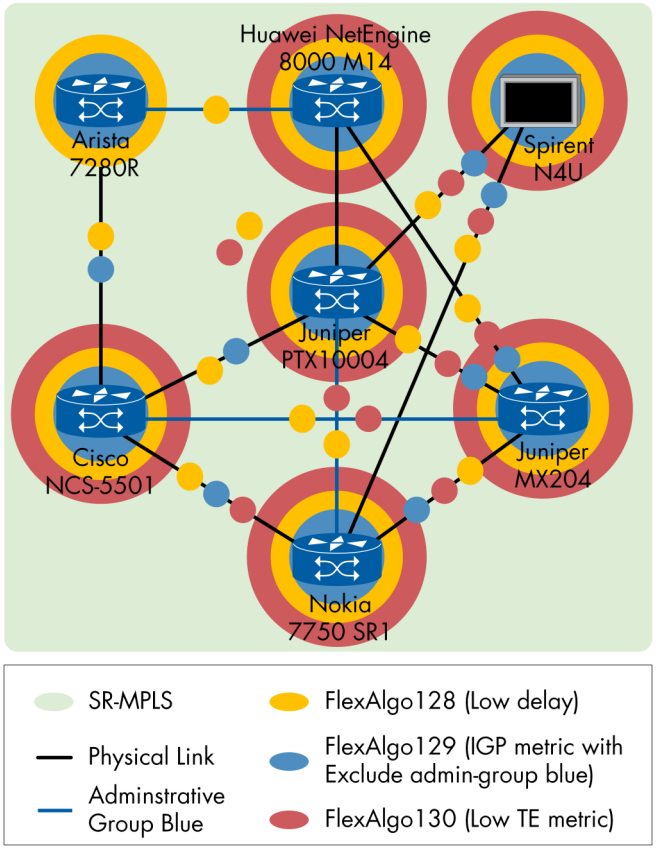


Figure 19: SR-MPLS Flex Algo

The following DUTs successfully participated in the Flex Algo 128, as

- PE with delay measurement: Cisco NCS 540, Juniper MX204, Nokia 7750 SR1
- PE with static delay value: Arista 7280R, Huawei NetEngine 8000 M14, Juniper PTX10004, Spirent N4U
- Impairment Device: Calnex SNE

The following DUTs successfully participated in the Flex Algo 129 (Affinity attribute), as

- PE: Arista 7280R, Cisco NCS 5501, Huawei NetEngine 8000 M14, Juniper MX204, Juniper PTX10004, Nokia 7750 SR1, Spirent N4U

The following DUTs successfully participated in the Flex Algo 130 (TE metric), as

- PE: Cisco NCS 5501, Huawei NetEngine 8000 M14, Juniper MX204, Juniper PTX10004, Nokia 7750 SR1, Spirent N4U

SRv6 with Flex Algo

The Flex Algo solution is applicable to both SR-MPLS as well as SRv6. We verified Flex Algo for SRv6 using the following definition.

Flex Algo	Link Metric
138	IGP metric with Exclude Blue

Table 7: Flex Algo Definition for SRv6

Traffic was flowing between nodes avoiding the blue links as expected. The following vendors successfully interoperated in the test:

- Cisco ASR-9901, Huawei NetEngine 8000 M14, Juniper MX204, Nokia 7750 SR1, Spirent N4U

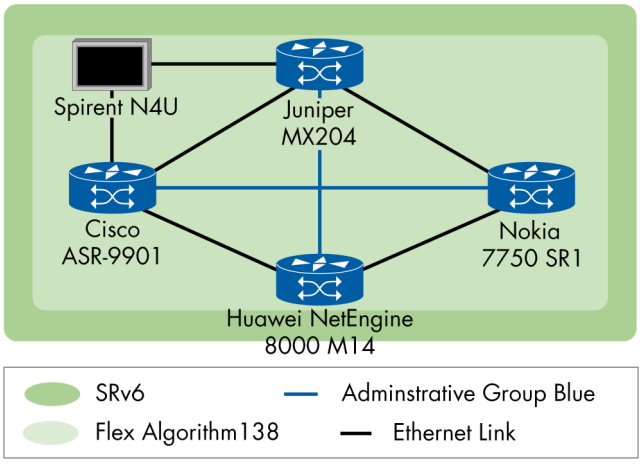


Figure 20: SRv6 Flex Algo with Affinity

Flexible Algorithm Prefix Metric

A service provider IGP may encompass multiple areas. Flexible Algorithm Prefix Metric (FAPM) is a sub-TLV type defined by the IETF draft (draft-ietf-lsr-flex-algo) for ISIS as well as OSPF, to allow the optimal end-to-end path for an inter-area or inter-domain prefix for any Flex-Algorithm to be computed.

Based on the draft definition, M-flag: When set, the Flex-Algorithm specific prefix metric MUST be used for inter-area and external prefix calculation. This flag is not applicable to prefixes advertised as SRv6 locators. We created two ISIS levels in the network for SR-MPLS. By adding a Flex Algo over the underlay topology, we verified that this Flex Algo 140 (based on delay metric) include both levels in the network.

In this scenario black metrics are IGP metrics with value of 10 for each link and red metrics are ISIS delay metrics. Original traffic is following lowest delay and due to inter-level FAPM, delay metric is correctly propagated between levels. This assures correct routing decisions between levels.

Above setup built the basic step of the test. A new step included a redundancy test. We performed a link failure in one of the levels, and expected Flex Algo to recover from the link failure and enter a new part of the network, but it should still include two levels.

Once Flex Algo 140 was up, as expected, SID prefix for the Flex Algo 140 was present in both ISIS levels through the prefix metric exchanged between the PEs. After the link failure introduced in ISIS level 1, we observed that Flex Algo recovered successfully from the link failure.

The following devices successfully participated in this test:

- Cisco NCS 5501, Huawei NetEngine 8000 M14, Juniper MX204, Nokia 7750 SR1, Spirent N4U

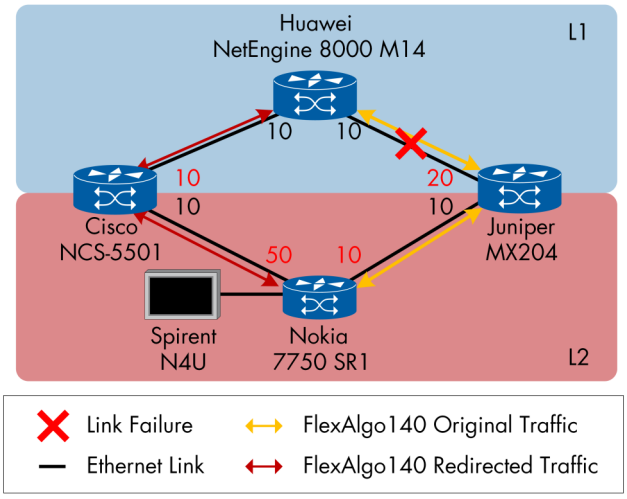


Figure 21: Flex Algo Prefix Metric

EVPN

Data Center and Interconnection

Current trends in data centers expose that the future will include larger data centers. Data centers do not exist isolated as they expand in number and scale. It is significant that applications can take advantage of end-to-end connections over geographically different locations. Reliable connections are critical, so that interconnected data centers, where access to data is critical, perform SLA.

We observed EVPN support in Carrier Ethernet services, providing E-Line and E-Tree services. The unique extension of EVPN to E-LAN services is that it supports integrated services delivery, consisting of Layer 2 and Layer 3 services over the same interface. We verified the Integrated Routing and Bridging (IRB) function which enables EVPN extension between subnets. We observed multi-homing and single homing setups for these network services. Additional EVPN capabilities over BGP as the unified control plane for simplifying network services were all on the agenda, inter-subnet multicast (OISM), IGMP proxy, and EVPN fault management over CCM (Continuity Check Message). One highlight must be to splice the network services together to form interconnected data centers. We observed EVPN and IPVPN interworking, EVPN VXLAN-VXLAN networking, and Seamless EVPN and VPLS.

E-Line Service

Historically, E-Line services are realized as Virtual Private Wire Services (VPWS) using point-to-point (P2P) Pseudo-Wires (PWs). With the growing demand for data center services, the IETF task force BESS built VPWS on EVPN, which provides a powerful VPWS framework suitable for data center design. The advantages of VPWS with EVPN mechanisms are single-active or all-active multihoming capabilities and support for Inter-autonomous system (AS) options associated with BGP-signaled VPNs.

Following the mature data encapsulation technology of MPLS from previous test events, SR-MPLS dominated in this test. It is a label-based encapsulation technology that can support EVPN. Its advantage is to reuse label-based data plane and it does not require changes to the signaling. We created all E-Line services over SR-MPLS. We observed a mix of multi-homing and single-homing PEs from the same site, as well as these mixed in multiple sites. In most combinations, the PEs built a multi-homing in a multi-vendor environment.

In such tests, we performed a redundancy test to prove the link failure protection. The following devices successfully participated in this test, as

- Single-homing PE: Arista 7280R, Cisco ASR-9901, Huawei NetEngine 8000 X8, Juniper MX480, Nokia 7750 SR1, and Spirent N4U
- RR: Arista 7280R and Cisco XRv 9000

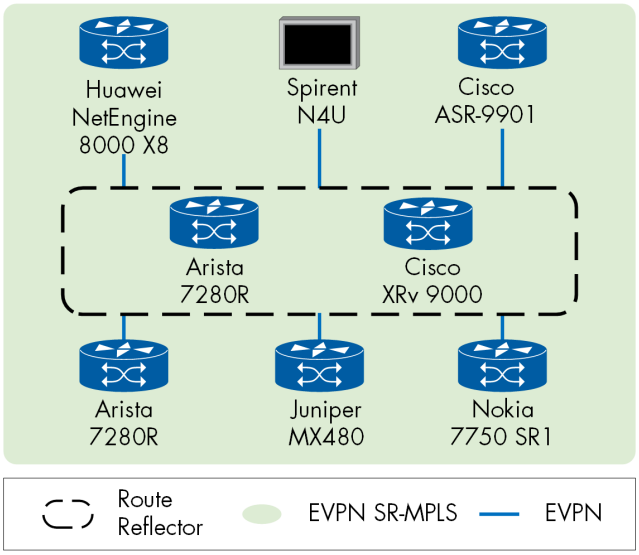


Figure 22: E-Line with EVPN over SR-MPLS

The following devices successfully participated in this test, with

- All-Active Multi-Homing PE: Arista 7280R, Juniper MX480, Cisco NCS 540, and Nokia 7750 SR1
- CE: Arista 7050SX3 and Cisco NCS 5501
- RR: Arista 7280R and Cisco XRv 9000

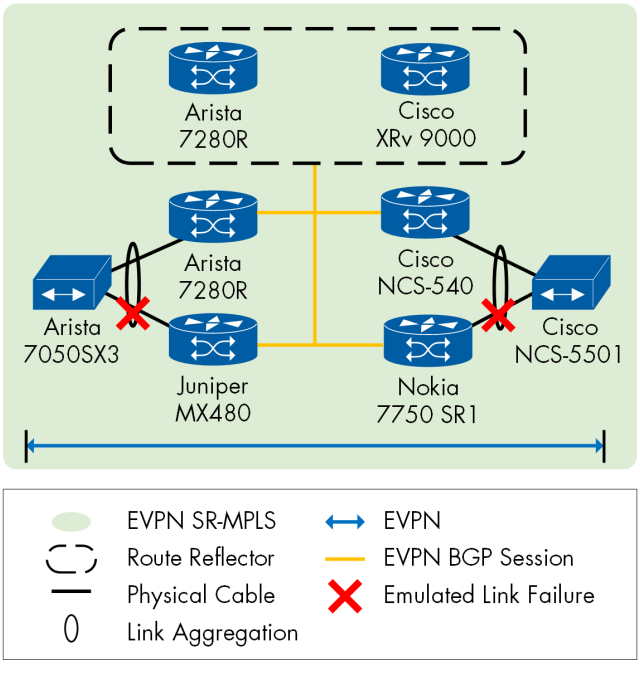


Figure 23: EVPN with All-Active to All-Active Multi-Homing PEs

The following devices successfully participated in the test, with

- All-Active Multi-Homing PE: Arista 7280R, Juniper MX480, Cisco NCS 540, and Nokia 7750 SR1
- CE: Arista 7050SX3 and Cisco NCS 5501
- RR: Arista 7280R and Cisco XRv 9000

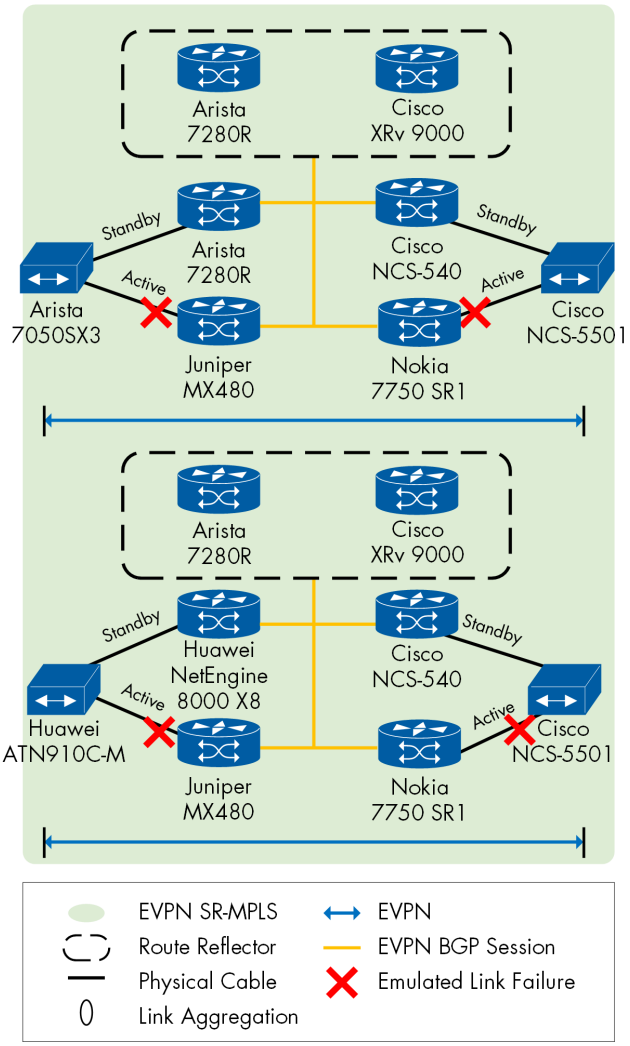


Figure 24: EVPN with Single-Active Multi-Homing Pes

The following devices successfully participated in the test, with

- All-Active multi-homing PE: Cisco NCS 540 and Nokia 7750 SR1
- Single-Active multi-homing PE: Huawei NetEngine 8000 X8 and Juniper ACX5448-M
- CE: Cisco NCS 5501 and Huawei ATN910C-M
- RR: Arista 7280R and Cisco XRv 9000

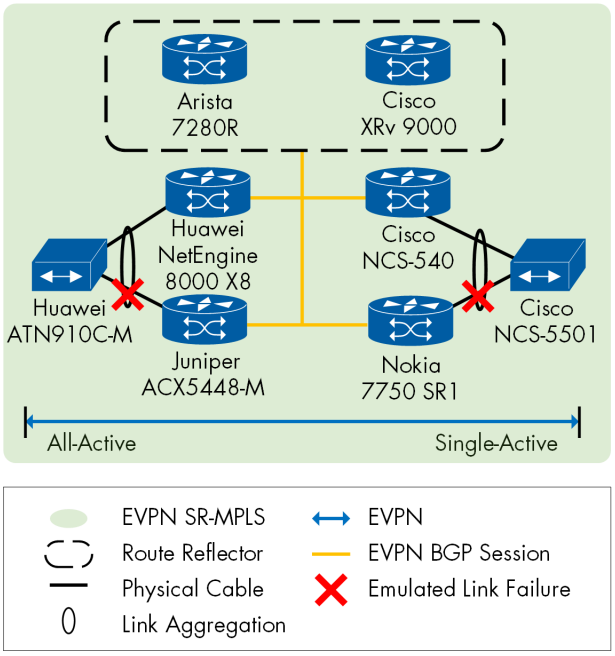


Figure 25: EVPN with a Mix of All-Active and Single-Active Multi-Homing PEs

Flexible Cross-Connect Service

This test verified that multiple attachment circuits (ACs) across different Ethernet Segments and physical interfaces are transported into a single EVPN VPWS service.

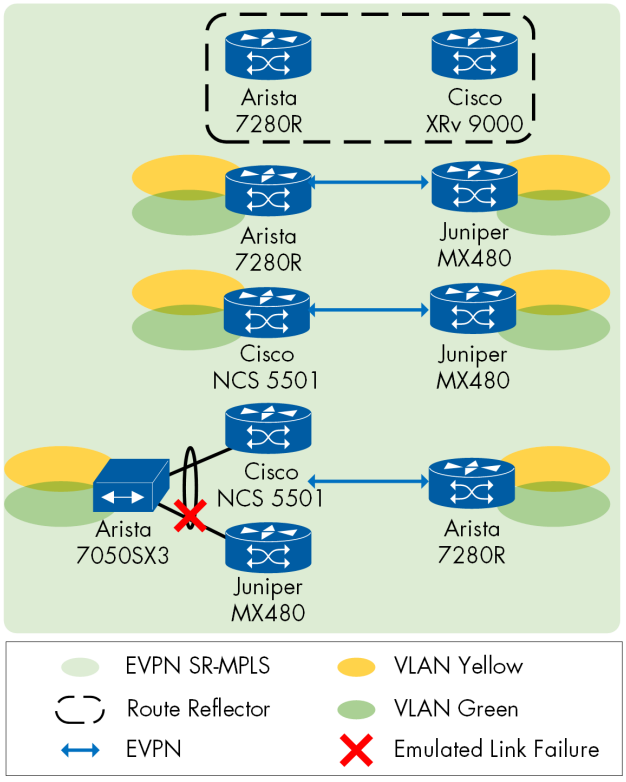


Figure 26: Flexible Cross-Connect Service

The following devices successfully participated in the test, as

- FCS multi-homing PE: Cisco NCS 5501, and Juniper MX480
- FCS single homing PE: Arista 7280R, Cisco NCS 5501, and Juniper MX480
- RR: Arista 7280R and Cisco XRv 9000

Integrated Routing and Bridging

A scalable data center design provides integrated Layer 2 and Layer 3 services, which allow end hosts across the overlay to communicate with each other within or across the EVPN subnets. The Integrated Routing and Bridging (IRB) provides a decentralized gateway approach that has proven itself with BGP route set (RT-2 and RT-5) many times in the event. The focus was on the multi-homing EVPN with IRB. We verified both the symmetric and asymmetric IRB. functionalities, using an EVPN VLAN-based service.

Symmetric IRB

In a multi-tenancy environment, a tenant may span on different leafs. EVPN provides Inter-subnet forwarding IRB to act as l2 extension and layer 3 routing between different Ethernet instances. We verified in this test inter-subnets forwarding with symmetric IRB. The symmetric IRB is one of the common modes of IRB, at where each tenant is assigned to a unique logical connection for IP-VRF. Two BGP route types: RT-2: the MAC and IP Route Type 2 is advertised with both Bridge-Domain/EVI label and IP VRF label with their respective route-targets. RT-5: IP prefix Route, is an alternative solution. A pure type-5 route operates without an overlay next hop or a type-2 route for recursive route resolution.

The identifier is VXLAN network identifier (VNI) in VXLAN data plane and needs to be the same on all peers participating the same tenant's symmetric IRB. In MPLS data plane, this identifier is the MPLS Label2 associated with the IP-VRF. Likewise, this mode is analogous to a Layer 3 routing interface between different switches. We created a full-meshed topology with the DUTs. All EVPN participating devices were present in the topology. We checked RT-2 routes exchanged between the PEs and identified the two VNI/labels for the MAC/IP routes as expected. We also looked at RT-5 routes with subnets exchanged between PEs. All traffic from all endpoints was received at the other endpoints. The following DUTs successfully participated in the symmetric IRB:

- Single homing PE: Arista 7050SX3, Juniper PTX10001, and Juniper QFX5210-64C, and Nokia 7750 SR1
- All-Active multi-homing PE (from a single vendor): Arista 7050SX3, Arista 7280R, Cisco Nexus 3100-V, and Cisco Nexus 9300-FX2
- RR: Arista 7280R and Cisco Nexus 9300-FX

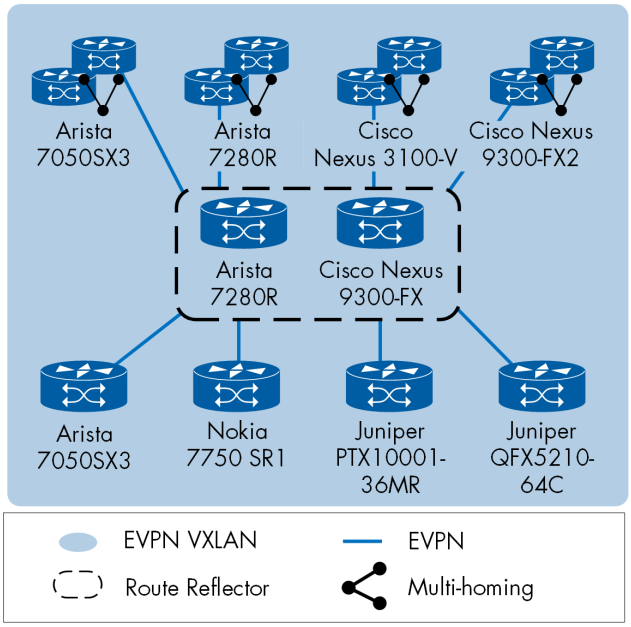


Figure 27: Symmetric IRB over VXLAN

The following devices successfully participated in the IRB test, as

- All-Active multi-homing PE (from a single vendor): Cisco Nexus 3100-V and Juniper QFX5210
- All-Active multi-homing PE (from multi-vendors): Arista 7280R and Nokia 7750 SR1
- CE: Arista 7050SX3
- RR: Arista 7280R and Cisco Nexus 9300-FX

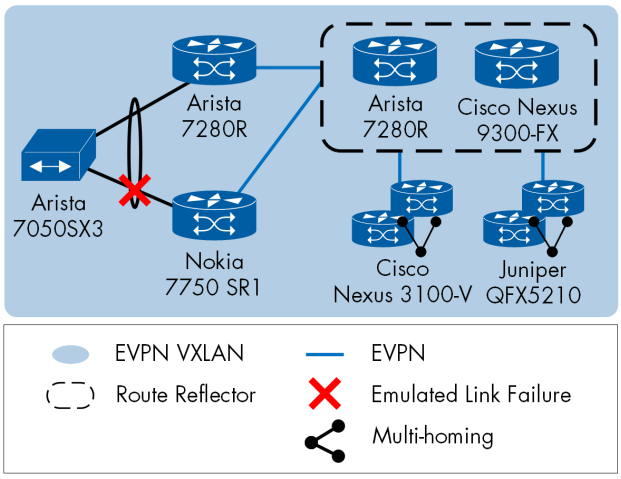


Figure 28: Symmetric IRB with All-Active multi-homing PEs

The following DUTs successfully participated in the IRB over SR-MPLS, as

- All-Active multi-homing PE (from a single vendor): Arista 7050SX3, Arista 7280R, Cisco ASR-9901, and Cisco NCS 5501
- Single homing PE: Huawei NetEngine 8000 X8, Juniper MX480-MPC10E, Nokia 7750 SR1, and Spirent N4U
- RR: Arista 7280R and Cisco XRv 9000

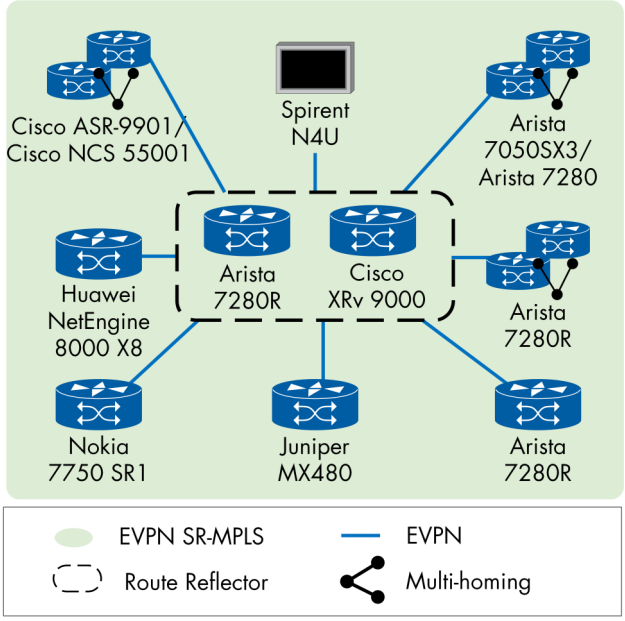


Figure 29: Symmetric IRB over SR-MPLS

The following devices successfully participated in symmetric IRB over SR-MPLS with all-active multi-homing, as

- All-Active multi-homing PE (from multi-vendors): Arista 7280R, Cisco NCS 5501, and Juniper MX480
- Single homing PE: Arista 7280R
- RR: Arista 7280R and Cisco XRv 9000

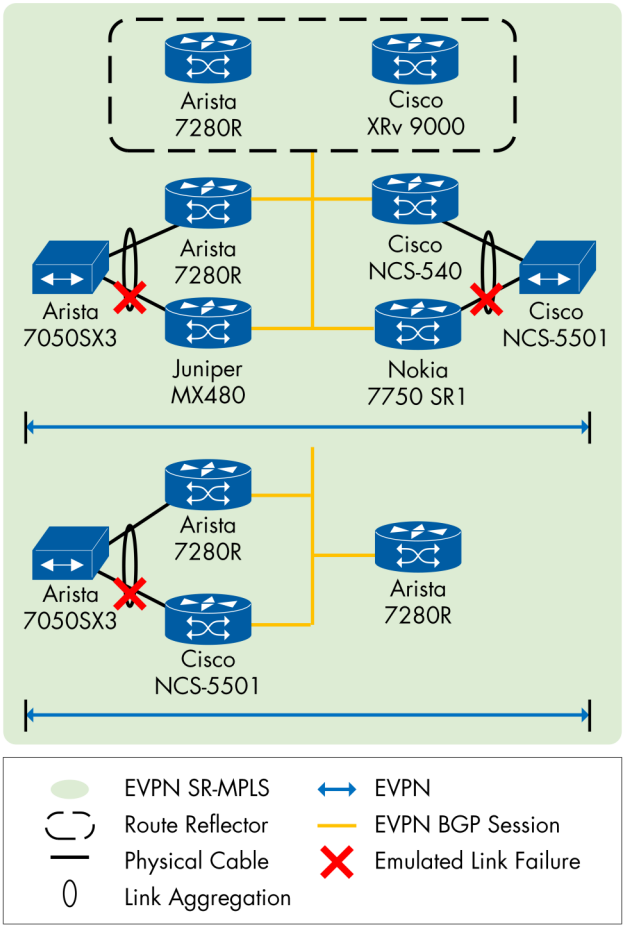


Figure 30: Symmetric IRB over SR-MPLS with All-Active Multi-Homing

Asymmetric IRB

We verified asymmetric IRB functionality. In the asymmetric IRB semantic, both IP and MAC lookups are required at the ingress PE, whereas only MAC lookup is needed at the egress PE. We created a full-meshed topology with the DUTs. We observed RT-2 routes and identified the one VNI/label carried in the route. We also checked that MAC addresses of remote sites were all learned in the MAC table of local PE. All traffic from all endpoints was received at the other endpoints as expected. The following devices successfully participated in PE:

- Arista 7050SX3, Arista 7280R, Cisco Nexus 3600-R, Cisco Nexus 9300-FX2, Juniper QFX10002-72Q, Juniper QFX5120-48Y, and Nokia 7750 SR1

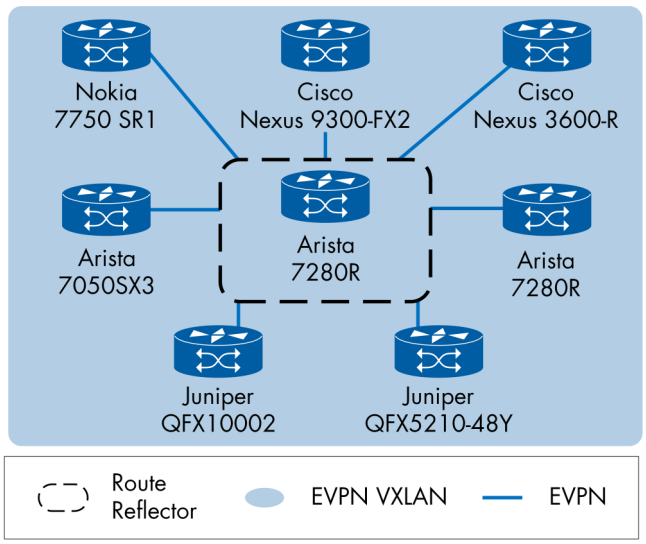


Figure 31: Asymmetric IRB

Proxy MAC-IP Advertisement

We sent traffic from all end-points and observed that it was load-balanced over all links of All-Active Multi-homing PEs. We observed the proxy bit sent in the RT-2 route advertised between these PEs as expected. In addition, MAC addresses were all learned in the All-Active Multi-homing PEs. We performed a link failure test with All-Active multi-homing PEs. As expected, that the EVPN service recovered from the link failure back to normal. The following DUTs successfully participated in the VXLAN setup, as

- All-Active multi-homing PE: Arista 7280R and Juniper QFX5120,
- Single homing PE: Cisco Nexus 3100-V
- CE: Arista 7050SX3
- RR: Arista 7280R

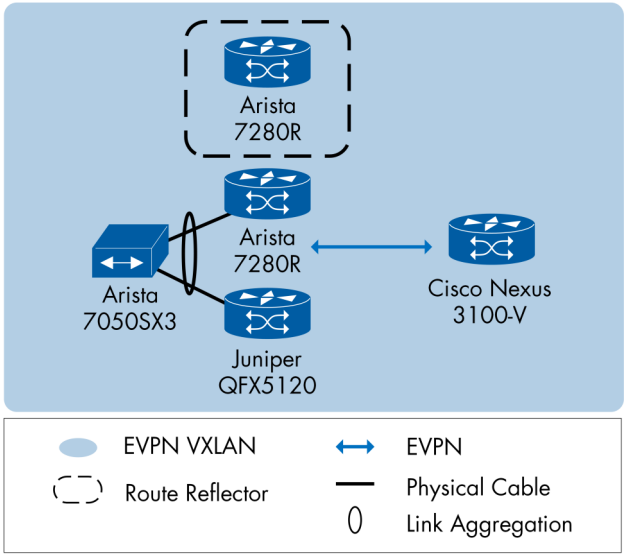


Figure 32: Proxy MAC-IP Advertisement over VXLAN

The following devices successfully participated in symmetric IRB, as

- All-Active multi-homing PE (from multi-vendors): Arista 7280R, Cisco NCS 5501, and Juniper MX480
- Single homing PE: Arista 7280R
- RR: Arista 7280R and Cisco XRv 9000

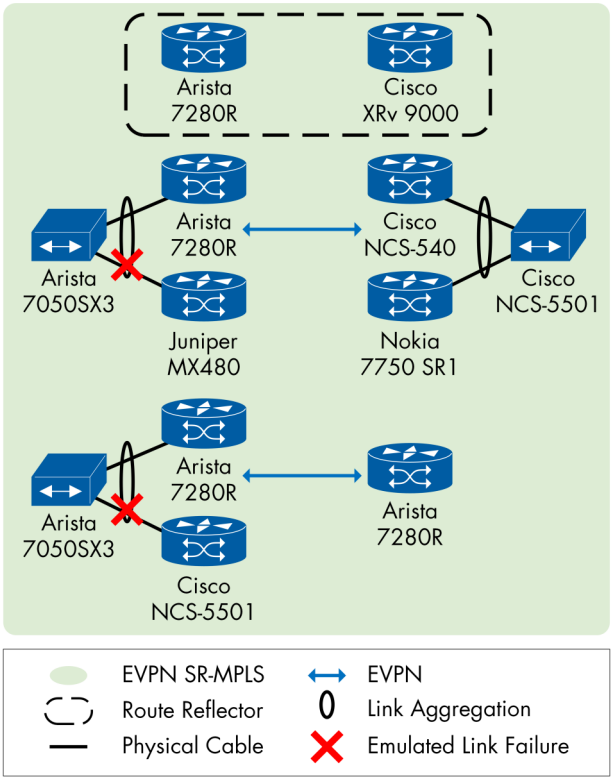


Figure 33: Proxy MAC-IP Advertisement over SR-MPLS

E-Tree Service

E-Tree service, a rooted-multipoint Ethernet service, has endpoints are labeled as either Root or Leaf sites. Root sites can communicate with all other sites. Leaf sites can communicate with Root sites but not with other Leaf sites.

EVPN provides E-Tree support including service's unicast, multicast, and broadcast forwarding as defined in RFC8317. BGP RT attribute supports the service model between the root and leaf endpoints.

We created four setups with the participating DUTs. In each configuration, a different DUT functioned as a root PE, and the other as leaf PEs. We sent traffic consisting of unicast, multicast, and broadcast (BUM traffic) and did not observe any loss between root and leaf. None of the traffic was received between leafs.

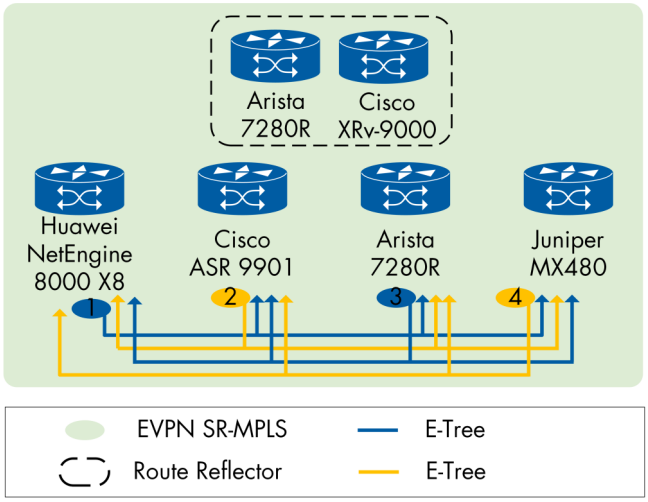


Figure 34: E-Tree with Leaf or Root site per PE

The following DUTs supported E-Tree:

- as root: Arista 7280R, Cisco ASR 9901, Huawei NetEngine 8000 X8, and Juniper MX480
- as leaf: Arista 7280R, Cisco ASR 9901, Huawei NetEngine 8000 X8, and Juniper MX480

IGMP Proxy

The hosts/VMs in a customer domain express their interests in multicast groups on a given subnet/VLAN by sending IGMP membership reports (Joins) for their interested multicast group(s). An IGMP router (e.g., IGMPv1) periodically sends membership queries to find out if there are hosts on that subnet still interested in receiving multicast traffic for that group.

The goal of the IGMP proxy mechanism is to reduce the flood of IGMP messages (both Queries and Reports) in EVPN instances among PE Routers. Furthermore, if there is no physical/virtual multicast router attached to the EVPN network for a given (*,G) or (S, G), it is desired for the EVPN network to act as a distributed anycast multicast router for all the hosts attached to that subnet.

We verified IGMP Proxy functionalities used to optimize the core network with Selective Multicast Ethernet Tag (Type 6: SMET) Routes and and synchronize the multicast state on Ethernet Segments with the EVPN routes type 7 (Multicast Join Sync Route) and type 8 (Multicast Leave Sync Route).

The leaf part of the setup includes emulated subscribers attached to two PEs. Both PEs shall register interest for the same multicast group and express their interest in the ingress PE, which performs ingress replication of multicast traffic.

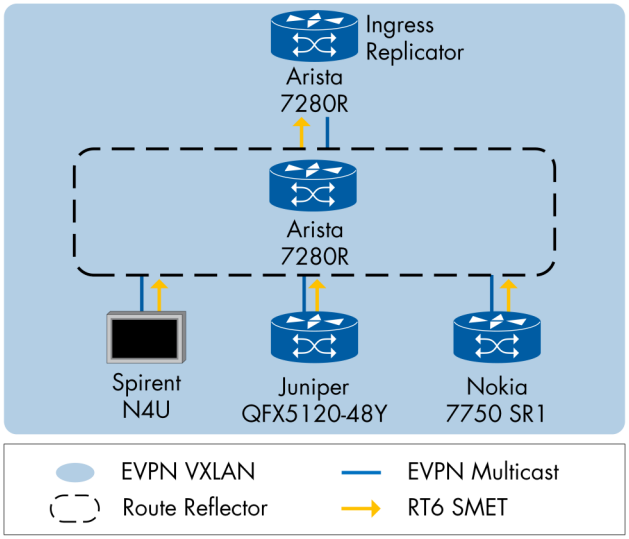


Figure 35: IGMP Proxy

The following DUTs successfully participated in the IGMP setup, as

- Ingress Replicator: Arista 7290R
- Leaf PE: Juniper QFS5120, Nokia 7750 SR1, and Spirent N4U

Optimized Inter-Subnet Multicast

A tenant may span across multiple networks, and the mechanism of unicast traffic forwarding from one network to another is defined by IRB draft (draft-ietf-bess-evpn-inter-subnet-forwarding).

Multicast traffic forwarding is Classified as BUM traffic in this document but is addressing unicast paths. The optimized inter-subnet multicast (OISM) draft (draft-ietf-bess-evpn-irb-mcast) defines the OISM forwarding with optimized multicast paths. OISM enables the creation of multicast distribution trees from upstream to downstream PEs and the efficient forwarding of layer-3 multicast traffic between subnets.

We used Ingress Replication mode in this multicast setup. The ingress PE (connected to the multicast source) duplicates multicast traffic based on interest in the multicast group received. The participating PEs established PMSI (P-Multicast Service Interface) tunnel with each other based on the RT-3 route (Inclusive multicast Ethernet Tag route). The PEs sent and learned RT6 SMET (Selective Multicast Ethernet Tag Route) in each domain for interested multicast groups. We sent multicast traffic from the emulated source, the ingress replicator forwarded the multicast traffic to all egress PEs. After receiving the multicast traffic, the PE re-encapsulated the traffic and forwarded it to the corresponding domain without any packet loss.

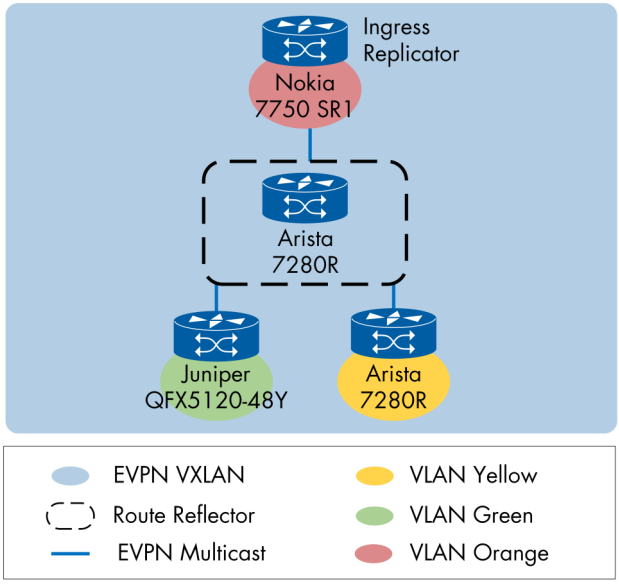


Figure 36: Optimized Inter-Subnet Multicast

The following DUT successfully participated in the OISM setup, as

- Ingress replicator: Arista 7280R
- Egress PE: Juniper QFX5120-48Y, Nokia 7750 SR1

EVPN and IP-VPN Interworking

Modern data center networks have evolved into a history of cross-domain interconnection collection, that can be seamlessly integrated. Gateway services simplify the border design and require more flexibility to adapt to the growth of border components to achieve interoperability. Individual gateway PEs handle not only basic functions of EVPN to extend, such as route selection, encapsulation conversion between different data planes like SR-MPLS, IP/MPLS and VXLAN, but also interworks in a group of PEs to provide service availability with service redundancy via multi-homing.

All gateway PEs under this test met these requirements in an inter-domain topology across SR-MPLS, IP/MPLS and VXLAN. In order to make this test more interesting, we observed another variation of the BGP community for loop prevention to see if the new attribute affects complexity, or perhaps accelerates to launch the test scenario.

The BGP path attribute faced challenges, with its Classic community attribute, which needed to quickly find peer support in an All-Active multi-homing gateway PEs scenario, but also to complete the configuration for loop prevention and launch test in a strict test hours span. Before D-PATH gains wide support with its lightweight configuration logic and dominates this role in the future.

D-PATH is optional and transitive BGP path attribute as specified in draft "EVPN Interworking with IPVPN" (draft-ietf-bess-evpn-ipvpn-interworking). Similar to AS_PATH, D-PATH is composed of a sequence of Domain segments. As an example, a BGP route received with a D-PATH attribute containing a domain segment of {length=2, <6500:2:IPVPN>, <6500:1:EVPN>} indicates that the route was originated in EVPN domain 6500:1, and propagated into IPVPN domain 6500:2. In ISF route received by a gateway PE with a D-PATH attribute that contains one or more of its locally associated domains for the IP-VRF is considered to be a looped ISF route and MUST NOT be installed in that IP-VRF.

Since D-PATH is a useful tool to provide end-to-end visibility across multiple domains, we observed shortened man-hours from design to completion of configuration to launch the test.

We verified the EVPN routes of different BGP families generated between all three domains to confirm the control plane interoperability. We also observed the in and out of traffic at each active PE, indicating that the traffic is load-balanced for all All-Active multi-homing PEs. We observed all-active multi-homing gateway PEs with the BGP community for loop prevention.

All DUTs successfully participated in the test.

- Gateway PE between SR-MPLS and IP/MPLS: Arista 7280R, Cisco NCS 5501, and Juniper MX480
- Gateway PE between IP/MPLS and VXLAN: Arista 7280R and Cisco ASR-9901

All PE and P devices successfully participated in the test

- RR: Arista 7280R, Cisco XRv 9000, and Juniper MX204
- SR-MPLS PE and VXLAN PE: Spirent N4U

Two participants successfully joined with all-active multi-homing PE using D-PATH for loop prevention.

- Gateway PE between SR-MPLS and IP/MPLS: Arista 7280R and Nokia 7750 SR1
- Gateway PE between IP/MPLS and VXLAN: Arista 7280R and Spirent N4U

The PE and P devices successfully participated in the test:

- RR: Arista 7280R
- PE: Spirent N4U

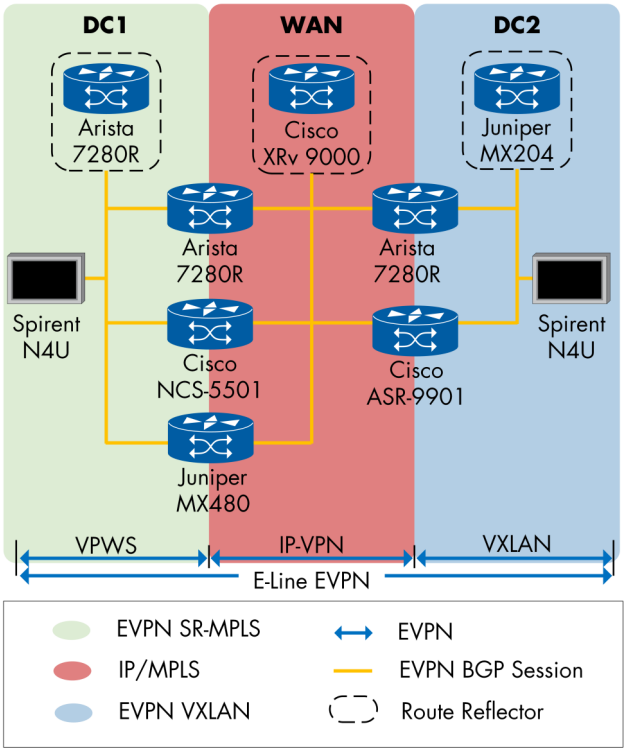


Figure 37: All-Active Multi-Homing Gateway PEs with Loop Prevention via BGP Community

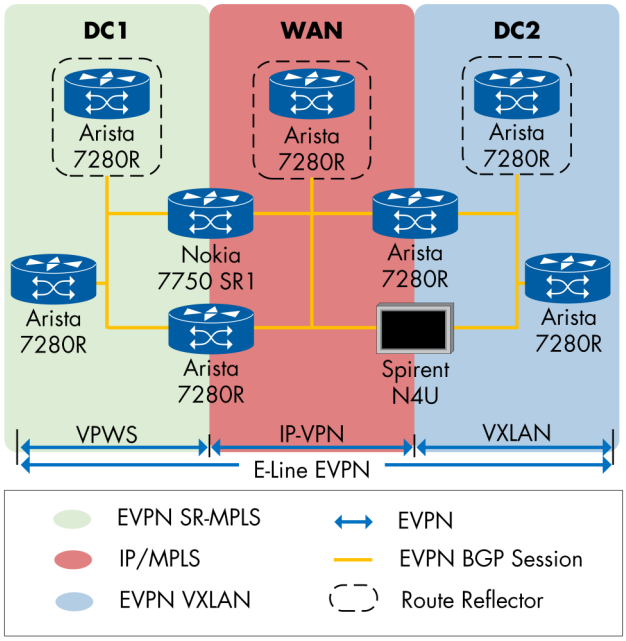


Figure 38: All-Active Multi-Homing Gateway PEs with Loop Prevention via D-PATH

EVPN VXLAN and VXLAN Interworking

We verified VXLAN and VXLAN interworking in this test. VXLAN is a widely supported data plane technology, which encapsulates a MAC frame in a UDP datagram for transport across an IP network. To be able to offer a regional or national EVPN network, service providers are seeking flexible approaches to extend the reach of EVPN beyond a single data center. One mechanism is the use of VXLAN in the Metro Area Network (MAN) to interconnect multiple EVPN domains.

We created four data center VXLAN segments consisting of AS1 to 4. We observed a total of three inter-domains EVPN services, each one served as an interconnection between every two data centers.

Num	Inter-Domain EVPN
1	AS1-AS2
2	AS2-AS3
3	AS3-AS4

Table 8: EVPN VXLAN and VXLAN Interconnection

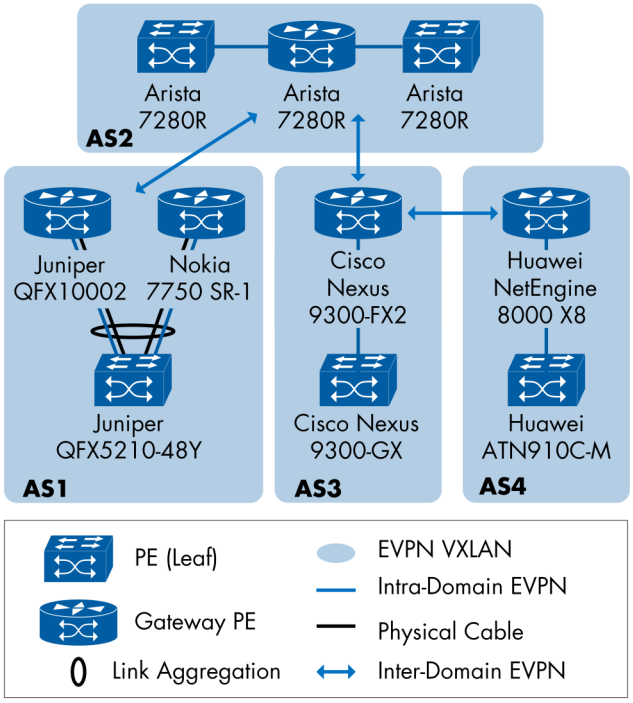


Figure 39: EVPN VXLAN and VXLAN Interworking

The following DUTs successfully participated in the setup, as

- Gateway PE: Arista 7280R, Cisco Nexus 9300-FX2, Huawei NetEngine 8000 X8, Juniper QFX10002, and Nokia 7750 SR1
- Leaf PE: Arista 7280R, Cisco Nexus 9300-GX, Huawei ATN910C-M, and Juniper QFX5210-48Y

The significance omitted was end-to-end EVPN. As end-to-end service requires interconnection through a AS, such as AS1-AS4 connection through AS2 or AS3. Thus we saw a bigger concept behind the design, which provides redundant interconnection of data centers if it succeeded. However, the configuration was not completed until the end of the test.

Seamless EVPN and VPLS

VPLS is a widely deployed l2VPN technology. Service providers who are looking at adopting EVPN want to pass the success of existing solutions to the new solution. EVPN provides backward compatibilities to VPLS PEs as defined in RFC8560. The solution must not require any changes to existing VPLS, not even a software upgrade. In order to support seamless integration with VPLS PEs, the RFC requires that VPLS PEs support VPLS A-D per [RFC6074], and it requires EVPN PEs to support both BGP EVPN routes per [RFC7432] and VPLS A-D per [RFC6074]. All the logic for seamless integration shall reside on the EVPN PEs. The EVPN PE establishes VPWS to VPLS PE. We verified end-to-end EVPN between VPLS and EVPN PEs.

All participating DUTs supported VPLS and established a VPLS service with each other. The DUTs with the role of EVPN PEs joined and fully discovered the VPLS PEs, then they established full-meshed VPWS devices with each other. We sent traffic through the VPLS. Once the services have been established, all traffic went through without any frame loss as expected.

The following DUTs successfully participated in the test as

- EVPN PE: Cisco ASR 9901, Huawei NetEngine 8000 X8, and Nokia 7750 SR1
- VPLS PE: Cisco ASR 9901, Huawei NetEngine 8000 X8, and Nokia 7750 SR1, Juniper MX480

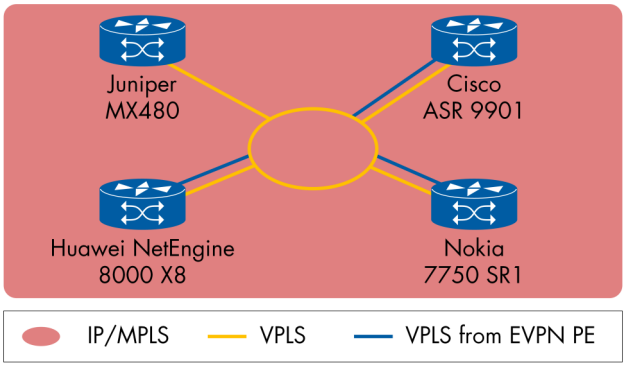


Figure 40: Seamless EVPN

Initially, we observed unexpected traffic loss for a few seconds while EVPN PEs joined. However, once the services have been established, all traffic went through without any packet loss. The latter result indicated that the established service ran as expected. We look forward to a solution in the configuration to exclude the impact from signaling, or to remove the signaling process from the SLA.

EVPN Fault Management

The test verifies that the OAM can be used to perform Connectivity Fault Management (CFM) with Continuity Check Messages (CCM) in an EVPN network.

The PEs established an EVPN service and sent CCM messages for the EVPN. We introduced a layer 2 failure with the impairment on the EVPN service and observed that both PEs discovered it with CCM messages. After that, we removed the failure from the link, we observed that CCM messages went through and reporting that service was up.

The following DUTs successfully participated as PEs with CCM messages:

- Juniper MX480 and Nokia 7750 SR1

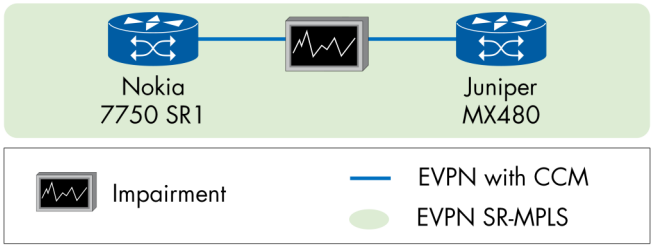


Figure 41: CCM for EVPN

SDN

Software-Defined Networks (SDN) is still the hot topic regarding separating the control plane from the data plane and centralizing the network control function. The wide use and implementations of SDN networks in different forms raise enormous challenges related to the best practices of implementations and integrating SDN with already in-use technologies.

Among all the challenges and difficulties facing the SDN implementations, we covered the multi-vendor interoperability aspect in our test event and this document. The SDN section doesn't include NETCONF/YANG tests this year as we separated the management northbound interface into another dedicated event.

We tested interesting test cases as colored SR policies and on-demand next hop. Nevertheless, we could not test PCEPv6 or SRv6 in this context as no vendor supported these features yet.

PCE-initiated Paths in a Stateful PCE Model

In environments where the LSP placement needs to change in response to application demands, it is helpful to support dynamic creation and tear down of LSPs. This ability of the PCE to trigger the creation of LSPs on demand can be used in SDN architectures.

A possible use case is where applications request network resources and paths from the network infrastructure. For example, an application can request a path with certain constraints between two network nodes by contacting the PCE. The PCE can compute a path satisfying the constraints and instruct the head-end network node (PCC) to instantiate and signal it. When the application no longer requires the path, the PCE can request a teardown for it.

The test was performed as follows:

- The DUTs started the IGP adjacencies between them, and the connectivity was verified.
  For this test, the DUTs established IS-IS as IGP.
- We verified the Stateful PCEP session.
- We verified PCE path instantiation.
- LSP state synchronization was verified.
- For this test we did not create VPN services to generate traffic, we used the pings to confirm transport paths were installed.

The following combinations successfully participated in the test, as

- PCE: Cisco Crosswork, Juniper Paragon Pathfinder + Paragon Insights, Nokia NSP, and Spirent N4U
- PCC: Cisco ASR 9901, Cisco NCS 540, Juniper MX204, and Nokia 7750 SR1
- P: Nokia 7750 SR1

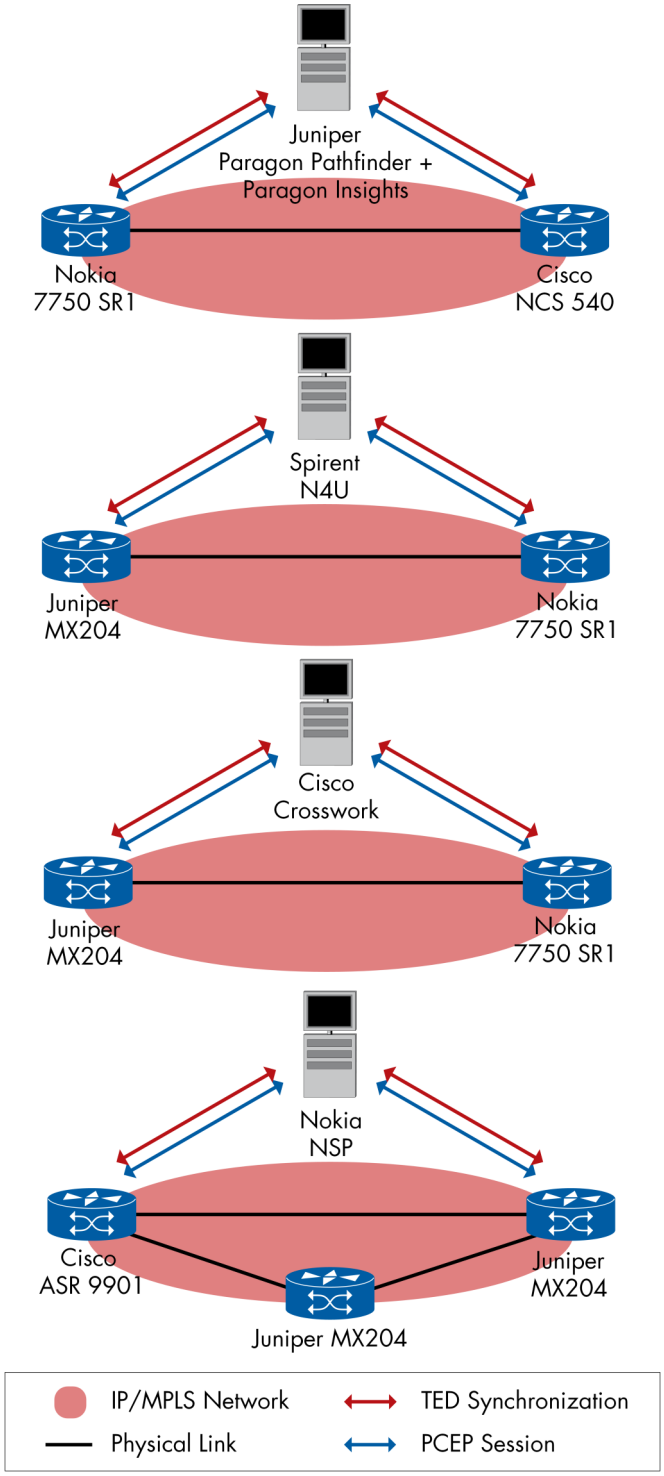


Figure 42: PCE-initiated Paths in a Stateful PCE Model Topology

PCC-initiated Paths in a Stateful PCE Model

The LSP state information allows the PCE to compute constrained paths while considering individual LSPs and their interactions. This requires a reliable state synchronization mechanism between the PCE and PCCs.

The updated PCEP mechanism allows PCE to modify the LSP attributes when the LSPs are delegated to the PCE.

We verified the creation, delegation, and revocation of LSP deletion of PCC-initiated LSP in this test. We also tested LSP re-optimization in the event of IGP cost is changed and wanted to test the re-delegation to the second PCE as an optional part.

This test could be performed for an MPLS network based on Segment Routing.

The test was performed as follows:

- The DUTs started the IGP adjacencies between them, and the connectivity was verified. For this test, the DUTs established IS-IS as IGP.
- We verified the Stateful PCEP session.
- We verified PCC path instantiation.
- LSP delegation and update.
- L3VPN was initiated between the PCCs.
- The paths were verified by generating bidirectional traffic using Spirent.

The following DUTs successfully participated in the test, as

- PCE: Cisco Crosswork, Nokia NSP, and Spirent N4U
- PCC: Cisco ASR 9901, Juniper MX204, and Nokia 7750 SR1

For one combination, the PCE did not have the ability to use general configurations for both PCCs, so we had to do the test sequentially between the PCCs.

The same PCE could not perform all the test steps, as it could not have the IGP knowledge.

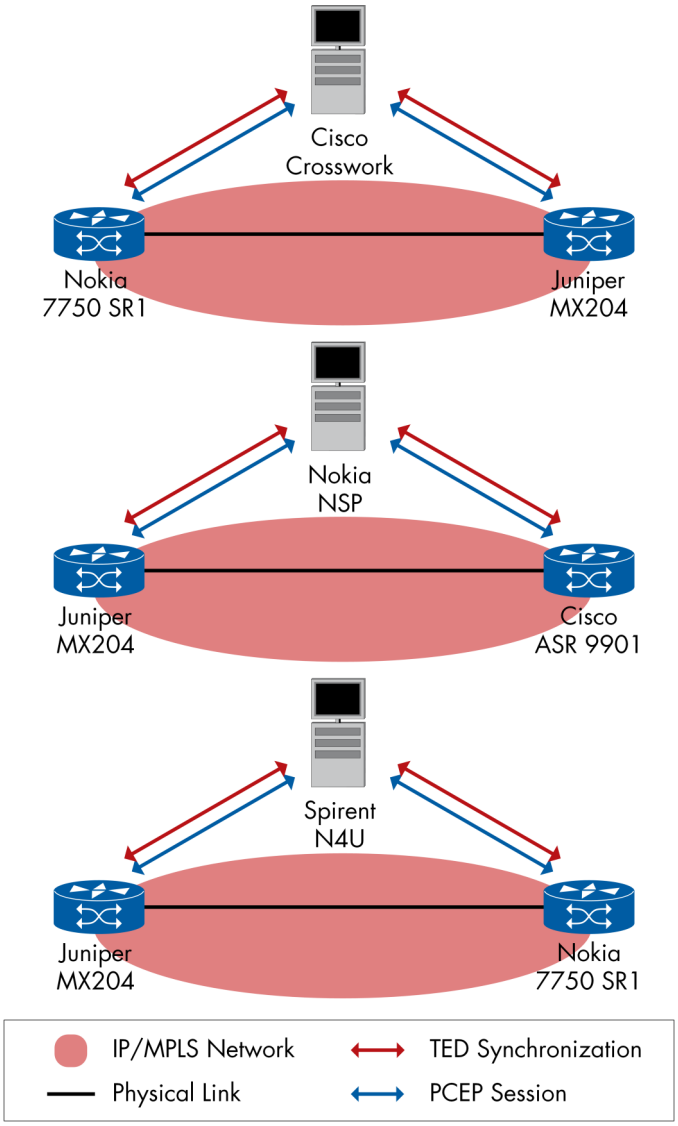


Figure 43: PCC-initiated Paths in a Stateful PCE Model Topology

PCE Path Computation with SR-MPLS

In this test, we verified the setup of end-to-end service using a standard service control plane. At the same time, the transport is derived using segment routing without utilizing hop-by-hop signaling technics (LDP or RSVP-TE). The SR path is derived from a PCE controller. The PCE controller knows network topology via Traffic Engineering Database (TED) and previously established paths via the LSP database.

During the test, the PCE pushed colored SR Policy via PCEP to the PCCs.

The test was done using SR-MPLS only, as there was no support yet for SRv6 or PCEPv6.

The test was performed as follows:

- The DUTs started the IGP adjacencies between them, and the connectivity was verified.
- We verified the Stateful PCEP session.
- Colored SR Policy via PCEP was pushed by the PCE.
- L3VPN was initiated between the PCCs.
- The paths were verified by generating bidirectional traffic using Spirent.

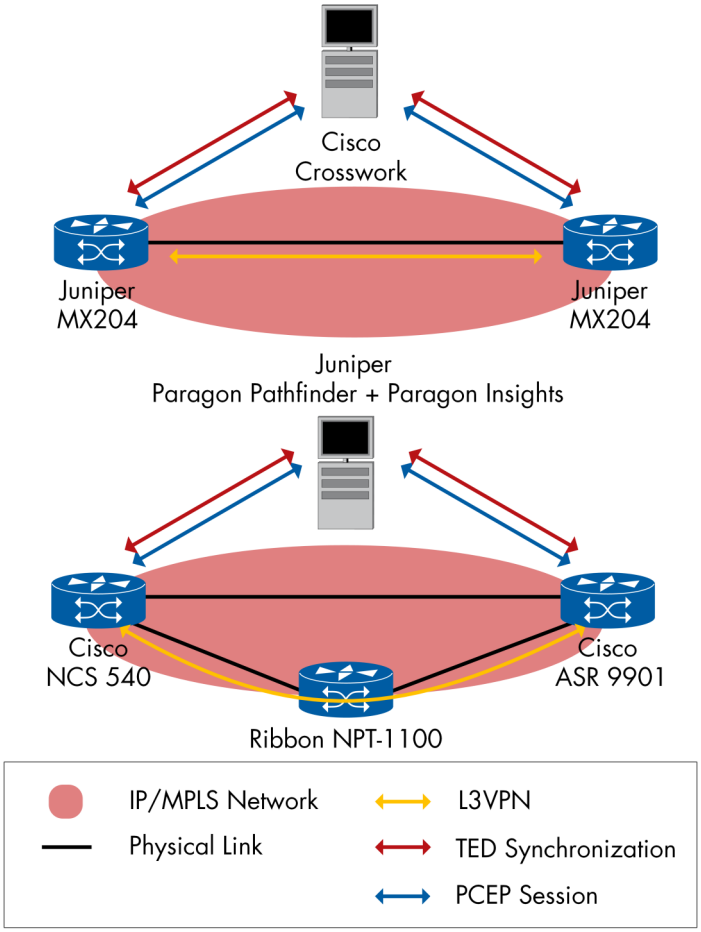


Figure 44: PCE Path Computation with SR-MPLS

The following DUTs successfully participated in the test, as

- PCE: Cisco Crosswork and Juniper Paragon Pathfinder + Paragon Insights
- PCC: Cisco ASR 9901, Cisco NCS 540, and Juniper MX204
- P: Ribbon NPT-1100

PCE Managing SR Policies via BGP SR-TE NLRI

BGP possesses the capability to provide SR policy, and it can also give a candidate path. Segment Routing with BPG SR-TE discusses an alternative between SDN controller and headend for path instantiation. The SDN controller provides computation of the path. The headend receives BGP the calculated path. In such a scenario, none of PCEP is required.

This test verified that PCE is managing SR policies with BGP SR-TE.

The test was performed as follows:

- BGP Sessions for Address Family SR-Policy were checked.
- VPN Routes for VRF configured on DUTs were checked.
- Started bidirectional traffic on Spirent through the configured VRF.
- Verified BGP next-hop resolution and SR Policies for VRF Prefixes.
- Triggered creation of SR Policies on PCE and advertise it via BGP to PCCs.

The following DUTs successfully participated in the test:

- PCE: Nokia NSP
- PCC: Cisco ASR 9901, Juniper MX204
- P: Cisco 8201

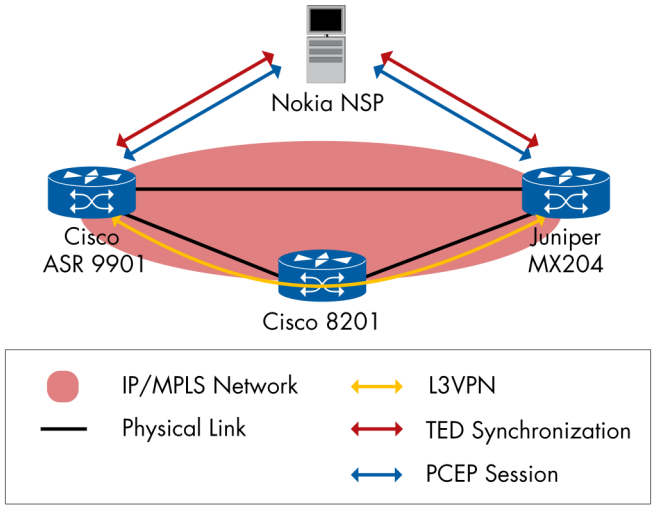


Figure 45: PCE Managing SR Policies via BGP SR-TE Topology

Dynamic Instantiation of SR Policy

Steering VPN service traffic over the SDN network introduces a new concept of automatic steering, which is essential for mapping packets for required network paths to meet SLA requirements. SR provides TE capabilities by using SR-TE policies associated with packets of a VPN service. The instantiation of SR policy can be dynamic as key to auto-steering. PCE implements a centralized control, PCC (ingress PE) identifies packets matching SR-TE policy for triggering of instantiation.

SR with PCE provides computation for instantiating any SRv6 policy that is associated with a VPN service. At the ingress PE (PCC) of the network, upon receipt of a VPN route (BGP routes) from CE or remote PE matching the SR-TE policy identifier, the PCC sends a request to PCE for triggering the instantiation of the SR policy.

In this test, we verified the instantiation of SR-TE policies upon receipt of VPN routes at the headend. A variety set of SR-TE policies were under test, such as IGP-TE metric as well as delay. We also tested path updates once the previous path to an SR-TE policy is no longer valid. Finally, we verified the deletion of a path.

For BGP routes mapped to an SR-TE policy, this test was based on the color community in BGP routes.

The following DUTs successfully participated in the test, as

- PE: Cisco ASR 9901, Juniper MX204, and Nokia 7750 SR1
- P: Cisco 8201

One device of the DUTs could not perform color-based prefix-steering by BFP so it was performed using locally configured SR policies.

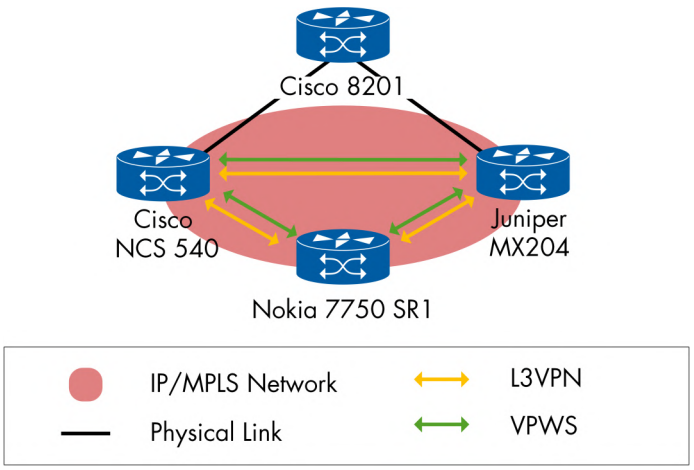


Figure 46: Dynamic Instantiation of SR Policy Topology

Flexible Ethernet

The idea behind FlexE technology is to channelize Ethernet and provide flexible granular bandwidth to optimize the use of fiber capacity. Initially, FlexE is a protocol published by the Optical Internetworking Forum (OIF). Furthermore, FlexE is now being adopted by ITU-T recommendation for Metro Transport Networks, G.mtn, as an integral part of ITU-T Study Group 15 (SG15) scheduled for 2020. Channelized techniques such as channel isolation, bonding, and sub-rating were all key parts of this test. We created SR-MPLS overlay over FlexE in a multi-vendor environment.

FlexE Channelization and Physical Isolation

The main purpose of this test is to perform FlexE channelization and physical isolation of 100G ports in a back-to-back FlexE scenario. In addition, we also verified the sub-rating functionality, through a full set of channels based on the FlexE interfaces with asymmetric bandwidth.

Sub-rating is an action that allows network elements to sub-divide physical interfaces in order to transport lower data pipes over partially filled Ethernet PHYs.

We created an SR-MPLS overlay network over the FlexE channelization. The setup includes two fully channelized 100GbE interfaces. Based on the 5 Gbit/s calendar slot of FlexE, building uniform channels at the same speed of 20 Gbit/s, to obtain visual balance, resulting in five equal channels.

Another interface maintained the same channel number to five, where sub-rating shall take place. The common channel speed at this interface is lowered down to 10 Gbit/s, resulting in four equal channels, and the remaining bandwidth allocated the last channel of 60 Gbit/s. The interface channel settings were asymmetric.

We expected traffic isolation between the channels. None of the traffic from one channel shall impact traffic in another channel. In direction of the 10 Gbit/s channel to the 20 Gbit/s channel, so-called sub-rating occurs, at where the rate on the other side became smaller, traffic shall not show any loss. In contrast, from the other direction (20 Gbit/s channel to 10 Gbit/s channel), when a flow is within 10G, there should be no loss. If it is exceeded, packet loss is required. During this process, none of the traffic from other channels shall be affected.

As expected, we did not observe any packet loss in all channels. The sent packets were all received at the same channel.

The traffic at a rate of 20 Gbit/s was received in 60 Gbit/s - 20 Gbit/s channel without any loss. The traffic exceeded the 20 Gbit/s was dropped as expected. The following DUTs successfully participated as PE over FlexE:

- Ciena 5164, Huawei NetEngine 8000 X8, and Spirent N4U

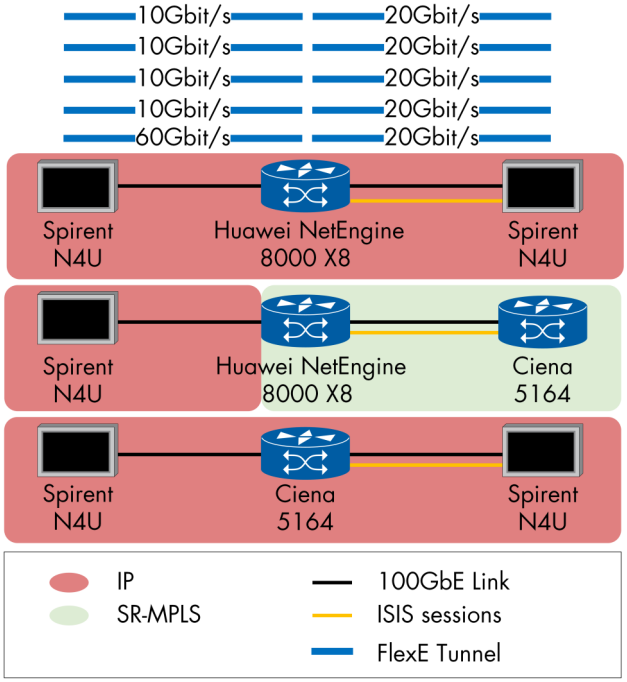


Figure 47: FlexE Channelization and Physical Isolation

Baseline Test

Channel (Gbit/s)	Frames Sent (fps)	Frames Received (fps)
10 - 20	1,000	1,000
10 - 20	2,000	2,000
10 - 20	3,000	3,000
10 - 20	4,000	4,000
60 - 20	5,000	5,000

Table 9: Baseline Test

Sub-Rating Test

Channel (Gbit/s)	Bitrate Sent (Gbit/s)	Bitrate Received (Gbit/s)
60 - 20	60	20

Table 10: Sub-Rating Test

FlexE Bonding

This test verified the FlexE bonding on 100G ports in a back-to-back FlexE scenario. We also verified the allocation of resources from two links to one of the individual channels.

Both bundled FlexE interfaces (2 x 100GbE) forwarded a full rate of traffic at 200 Gbit/s without any loss. We observed one selected channel at 120 Gbit/s over two links as expected.

The following DUTs successfully participated in the tests:

- Ciena 5164, Huawei NetEngine 8000 X8, and Spirent N4U

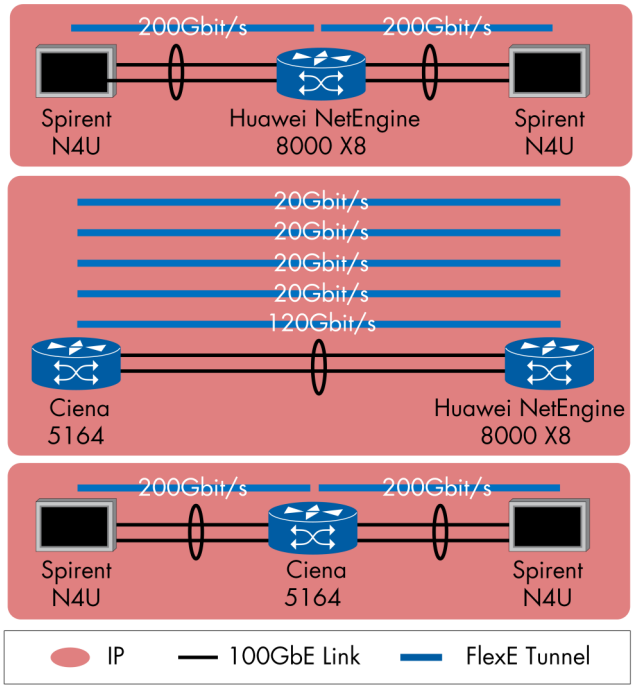


Figure 48: FlexE Bonding

FlexE Tunnel	Traffic Sent (Gbit/s)	Traffic Received (Gbit/s)
All channels	200	200

Table 11: Bonding Traffic

FlexE Dynamic Bandwidth Adjustment

The purpose of this test was to verify the FlexE capability of dynamic bandwidth adjustment on 100G ports in a back-to-back FlexE scenario.

We reused the same setup from the baseline test and selected the last two channels. We sent traffic for all selected channels. While traffic was running, we modified via CLI the bandwidth settings for the last channel from 20 Gbit/s to 30 Gbit/s. During this process, we expect that none of the traffic from the existing channel shall be affected. Once the configuration was applied, we did not expect any traffic loss in the channel with modified bandwidth at 30 Gbit/s.

All setups showed the expected behavior as described.

The following DUTs successfully participated in this test:

- Ciena 5164, Huawei NetEngine 8000 X8, and Spirent N4U

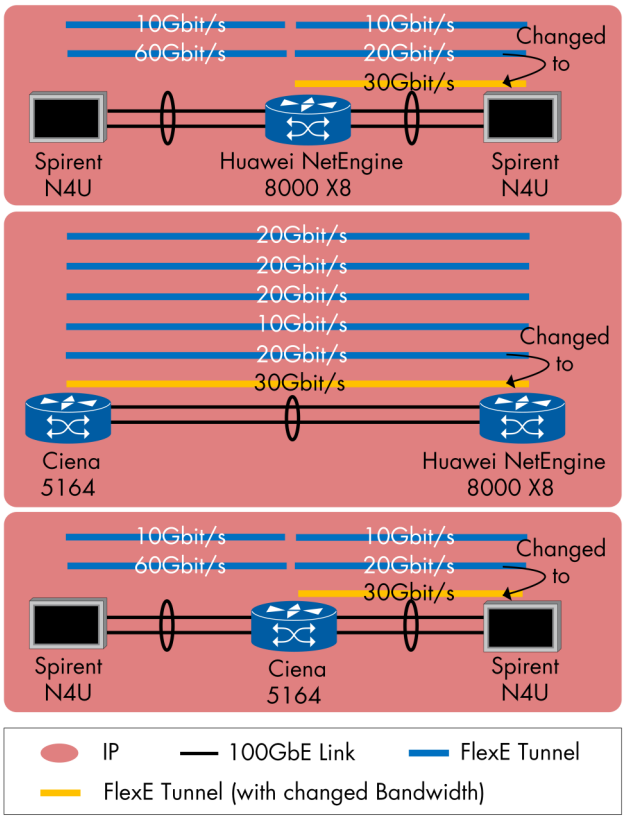


Figure 49: FlexE Dynamic Bandwidth Adjustment

Clock Synchronization

The industry is going towards 5G and its standards, and as such the synchronization requirements are becoming tighter and harder to achieve, which points out the need for very accurate, modern, intelligent and reliable test scenarios. This is the role of EANTC's MPLS SDN interoperability tests to provide highly reliable, accurate, and independent testing.

With the growing importance of time synchronization in the industry, and the crucial role it plays in modern technologies like 5G, Clocking tests have become very essential, not only regarding time accuracy and error but also failover scenarios and security.

This year's event presented tests covering 5G synchronization requirements, ITU-T performance requirements e.g., Boundary Clock Class C/D tests, resiliency scenarios, and PTP implementations. We tested the behavior of the time signal delivery in optimal and suboptimal conditions: Hold-over performance, source failover between two Grandmaster (GM) Clocks with high precision Clocking. As a result, a total of 29 test combinations were successfully completed. A range of interface rates was used during these test combinations covering 1GbE, 10GbE, 25GbE, and 100GbE.

For those tests involving a network of devices, we defined a minimum acceptable accuracy level of ± 260ns (ITU-T recommendation G.8271 accuracy level 6A). However, where individual devices were tested in isolation, the G.8273.2 Telecom-Boundary Clock (T-BC)/ Telecom-Time Slave Clock (T-TSC) Class C and D limits were applied. In other cases, we defined the accuracy level of ±1.5µs (ITU-T recommendation G.8271 accuracy level 4) as our end-application goal, with 0.4µs as the phase budget for the air interface. Therefore, the requirement on the network limit, the last step before the end application, had to be ±1.1µs.

EANTC used the Calnex Paragon suite of products for both measurement and impairment scenarios. The Paragon-X was used to generate the network impairments and perform captures that were analyzed with the Calnex Analysis Tool (CAT) and PFV tools to confirm T-BC/T-TSC behavior.

The Calnex Paragon-T was used to provide multiple measurements and was very valuable in having four possible ports to take the measurements at the same time which helped us perform multiple tests in parallel. In addition, the Calnex Paragon-Neo was able to act simultaneously as both an emulated PTP Master and Slave to perform accurate time error measurements of T-BC and T-TSC devices to Class D levels of accuracy.

The CAT was our analysis and reports generation tool, providing all that we needed to apply masks or calculate the time error in all its forms (maximum absolute time error, or constant time error,...) as well as reporting against the 5G network limits and Clock mask requirements.

The Primary Reference Time Clock (PRTC) was GPS using an L1 antenna located on the roof of our lab. The synchronization test team tested brand new software versions, products, and interface types, including PTP over 100GbE.

Our tests helped to discover several small issues, but the R&D departments of the vendors reacted quickly providing patches and troubleshooting support.

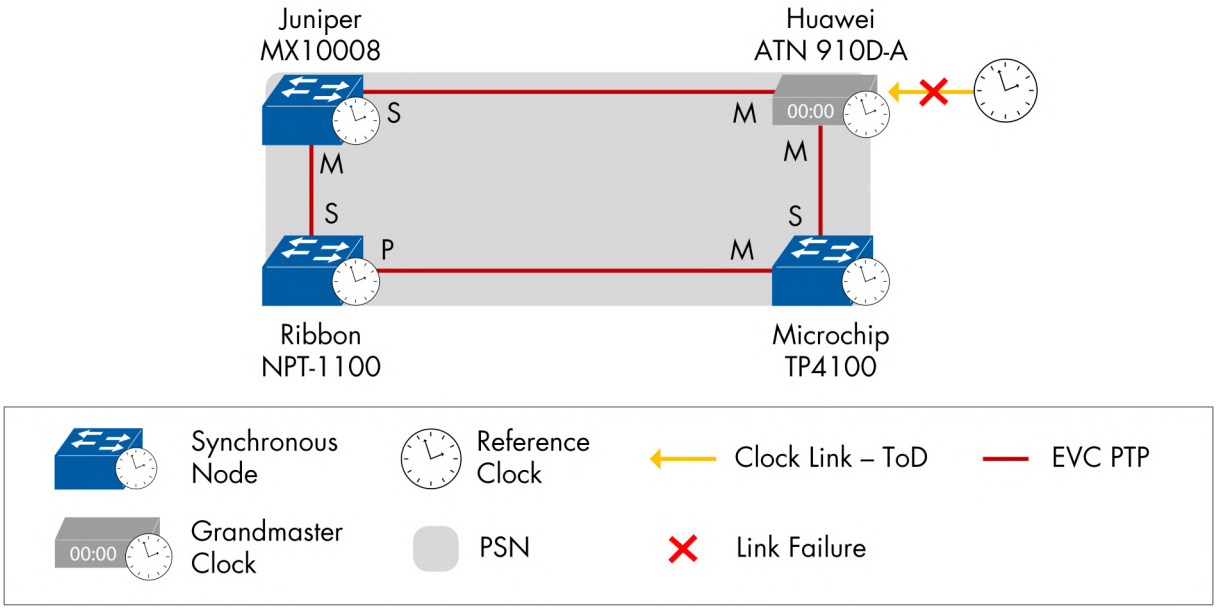


Figure 50: Phase/Time Passive Port Monitoring

Phase/Time Passive Monitoring — Relative Time Error Monitoring Option as per G.8275.1 Annex G

This test case aims to verify the relative time error monitoring option as per G.8275.1 Annex G testing the correct usage of the alternateMasterFlag usage and potential accuracy of observed TE difference.

The test configuration was set up and shown with one Grandmaster and three BCs, only one of which performed passive monitoring, each configured to use the G.8275.1 Telecom Profile and SyncE in hybrid mode.

As per the standard, the Calnex Paragon-X instrument was used to capture the PTP messaging from the passive port and used its integrated PFV tool verify the contents of the PTP messaging from the Passive port to ensure that the alternateMasterFlag was set to FALSE as required.

The Paragon-X then added a 250µs asymmetry into the timing path in one direction to add asymmetry into the timing path. Logs were checked to ensure that this was detected.

The following DUTs successfully participated in the test, as

- T-BC: Juniper MX10008, Microchip TP4100, and Ribbon NPT-1100
- GM: Huawei ATN 910D-A

The test was performed as expected, with the correct setting for the flag being detected at the Paragon-X and the asymmetry being flagged, and the alarm raised as required at the vendors BC.

Phase/Time Holdover with Enhanced Sync-E Support in the Core — Measuring Holdover Performance using Enhanced-Sync-E Frequency Lock

The test was performed using the G.8275.1 Telecom Profile with eSyncE in hybrid mode to illustrate the use of the QL levels defined for Enhanced SyncE.

Using the configuration shown, the Slave Clock was connected to a Boundary Clock, which was GNSS locked for time reference and receiving PTP from a GNSS locked Grandmaster. The Slave Clock in turn was PTP and eSync-E locked to the Boundary Clock.

At each test stage, the time error output was measured using a Calnex Paragon-T measurement analyzer.

During the test, the GNSS reference to the Boundary Clock was disabled, causing the Boundary Clock to take its frequency reference from the eSync-E provided by the Grandmaster. At this time, the BC remained PTP locked to the Grandmaster and received PRTC packets. After the transition due to the frequency reference change, the Slave, we again analyzed the time error performance at the Slave output.

Finally, the PTP output from the Boundary Clock was disabled, causing the Slave Clock to go into holdover but maintaining its frequency lock to eSync-E. Again we measured time error at its output to determine any performance loss due to holdover.

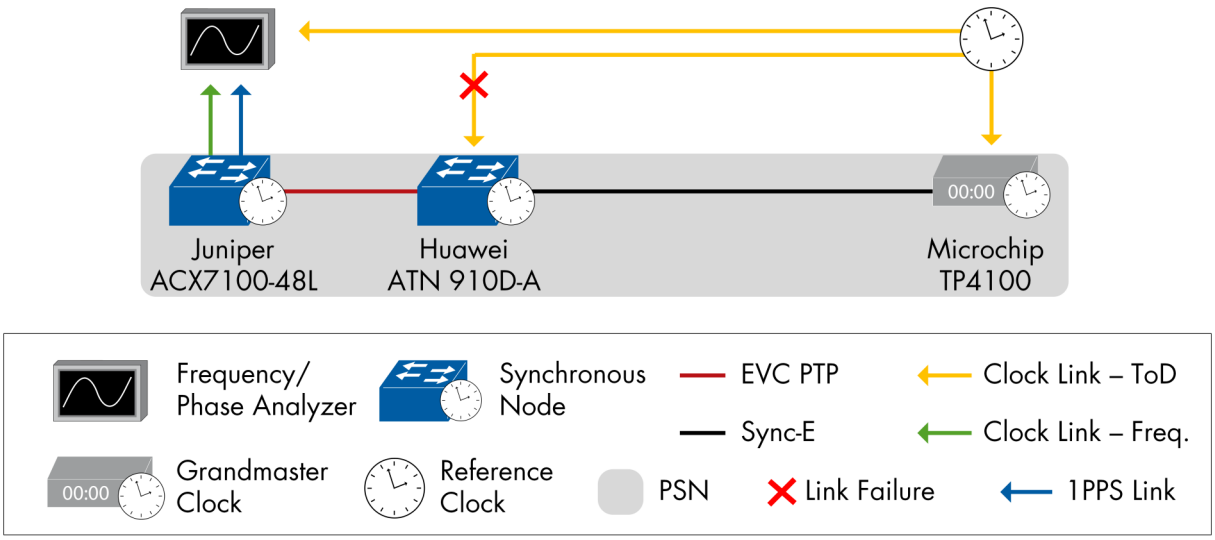


Figure 51: Phase/Time Holdover with Enhanced Sync-E Support in the Core

The following DUTs successfully participated in the test, as

- BC: Huawei ATN 910D-A
- SC: Juniper ACX7100-48L
- GM: Microchip TP4100

We did not find real issues. Logs were captured from Boundary Clock and Grandmaster, showing the use of Enhanced Sync-E status. The holdover performance passed Class 6B.

Phase Synchronization when PTP Carried over FlexE Transport –

Maintaining Phase Performance when using FlexE

To ensure that the impact of using FlexE as a transport container does not impact the inherent timing performance of a Class 6 C/D Boundary Clock.

The team used the Calnex Paragon-Neo test and measurement instrument in this test to emulate a PTP Master and Slave and to accurately measure the time error output of the Boundary Clock to ITU-T G.8273.2 Standard Class C/D limits.

The test used the G.8275.1 ITU-T Telecom Profile carried over 100GbE links using a FlexE transport.

The Boundary Clock was connected to the Paragon-Neo Master and Slave and configured to acquire frequency and PTP. PTP was started on the Paragon-Neo Master and Slave with SyncE (QL-PRC) generated at the Paragon-Neo Master.

Once we attained the lock at the BC, a time error measurement was performed on the Paragon-Neo for 1000s. The resulting capture was then analyzed using the Paragon-Neo analysis tool to examine the Time Error output of the Boundary Clock.

The 2-way time error value was subjected to the G.8273.2 T-BC/T-TSC limits for a Class D Clock and the dynamic TE MTIE LF was measured against the G.8273.2 T-BC dynamic MTIE mask.

NOTE: The dynamic MTIE mask is provisional at this stage for Class D Clocks and was used here as an indicate of performance only.

The following DUTs successfully participated in the test, as

- BC: Ciena 5164
- SC: Calnex Paragon-Neo
- GM: Calnex Paragon-Neo

There were no significant issues seen during the test. Configuration, as always, took some time to resolve for compatible optics and FlexE settings, but once sorted, the test proceeded as expected.

The Boundary Clock passed both the Class 6D limit and the MTIE mask.

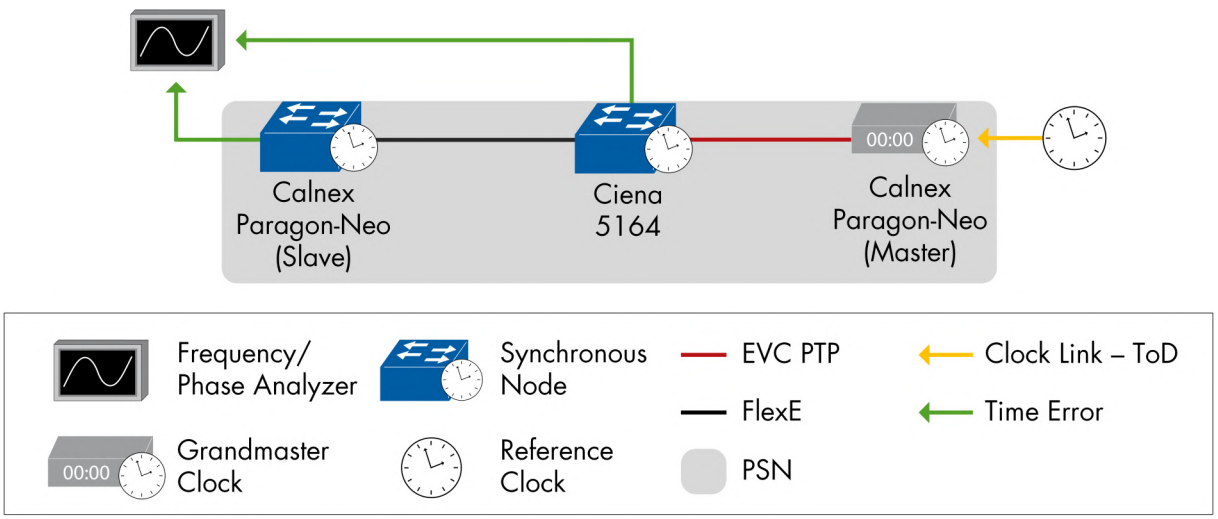


Figure 52: Phase/Time Synchronization over FlexE

Conformance Test Class C/D Boundary Clock –

Measuring to G.8273.2 Standards

The migration to 5G has imposed ever-stricter limits on synchronization timing in networks. The ITU-T G.8273.2 T-BC/T-TSC Timing Characteristics standard has introduced tighter limits for devices operating in this environment. This test is designed to determine whether or not vendors' T-BC or T-TSC devices conform to these limits. It should be noted that while the use of Class C/D Clock devices is targeted at Fronthaul networks, the Backhaul networks will face the existing budgets of +/- 1.5µs.

We used the Calnex Paragon-Neo test and measurement instrument in this test to emulate a PTP Master and Slave and to accurately measure the time error output of the Boundary Clock to ITU-T G.8273.2 T-BC limits for Class C/D Clocks. The test used the G.8275.1 ITU-T Telecom Profile carried over 1GbE, 10GbE, or 100GbE links while simultaneously sending Sync-E in hybrid mode.

The Boundary Clock was connected to the Paragon-Neo Master and Slave and configured to acquire frequency and PTP. PTP was started on the Paragon-Neo Master and Slave with SyncE (QL-PRC) being generated at the Paragon-Neo Master.

Once we attained the lock at the BC, a time error measurement was performed on the Paragon-Neo for 1000s. The resulting capture was then analyzed using the Paragon-Neo analysis tool to examine the time error output of the Boundary Clock.

The 2way time error value was subjected to the G.8273.2 T-BC Clock Class D limits and the dynamic TE MTIE LF was measured against the G.8273.2 T-BC dynamic MTIE mask.

Note: The dynamic MTIE mask is provisional at this stage and as such was used here only for informational purposes.

The following DUTs successfully participated in the test, as

- T-BC: Ciena 5164, Juniper ACX710, Juniper ACX7100-48L, Microchip TP4100, Nokia IXR-e, and Ribbon NPT-1100
- SC: Calnex Paragon-Neo
- GM: Calnex Paragon-Neo

DUT	GbE	Clock Class
Ciena 5164	100	D
Juniper ACX710	100	D
Juniper ACX7100-48L	10	D
Microchip TP4100	1	D
	10	C
Nokia IXR-e	10	D
	100	D
Ribbon NPT-1100	1	D
	10	D
	25	C
	100	D

Table 12: DUT, GbE, and Clock Class

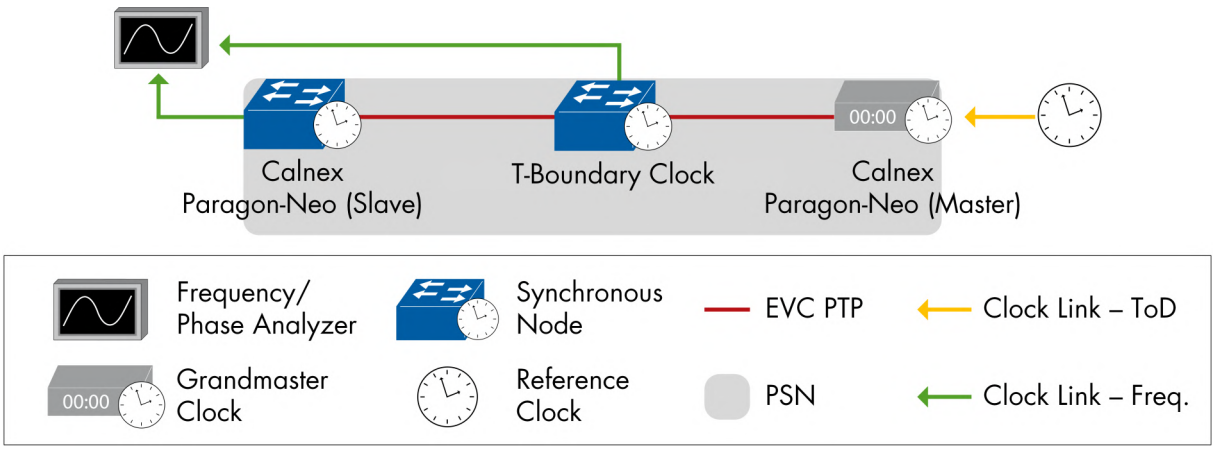


Figure 53: Conformance Test Boundary Clock Class C/D 1GbE, 10GbE, 25GbE, 100GbE

High-Precision Clocking Source Failover

Verify that a Slave Clock maintains the required Clock synchronization frequency, phase, and time quality when it switches from its primary to its secondary Grandmaster following a signal degradation. Verify the level of accuracy can reach ITU-T G.8271 Level 6 precision.

The Slave and Boundary Clocks may be provided with a primary and secondary Grandmaster for resiliency. In this test, both Grandmasters are provided with a GPS signal. All devices are configured to use the G.8275.1 Telecom Profile with SyncE frequency reference in hybrid mode.

Time of day measurements were taken at each failover using the Calnex Paragon-T measurement analyzer after achieving the stable lock.

We configured the Slave Clock to favor the primary Grandmaster. When we started the test, it locked to that Grandmaster. The GPS signal was then disconnected from the primary Grandmaster, and the Slave Clock locked to the secondary Grandmaster.

After the team achieved stability, we disconnected the GPS connection to the secondary Grandmaster. The Slave Clock then locked back to the primary Grandmaster. The GPS signal was then reinstated to the secondary Grandmaster, causing the Slave Clock to reacquire its lock to this Grandmaster.

Finally, the GPS signal was reinstated to the primary Grandmaster, resulting in the Slave Clock reacquiring lock to the primary again.

During the transition from one Grandmaster to the other, the Paragon-T was used to capture the transition period, and the time error output was measured against MTIE G.813 Short Term transient mask. When we acquired the lock, we measured the time error output of the Slave against the Class 6 limits and MTIE against the G.823 SEC Mask.

The following DUTs successfully participated in the test, as

- BC: Ciena 5164, Juniper ACX710, Juniper ACX7100-48L, Juniper MX10008, Ribbon NPT-1100
- GM: Calnex Paragon-Neo, Microchip TP4100, and Huawei ATN 910D-A

The biggest issue in this test was the cable lengths used for both the 1pps output and the GPS antenna to the Grandmaster GNSS ports. The 1pps cable lengths are significant when testing to Class 6 as there is little margin for error at these limits, so knowing the exact cable length is important.

This is also true when it comes to switching between Grandmasters. Ideally, all Grandmasters should be connected to the GNSS antenna with the same cable length. However, at one point, we discovered a 20m difference that accounted for 100ns of error when a switchover occurred. Once this was compensated for in the Grandmaster, the results fell within the Class 6 limits.

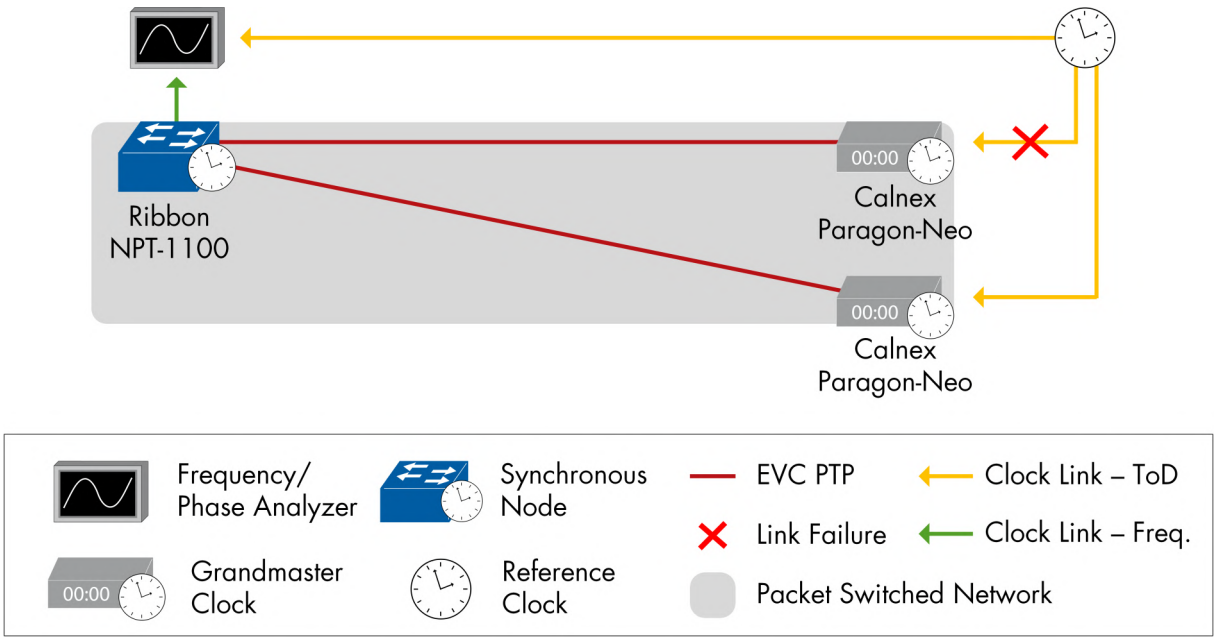


Figure 54: High-Precision Clocking Source Failover 1GbE

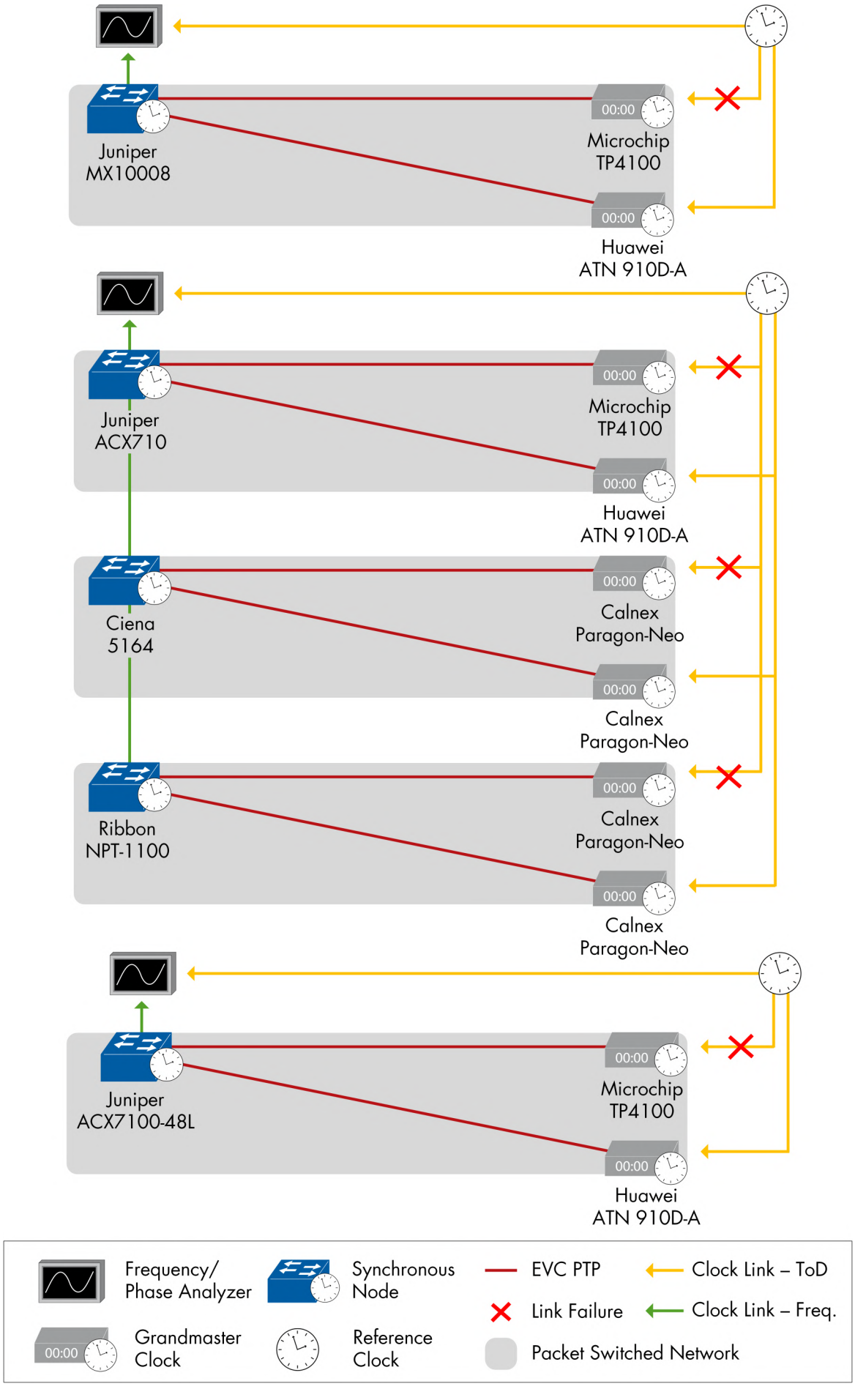


Figure 55: High-Precision Clocking Source Failover 10GbE

Phase/Time Synchronization Source Failover

Verify that a Slave Clock maintains the required Clock synchronization phase/time quality when the Boundary Clock switches from its primary to its secondary Grandmaster following a signal degradation.

The Slave and Boundary Clocks may be provided with a primary and secondary Grandmaster for resiliency. In this test, both Grandmasters are provided with a GPS signal. All devices are configured to use the G.8275.1 Telecom Profile.

Time of day measurements were taken at each failover using the Calnex Paragon-T measurement analyzer after achieving the stable lock. We configured the Boundary Clock to favor the primary Grandmaster. When we started the test, it locked to that Grandmaster. The GPS signal was then disconnected from the primary Grandmaster, and the Boundary Clock locked to the secondary Grandmaster. After we achieved stability, we disconnected the GPS connection to the secondary Grandmaster. The Boundary Clock then locked back to the primary Grandmaster. The GPS signal was then reinstated to the secondary Grandmaster, causing the Boundary to reacquire its lock to this Grandmaster.

Finally, the GPS signal was reinstated to the primary Grandmaster, resulting in the Boundary Clock reacquiring lock to the primary again.

After we acquired the lock at each stage, we measured the time error output of the Slave against the Class 6 limits and MTIE against the G.823 SEC Mask.

The following DUTs successfully participated in the test, as

- BC: Ribbon NPT-1100
- SC: Huawei ATN 910D-A
- GM: Microchip TP4100

There were no issues with this test. At some points, the Slave device passed the Class 6C limits as others Class 6B.

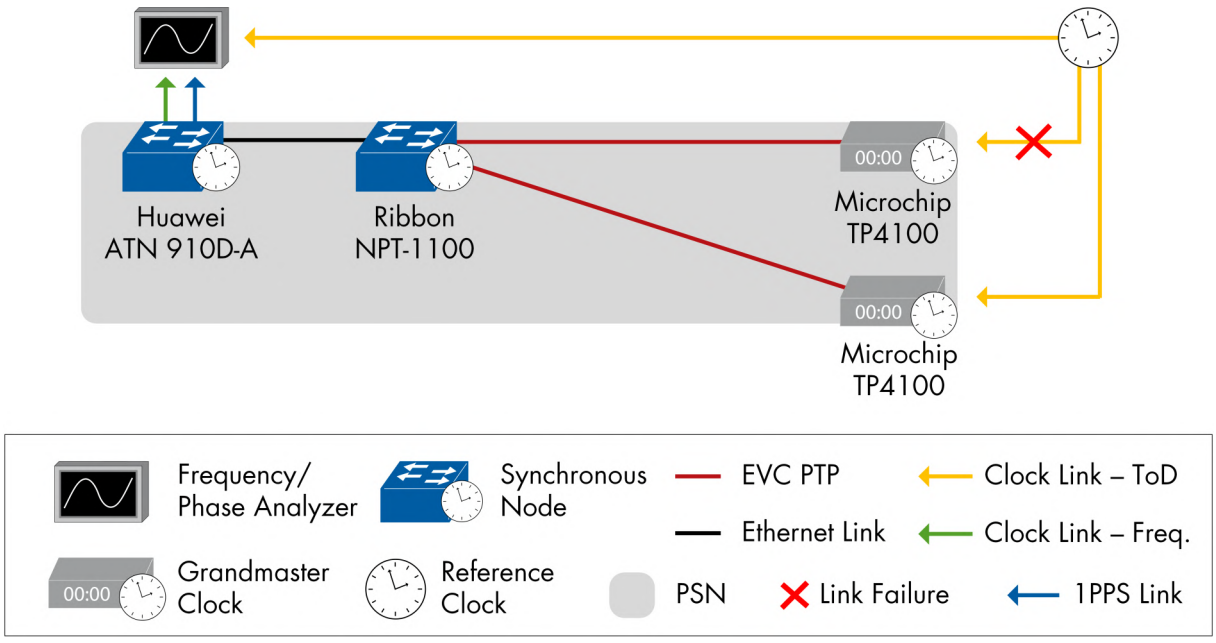


Figure 56: Phase/Time Synchronization Source Failover

Phase/Time Synchronization: Loss of Primary Source —

Measuring the Effect of Source Failover to Secondary Source

The purpose of this test is to ensure that a Boundary Clock can maintain its phase/time synchronization when it loses its GPS connection and switches to use another GPS-led source of PTP. In this test, the Boundary Clock was locked to GPS as its primary source and received PTP from another Boundary Clock connected to a GPS connected Grandmaster.

The test was performed using the G.8275.1 Telecom Profile, all devices are configured to use it with SyncE frequency reference in hybrid mode.

During the execution of the test, the performance when using its primary reference was recorded and measured against G.8271 Accuracy Level 4 limits.

We then disconnected that primary source such that Boundary Clock switched to use its secondary source, that of the PTP flow from the other Boundary Clock. Performance during this transition whilst acquiring lock and performance once locked were both measured and compared against the limits defined in the G.8271 ITU-T standard using the Calnex Paragon-T measurement analyzer.

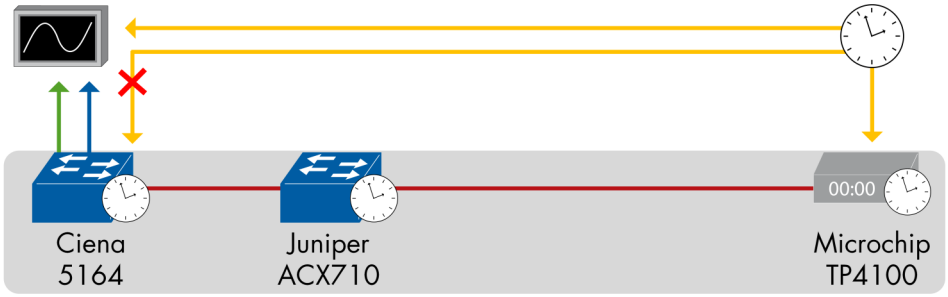


Figure 57: Phase/Time Synchronization Degradation of Primary Source 1GbE

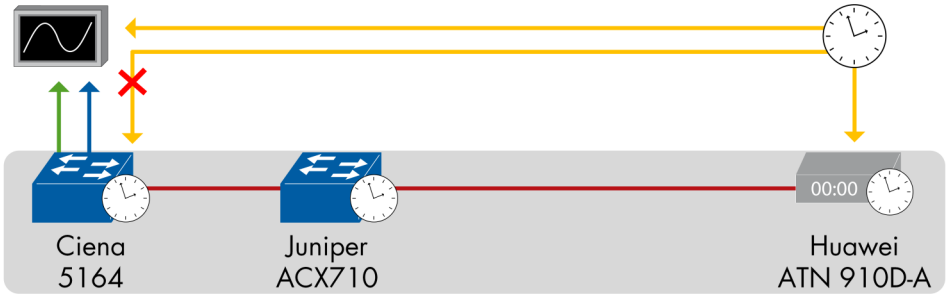
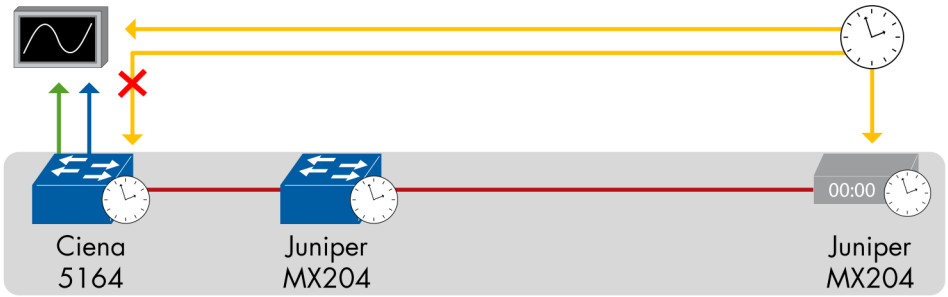


Figure 58: Phase/Time Synchronization Degradation of Primary Source 10GbE



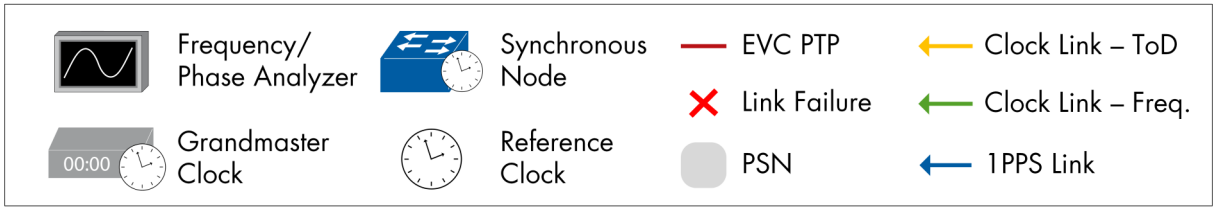


Figure 59: Phase/Time Synchronization Degradation of Primary Source 100GbE

The following DUTs successfully participated in the test, as

- BC: Ciena 5164, Juniper ACX710, Juniper MX204
- GM: Huawei ATN 910D-A, Juniper MX204, and Microchip TP4100

All devices in the test comfortably passed the 1.1µs limit defined in the G.8271 standard.

Measuring Phase/Time Accuracy of 5G Devices using MACsec

To ensure that 5G PTP devices that are using MACsec security protocols can maintain the timing performance when deployed in 5G networks.

Using the G.8275.1 profile on a 1GbE physical connection, the Boundary Clock was connected to a Paragon-Neo Master port, and the Slave Clock was connected to the Paragon-Neo Slave port.

The 1pps output was connected to the 1pps measurement port of the Paragon-Neo.

The test was started with the Boundary Clock waiting on PTP and SyncE lock from the Paragon-Neo Master.

Initially, MACsec was disabled between the Boundary Clock and the Slave Clock.

Both PTP and SyncE (QL-PRC) were started on the Paragon-Neo Master, and when PTP and SyncE lock were attained at the Boundary Clock a Slave Clock a measurement was performed for 1000s.

The 1pps output was measured using the CAT with the TE measured against the G.8271 Level 4 limit of +/- 1.5µs, and the dynamic TE MTIE LF measured against then G.823 SEC Wander Limit.

MACsec was then switched on and we repeated the test. Testing was carried out using interface rates of 1GbE and 10GbE.

The following DUTs successfully participated in the test, as

- BC: Ribbon NPT-1100
- SC: Ribbon NPT-1100
- GM: Calnex Paragon-Neo

The Slave Clock passed the TE limits and MTIE masks in both cases when MACsec was not enabled and when it was enabled. There was, therefore, no significant impact when using MACsec.

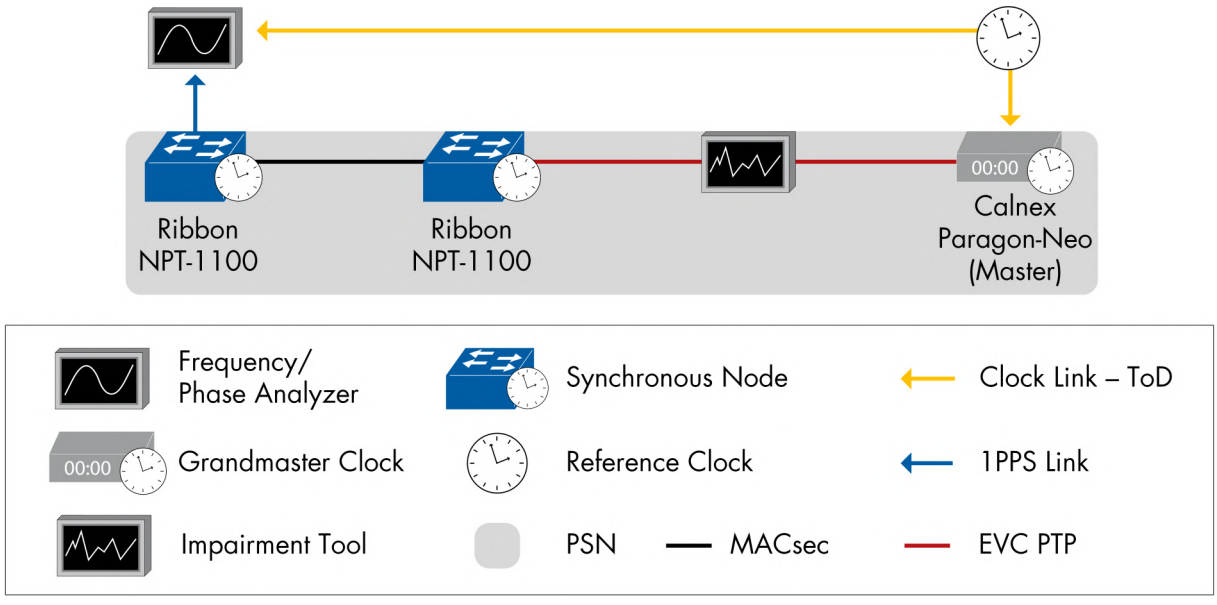


Figure 60: Phase/Time Partial Timing Support over MACsec

Phase/Time Full Timing Support: Boundary Clock Class C/D Test — Measuring T-TSC to Class 6 in a Network Chain

With ever-tighter limits imposed on timing synchronization within modern networks, it is important to ensure that equipment used in these networks conforms to the performance limits defined for those networks.

This test aims to verify that a Slave Clock can maintain its synchronization quality when using the G.8275.1 Telecom profile and SyncE in hybrid mode with a chain of Class D Boundary Clocks that themselves conform to the performance limits defined for Boundary Clocks in the G.8273.2 Standard.

The latest revision of this standard includes two new high-accuracy Clocks (Class C and D) that are subject to tighter performance constraints to existing Class A and B Clocks. To ensure that the performance of the Slave Clock in this test configuration matched that of a real-world scenario, the test involved a series of source failover events.

Such events are used to stress the ability of the Slave Clock to cope with such switching and that its performance was not adversely affected, enabling it to maintain its required performance levels.

The team used a chain of Clock Class D Boundary Clocks coupled to 2 Grandmasters to provide the stimulus to this test. We involved multiple combinations of such configurations to ensure that a valid mix of devices is tested.

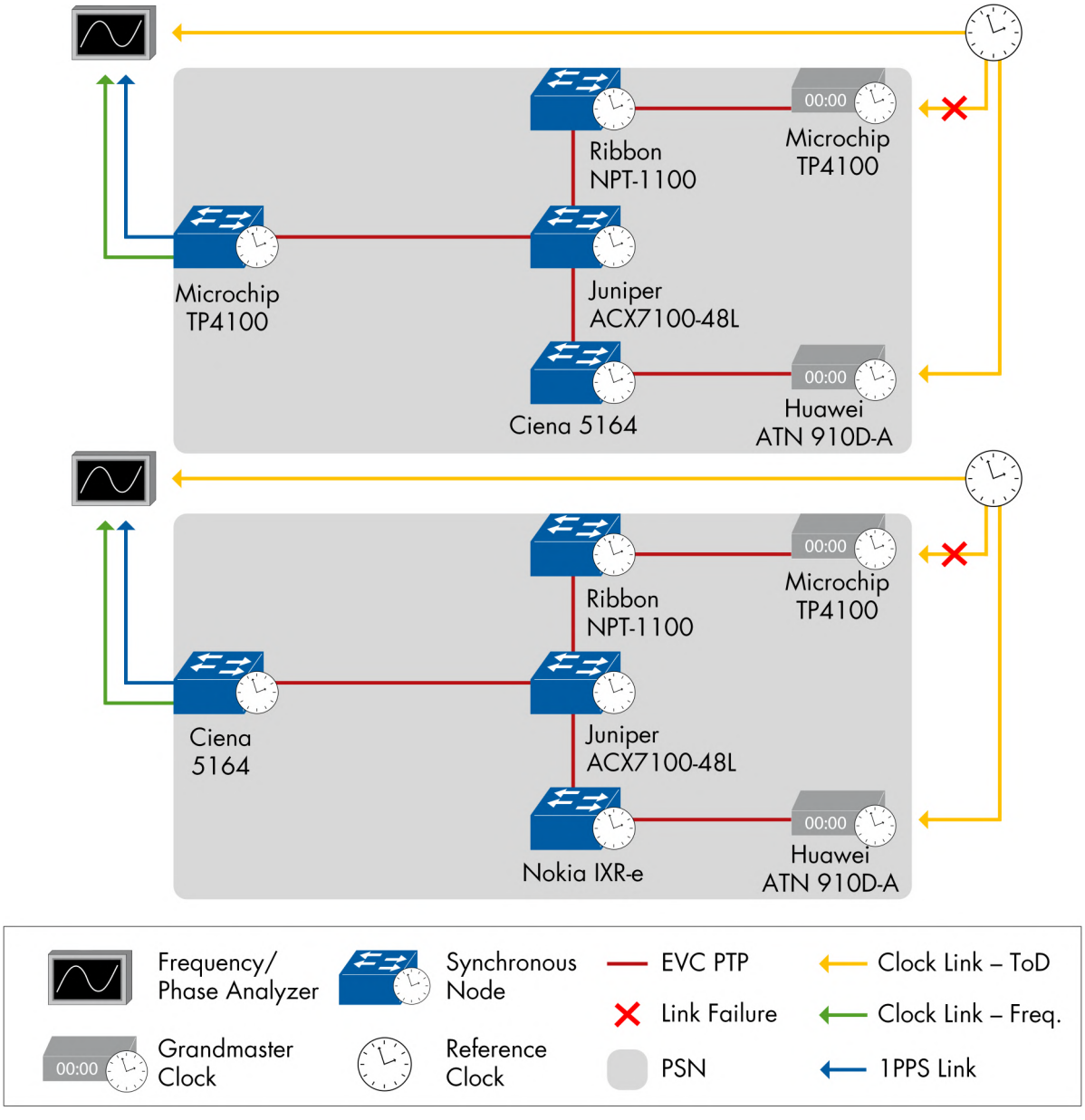


Figure 61: Phase/Time Full Timing Support Boundary Clocks Class-C/D Test

In each case, we used the Calnex Paragon-T measurement analyzer to accurately analyze the Slave time error Performance. In each of the test configurations, the Slave met the CLASS 6A performance limits.

The following DUTs successfully participated in the test, as

- T-BC: Ciena 5164, Nokia IXR-e, Juniper ACX7100-48L, and Ribbon NPT-1100
- T-TSC: Microchip TP4100
- GM: Microchip TP4100, Huawei ATN 910D-A

We did not find real issues other than ensuring that the GNSS antenna cable compensation was correctly entered into each of the GMs to ensure timing alignment. This required the tests to be run in some instances after the degradation of the primary source.

Conclusion

MPLS, SDN, Segment Routing, and EVPN services are maturing, as the results of this year's interoperability tests show once more. Service provider transport networks benefit from a wide range of standardized protocols and design blueprints these days. At the same time, network design need to take more complex and diverse requirements into account. The purpose of our series of interoperability events is to document well proven ways to establish multi-vendor end-to-end SDN networks. We hope that the wealth of information shared this year provides some reasonable advice for service provider design and operations teams.

As a result of this event, we can conclude: For the many transport network configuration challenges to come with 5G standalone network deployments and with further cloudification, the industry is already well equipped with a powerful SDN toolbox. At least, as far as the vendors frequently participating in our interop showcases are concerned. At EANTC, we are proud to help improve the multi-vendor interoperability and industry openness with this event again! We hope that we will be able to conduct a physical event in Paris next spring again, and look forward to seeing many vendors and network operators again soon!

EANTC AG
Salzufer 14, 10587 Berlin, Germany
info@eantc.de, https://www.eantc.de/
[v1.1 20210929]