MPLS+SDN
+NFVWORLD
@PARIS2018

White Paper 2018

INTEROPERABILITY SHOWCASE

EANTC

Editor's Note

Carsten Rossenhövel
Managing Director, EANTC

Packet transport networks are amid a quiet yet successful evolution cycle. After years of software-defined networks (SDN) revolution noise without many actual deployments, the industry converges on evolutionary solutions for SDN. Today's SDN-WAN solutions are interoperable, integrate pre-existent MPLS implementations, and adopt the diversity and complexity of service provider network routing and services. Admittedly, the MPLS/SDN use case scenarios we tested this year were some of the most complex so far; yet, it seems this is the most viable evolutionary path to fulfill future service requirements.

21 vendors participated in the EANTC interoperability event this time – one of the largest numbers ever. The level of interoperability for Ethernet VPN (EVPN) services and Segment Routing (SR) implementations over MPLS (SR-MPLS) has been very reassuring. We have seen tangible progress in the number of successful multi-vendor combinations, the maturity of implementations as expressed by more complex test cases, and the efficiency of configuration and troubleshooting. There is a major evolution going on quietly: Legacy signaling protocols LDP and RSVP-TE will no longer be needed in the future, greatly improving the scalability and efficiency of core and aggregation networks.

Path-Computation Element Protocol (PCEP) tests were much more promising than last year as well; router ("PCC") and controller ("PCE") implementations are increasingly aware of multi-vendor scenarios and ready to interoperate with each other in a collaborative way. In other areas such as Segment Routing over IPv6 ("SRv6"), a few vendors are doing ground breaking work; it might take a little more time until SRv6 becomes widely deployable.

What is the secret sauce of these successful developments? It is about continuity and consistent standardization along the lines of industry requirements. At the 101st IETF meeting in London a few weeks ago, George Swallow retired. As one of the main inventors of MPLS standards, he and many other key contributors have spent two decades defining, expanding, and maturing today's MPLS-based service provider packet transport networks. As a result, MPLS is still around, running virtually all service provider transport networks worldwide today and morphing into SDN-WAN, ready to serve future requirements. Continuous expansion and re-invention of data forwarding, signaling and routing techniques has secured its long and continuing lifetime.

By the way – the number of vendors participating in our event keeps growing continually due to another aspect of the secret sauce: It is not trivial but rewarding to implement MPLS/SDN, and there are many different viable approaches. Microwave vendors showed full MPLS data plane and control plane integration this time; a white-box router vendor together with a company producing protocol stacks for white boxes demonstrated SDN integration; and a first-time participant showed how to jump-start into SRv6 successfully. All these are great success stories, expanding service providers' choices and solution diversity.

Fast forward to the future: 2018 is the year that we will remember as "the year before 5G deployments took off." Things need to get real now: Operators discover that there is a lot to do which cannot be delayed any further – backbones need to scale in anticipation of the 5G traffic growth and much higher number of base station sites; software-defined network controllers need to calculate optimal paths for each of the new slices (service classes); mobile edge computing and service virtualization will create much more East-West traffic as well. Multiple network parts and elements need to be integrated to form a consistent end-to-end 5G transport network.

It has been great to witness how the participating vendors are getting ready to fulfill these 5G transport network requirements. One of the 5G-relevant test areas is network clock synchronization. It was supported by a large group of vendors again – participants who have continually improved multi-vendor network clocking at our events for about ten years now. The solutions are rock-solid meanwhile, and this test area is always one with most reliable results in our interop events.

Next year, we will increase clock synchronization precision requirements further for some of the 5G Release 15 requirements. In addition, we plan to expand the SDN tests, revisiting areas that showed only limited participation or success this year. In the coming years we will focus much more on domain and potentially even service orchestration: Auto-mated service provisioning, fault management, performance monitoring, and other management aspects become increasingly important. It seems this has always been the successful motto of MPLS:

Table of Content

Editor's Note 3
Introduction 4
Participants and Devices 4
Interoperability Test Results 5
Segment Routing 5
Ethernet Virtual Private Network 13
Topology 20
Software Defined Networking 23
Microwave 30
Clock Synchronization 32
Summary 38

Never rest – always aspire to improve further. There is a lot to do to get 5G services off the ground; let's get started!

Introduction

For the last couple of years, we have been focusing on testing Ethernet VPNs, Segment Routing (SR) and the use of Path Computation Elements (PCE) to influence traffic forwarding and routing policies.

This year it was the time to probe the maturity of these technologies by encouraging the participation of more vendors than ever in our interoperability hot-staging. Our stretch goal for this year's event was integration.

We had the pleasure of performing tests with a total of 21 vendors, with test scenarios of up to 10 vendors in the same topology, performing any-to-any interoperability.

It was definitely a year of consolidation for Segment Routing as the new standard for MPLS-enabled networks. All our test scenarios involving MPLS in the Segment Routing, Ethernet VPNs and Software Defined Networking sections were carried out using Segment Routing. Therefore, it showed us how mature vendor implementations are and a clear view, whereto the industry is moving forward.

We additionally tested for the first time SR implementations using IPv6 in the data plane (SRv6) and we verified vendor interoperability of some new proposed standards:

- "BGP Signaling of IPv6-Segment-Routing-based VPN Networks"
- "Segment Routing Prefix SID extensions for BGP"
- "IPv6 Segment Routing Header (SRH)"
- "SRv6 Network Programming"
- "Topology Independent Fast Reroute using Segment Routing"

In the EVPN section, we saw most of the new vendor faces, with broad support for EVPN bridging and EVPN Integrated Routing and Bridging (IRB). EVPN is already a well established technology for Data Center use-cases but we are seeing it more and more present as a unified control-plane technology to provide L2VPN/L3VPN services across WAN and Core network deployments.

Implementations of Path Computation Element Protocol (PCEP) also showed a good level of maturity. This year we could test a total of 31 combinations of different vendor/products interoping as PCE and Path Computation Clients (PCC).

This was also the first year where we could test disaggregated hardware/software, with some white-box and Network Operating System (NOS) vendors participating.

Furthermore, in the NETCONF/YANG section we reached a notable milestone by provisioning an end-to-end L3VPN service in a multi-vendor environment by using standardized IETF YANG models.

As always, packet clock synchronization area showed consistent results with many vendors supporting the latest PTP profiles for time/phase synchronization.

In the microwave section we saw a great deal of effort to integrate the wireless transport into IP/MPLS transport networks. We find this to be a critical requirement to support next-generation mobile networks, where end-to-end network slicing will play a key role for the diverse 5G use-cases.

Participants and Devices

Participants	Devices
Adva	FSP150 ProVMe
Arista Networks	7050SX2-72Q 7280SR-48C6
BISDN GmbH	Basebox controller (external) Switch AG7648 (Delta Electronics)
Calnex	Calnex Paragon-T Calnex Paragon-X
Cisco	ASR 9000 CSR1kv IOS XRv9000 NCS 5500 Network Services Orchestrator (NSO) Nexus 7702 Nexus 93180-FX
Delta Electronics	AGC7648A
ECI Telecom	NPT-1800
Ericsson	Baseband 6620 Baseband 6630 MINI-LINK 6651 MINI-LINK 6352 MINI-LINK 6366 MINI-LINK 6691 MINI-LINK 6693 Router 6371 Router 6471 Router 6672 Router 6675
Huawei	CX6608 CX600-X2-M8A NE40E-X2-M8A NE40E-M2K NE9000-8 Network Cloud Engine (NCE)
IP Infusion	OcNOS-AS7712-32X OcNOS-Virtual Control Machine

Participants	Devices
Ixia	IxNetwork Novus One
Juniper Networks	MX80-P MX104 MX240 QFX10002-72Q QFX5110-48S
Meinberg	LANTIME M1000S LANTIME M4000
Metaswitch Networks	Metaswitch CNRouter
Microsemi	TimeProvider 2300 TimeProvider 2700 TimeProvider 4100 TimeProvider 5000 TimeProvider 5000 Expansion 10
NEC	iPASOLINK VR
Nokia	7750 SR-7 Network Services Platform (NSP)
Oscilloquartz	OSA5421 HQ++
Spirent Communications	Attero-100G TestCenter (STC) TestCenter Virtual (STCv)
UTStarcom	SOO Station UAR500
ZTE Corporation	ZENIC WAN Controller ZXR10 M6000-3S ZXR10 M6000-5S ZXR10 M6000-8S PLUS ZXR10 M6000-18S ZXR10 T8000-18

Table 1: Participants and Devices

Interoperability Test Results

As usual, 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 proceed to solve interoperability issues after our test so there is no point in punishing them for their willingness to learn by testing. Confidentiality is vital to encourage manufacturers to participate with their latest - beta - solutions and enables a safe environment in which to test and to 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, Ixia and Spirent Communications for their test equipment and support throughout the hot staging.

Segment Routing

Segment Routing is becoming the de-facto SDN architecture. Leveraging the source routing paradigm, SR brings scalability, simplicity and end-to-end traffic engineering to MPLS and native IPv6 networks

Its architecture allows the use of different control-planes models. In a distributed model, Network Elements (NEs) use a dynamic routing protocol to allocate and distribute Segment Identifiers (SIDs). The routing protocol used for this purpose could be an IGP, such as IS-IS or OSPF with SR extensions, or BGP with SR extensions.

In a centralized model, external controllers can be leveraged for computation of paths that are then encoded in a SID list. A variety of methods including PCEP, BGP, NETCONF could be used to signal these SR policies to the NEs. For the former, results are covered in the PCEP section of this white paper.

Additionally, the SR architecture can be instantiated over various data planes: SR over MPLS (SR-MPLS) and SR over IPv6 (SRv6).

Throughout the SR section we will present several executed test cases using different combinations of control plane protocols and data plane encapsulations.

Segment Routing over IPv6 (SRv6)

IPv6 Routing over SRv6

Segment Routing uses network programming function concepts to enable packet forwarding through a specific path, different from the default IGP shortest path. A number of standard SRv6 functions are specified in the SRv6 Network Programming IETF draft (filsfils-spring-srv6-network-programming).

In this scenario we tested both the END and END.X functions.

The END function is the most basic function. It is used to steer traffic along the shortest-path to the advertising node.

On the other hand, the END.X function is used to steer traffic along the shortest path to the advertising node and then cross-connect it to a particular neighbor.

During our test, we verified the expected END and END.X data plane forwarding behavior and IPv6 SR Header (SRH) handling by Cisco and UTStarcom devices for SRv6 encapsulated traffic generated by Spirent TestCenter (STC) and Ixia IxNetwork.

In our first scenario, UTStarcom's UAR500 performed the END function and Cisco's NCS 5500 performed the END.X function. In the second scenario, the roles were inverted accomplishing the same results.

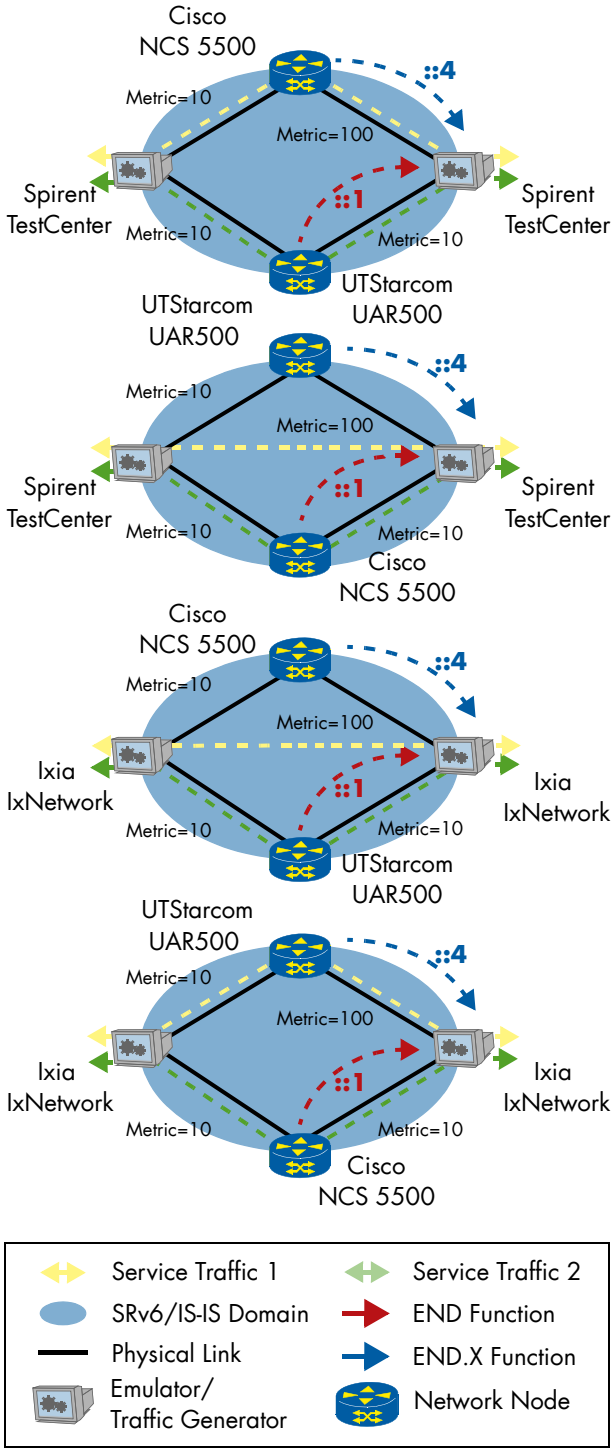


Figure 1: SRv6 IPv6 Routing

IPv4 VPN over SRv6

The draft "dawra-idr-srv6-vpn" defines procedures and messages for BGP SRv6-based EVPNs and L3 VPNs in order to provide a migration path from MPLS-based VPNs to SRv6 based VPNs.

In order to provide an SRv6 based VPN service, the egress PE signals an SRv6-VPN SID with the VPN route via MP-BGP. SRv6-VPN SID refers to an SRv6 SID that may be associated with one of the END.DT or END.DX functions defined in the IETF draft "filsfils-spring-srv6-network-programming".

In our test, vendors configured an IPv4 L3VPN and the egress node performed the END.DT4 function, which performs the Endpoint (END) function with decapsulation and IPv4 table lookup.

The ingress PE encapsulates the VPN packets in an outer IPv6 header where the destination address is the SRv6-VPN SID provided by the egress PE. Additionally, the ingress PE inserts the Segment Routing Header (SRH) which allows traffic engineering based on the SIDs listed in the Segment-list field (SID-list).

Additionally, during this test case execution, the transit node (P) performed the END function, updating the packet's IPv6 destination address with the next SID. This behavior was possible due to inclusion of Node P to the SID-list of the aforementioned SRH.

During the test, unidirectional traffic from TG1 to TG2 was sent by the test generator.

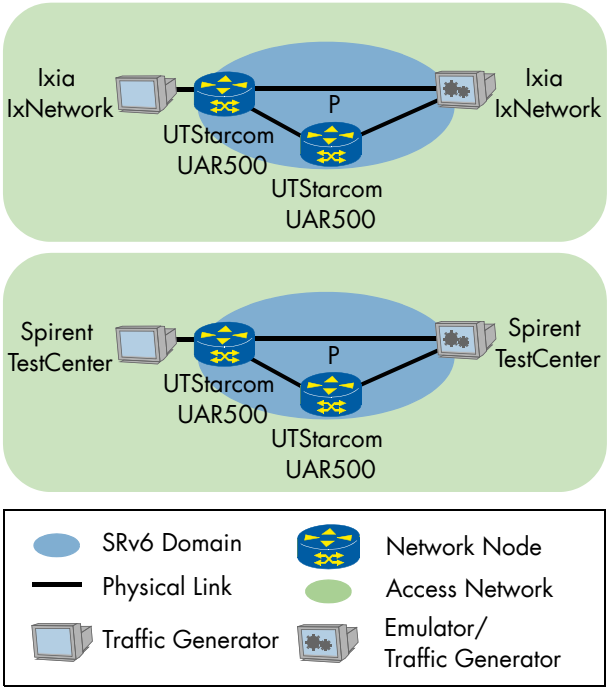


Figure 2: IPv4 VPNs with SRv6 Core

In this test we additionally verified that BGP can be used to advertise the reachability of prefixes in a particular VPN from an egress Provider Edge (egress-PE) to ingress Provider Edge (ingress-PE) nodes.

Segment Routing and LDP Interworking

We tested that an SR mapping server can be used to provide interworking between SR and LDP networks. The mapping server advertises a remote-binding segment id for prefixes attached to non-SR capable LDP nodes.

Additionally, we verified that an end-to-end LSP can be built when one part of the network is Segment Routing enabled and the other part relies on LDP exclusively for label allocation and distribution.

First, we verified that the mapping server advertises the range of prefixes corresponding to non-SR capable nodes and their associated SIDs/Labels.

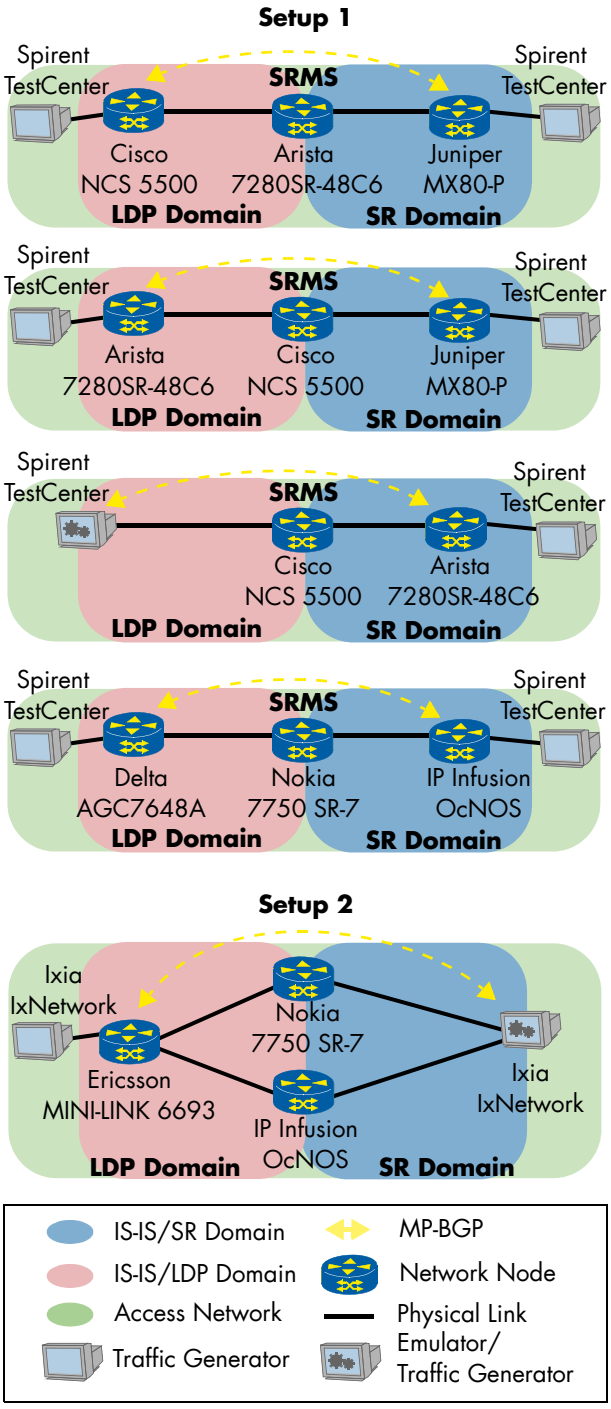


Figure 3: SR/LDP Interworking

Then, we verified that the SR nodes (mapping clients) process the advertised mapping and program their MPLS forwarding accordingly.

Lastly, we verified that edge network nodes do a proper encoding of the data path and that services function appropriately between SR and LDP-only nodes.

The following vendors participated in setup 1 as:

- LDP Node: Arista 7280SR-48C6, Cisco NCS 5500, Delta AGC7648A, Spirent TestCenter
- SR Node: Arista 7280SR-48C6, IP Infusion OcNOS (AS7712-32X), Juniper MX80-P
- SRMS Node: Arista 7280SR-48C6, Cisco NCS 5500, Nokia 7750 SR-7

The following vendors participated in setup 2 as:

- LDP Node: Ericsson MINI-LINK 6693
- SR Node: Ixia IxNetwork
- SRMS Nodes: IP Infusion OcNOS (AS7712-32X), Nokia 7750 SR-7

Segment Routing Anycast Segment — Disjointness in Dual Plane Networks

In this section we verified that Segment Routing Anycast Segment could be used to disjoint traffic forwarding paths within dual plane networks.

Segment Routing provides a new solution for disjoint paths within dual plane networks. Disjointness allows to transport different traffic services across disjoint paths. This can be achieved by using SR Anycast segment in SR routers.

Anycast Segment Identifier (Anycast SID) is specified for a set of routers within the same data plane. Each SR router in the Anycast set advertises the same Anycast-SID, which represents ECMP-aware, shortest-path IGP route to the closest node of that Anycast set.

In this test, we tested that the service traffic can be diverted to specific data plane based on a configured policy in the ingress PE.

The following vendors participated as:

- Ingress PE node: Ixia IxNetwork, Juniper MX104
- P Node: Ericsson Router 6675, Juniper MX80-P
- Anycast Nodes: Arista 7280SR-48C6, Cisco ASR 9000, ECI NPT-1800, Ericsson Router 6675, Huawei CX6608, Nokia 7750 SR-7
- Egress PE Node: Ixia IxNetwork, Juniper MX80-P

Segment Routing Per-CoS Steering into Multi-Plane Network

With the same test bed as the previous test, depicted in Figure 4, we verified that a policy in the ingress PE can divert traffic into different data planes by leveraging DSCP marking to differentiate traffic flows and mapping them two different data planes.

Figure 5 shows the participants who tested this scenario.

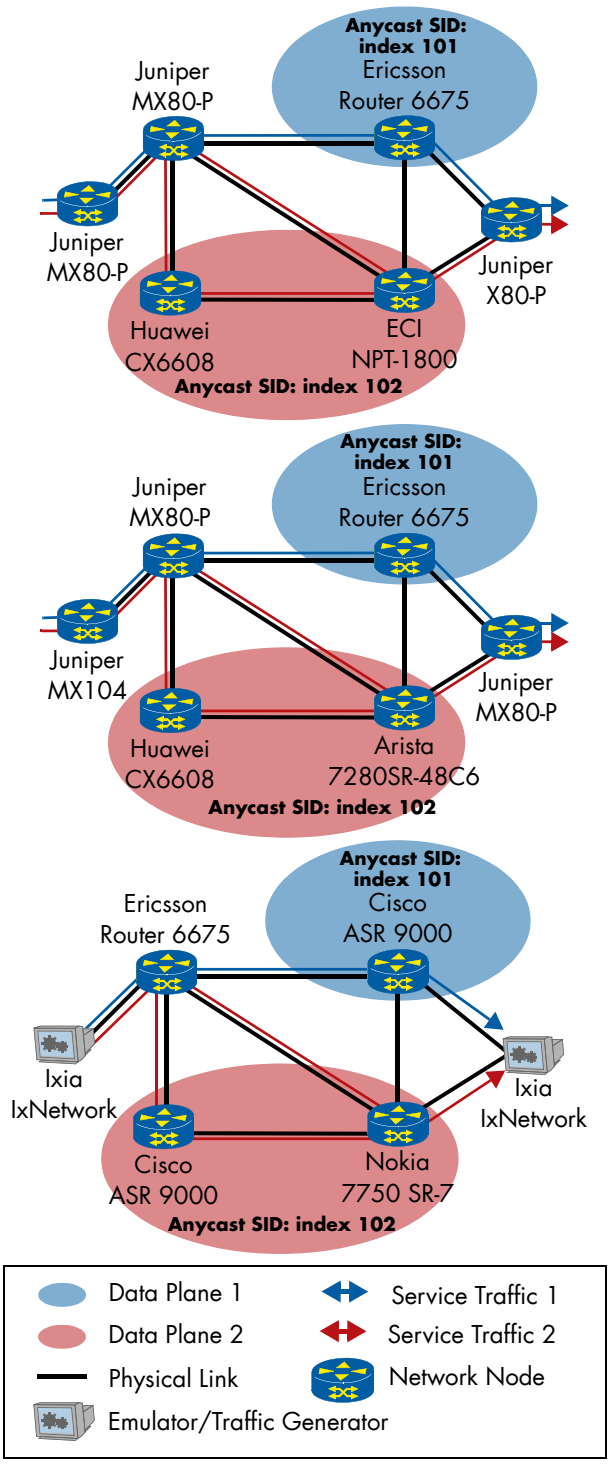


Figure 4: Segment Routing — Anycast Segment Disjointness in Dual Plane Networks

The following vendors participated as:

- Ingress PE node: Arista 7280SR-48C6, Juniper MX104
- P Node: Juniper MX80-P
- Anycast Nodes: Arista 7280SR-48C6, ECI NPT-1800, Ericsson Router 6675, Huawei CX6608
- Egress PE Node: Juniper MX80-P, Spirent TestCenter

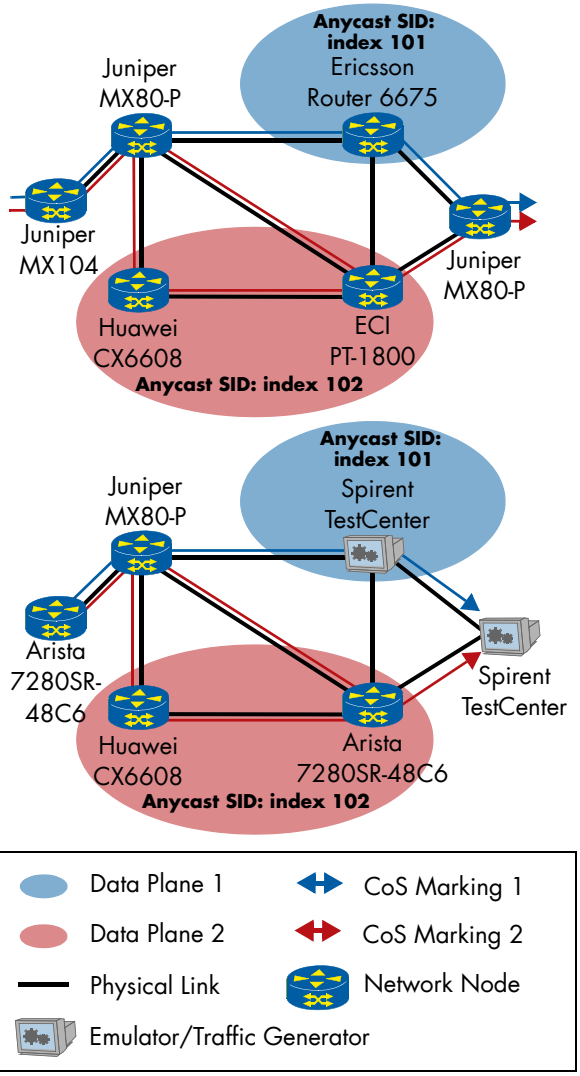


Figure 5: Segment Routing Per-CoS Traffic Steering

BGP Segment Routing

Segment Routing can be used in large scale Data Centers as a simple solution to provide traffic engineering and fast re-route capabilities in the DC fabrics. BGP is a popular choice as routing protocol in Clos topologies due to its scalable intrinsec properties.

For these reasons, we arranged two topologies using BGP-LU. In the first one, we tested 5 Leaves and 1 Spine exchanging IPv4 prefixes and their associated labels using BGP Labeled Unicast (BGP-LU) NLRI. Then we tested a similar topology with one Spine and three leaves but this time vendors enabled the BGP Prefix-SID attribute in the BGP-LU NLRI.

In the latter test bed, we additionally validated the correct SR forwarding and SID advertisement.

We tested full-mesh traffic forwarding between all Leaf nodes, and between the Leaf nodes and the node emulators—Ixia and Spirent (were applicable).

The following vendor equipment participated in the BGP-LU scenario: Arista (7280SR-48C6), Cisco (NCS 5500), Ixia (IxNetwork), Juniper (MX104) and Spirent (TestCenter).

Ixia IxNetwork and Spirent TestCenter acted as traffic generators and emulated Leaves.

For the Prefix-SID Label in Labeled Unicast NLRI (BGP-LU + SID) variant of the scenario; Arista, Cisco and Ixia participated with the same equipment as in the BGP-LU setup.

In this case, Ixia IxNetwork was used as a traffic generator and emulated Leaf.

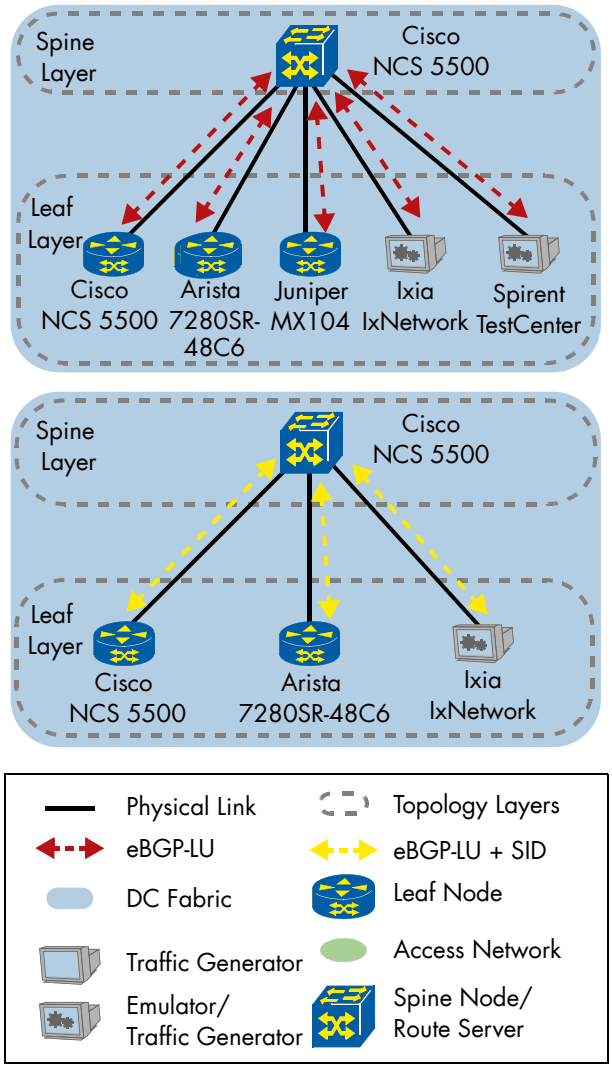


Figure 6: BGP Segment Routing

During this test we found out that one of the vendors acting as Route Server was not able to process/ propagate BGP updates with Prefix-SID TLVs, occasioning BGP session flaps upon reception. Due to this we removed the vendor from the Route Server function.

Resiliency and Monitoring

Segment Routing FRR/TI-LFA

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 link failures. The LFA approach is applicable when the protected node or Point of Local Repair (PLR) has a direct neighbor that can reach the destination without looping back traffic to the PLR. When the protected path fails, the traffic was sent to the neighboring node which in turn forwards the traffic to the destination.

When above conditions are not met, Topology Independent Loop-free Alternate (TI-LFA) can be used instead. It relies on segment routing to build a protection mechanism based on proven IP-FRR concepts. TI-LFA does not require any additional signalling between the Point of Local Repair (PLR) and the repair node — typically called PQ node.

In both cases, the destination is protected against the failure of a link. Additionally, the SRLG protection describes the situation, in which the destination is protected assuming that a configured set of links share fate, with the primary link which has failed.

During the test, initially we performed baseline measurement for packet loss using bidirectional traffic from a traffic generator.

Afterwards, vendors configured LFA/TI-LFA on the network nodes and verified that the network nodes installed backup forwarding entry in FIB. While the traffic was running via the primary path we disconnected the link and measured the service interruption time based on the packet loss. We saw that the traffic was taking the backup path.

Finally, vendors configured a new link (link 2) and configured it with the same SRLG as the failed link. We tested TI-LFA with SRLGs in this case.

We tested the following three scenarios: FRR / LFA link protection, TI-LFA link protection and TI-LFA local SRLG protection for both SR-MPLS and SRv6 data plane options.

Vendors participating in the IP FRR/LFA link protection tests with MPLS data plane (Figure 7) were:

- PQ Node: ECI NPT-1800, Ericsson Router 6675, Huawei CX6608
- PLR Node: ECI NPT-1800, Ericsson Router 6675, Huawei CX6608, Juniper MX80-P
- P Node: ECI NPT-1800, Huawei CX6608, Juniper MX80-P
- Egress PE: ECI NPT-1800, Ericsson Router 6675, Huawei CX6608, Juniper MX80-P

Vendors participating in the TI-LFA link protection tests with MPLS data plane (Figure 8) were:

- PQ Node: Cisco ASR 9000, ECI NPT-1800, Huawei CX6608, Juniper MX-80-P
- PLR Node: Cisco ASR 9000, ECI NPT-1800, Ericsson Router 6675, Juniper MX80-P
- P Node: ECI NPT-1800, Ericsson Router 6675, Huawei CX6608, Juniper MX80-P
- Egress PE: Ericsson Router 6675, Huawei CX6608, Juniper MX80-P, Nokia 7750 SR-7

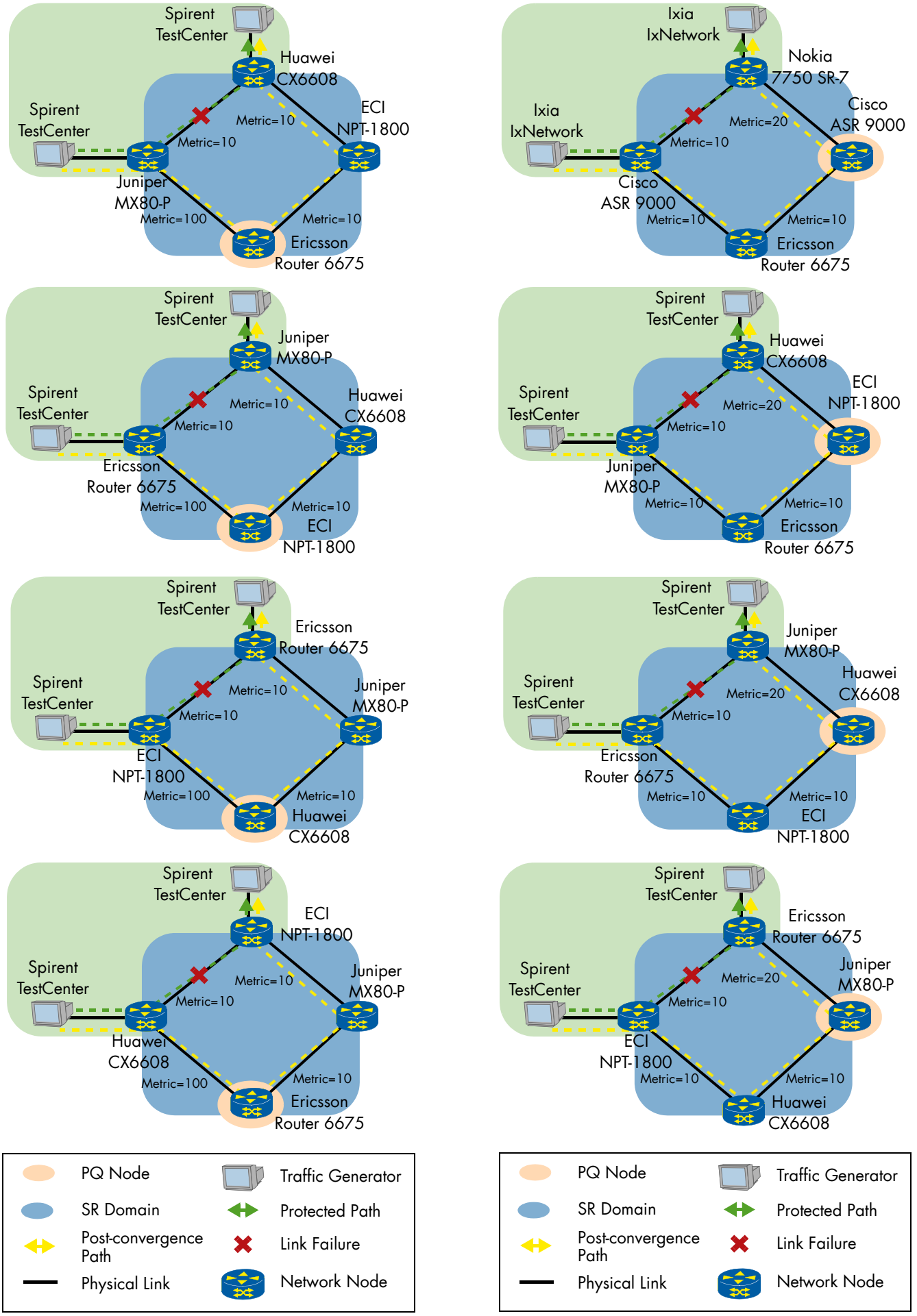


Figure 7: SR FRR/LFA Link Protection

Figure 8: SR-MPLS TI-LFA Link Protection

Vendors participating in the TI-LFA local SRLG protection tests and MPLS data plane were: Cisco ASR 9000 as PLR Node, Ericsson Router 6675 as P node, Cisco ASR 9000 as PQ node and Nokia 7750 SR-7 as egress PE node.

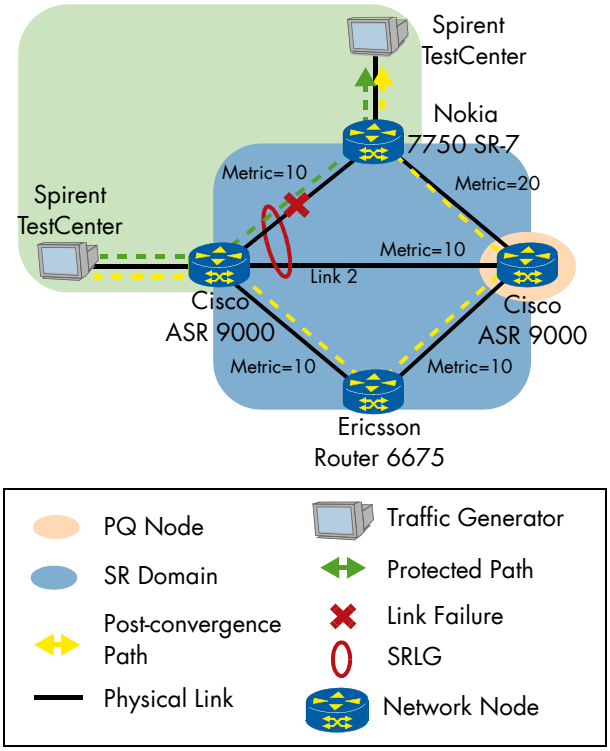


Figure 9: SR-MPLS TI-LFA Local SRLG Protection

During the SR-MPLS tests we observed that many vendors could Fast Reroute but not all of them were able to test TI-LFA, and even fewer were able to test TI-LFA with SRLG groups. We encourage vendors to work on these features so we can test more combinations next year.

SRv6 FRR

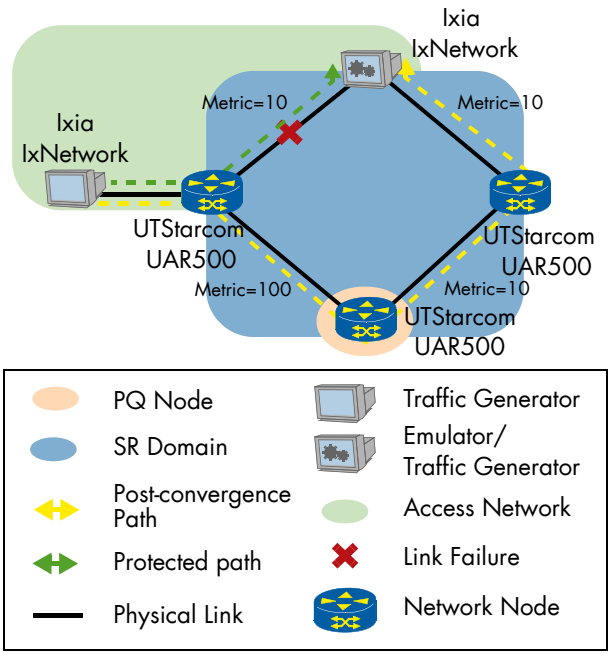
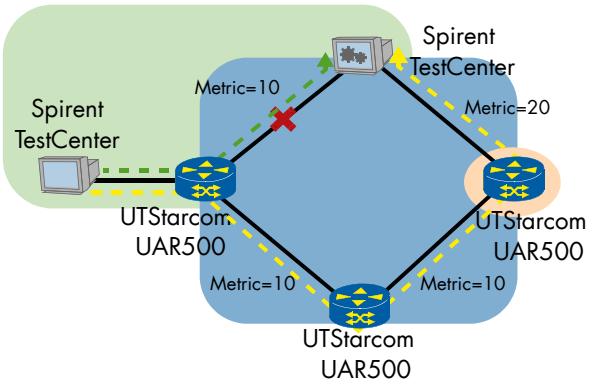


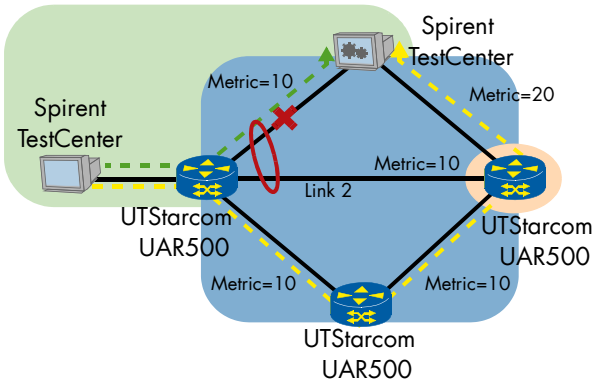
Figure 10: SRv6 FRR/LFA Link Protection

For SRv6, the TI-LFA implementation of UTStarcom is based on centralized controller SOO Station that runs PQ algorithm and calculates post-convergence path.

SRv6 TI-LFA



SRv6 TI-LFA with SRLG



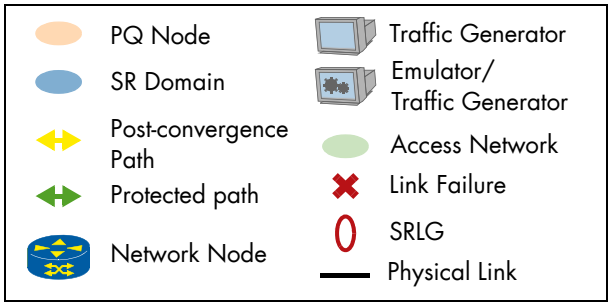


Figure 11: SRv6 TI-LFA Protection

Vendors participating in the SRv6 tests for FRR, TI-LFA and TI-LFA with SRLGS were:

- PQ Node: UTStarcom UAR500
- PLR Node: UTStarcom UAR500
- Egress PE: Ixia IxNetwork, Spirent TestCenter

We were glad to test the same set of features for SR-MPLS and SRv6, it shows that the gap between the two implementations is closing.

Label Switched Path (LSP) Ping/Trace

The IETF standard RFC 8287 defines the LSP ping and traceroute method for Segment Routing with MPLS data plane. Similar to conventional LSP ping/traceroute, the SR fault detection and isolation tools are also based on MPLS 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 MPLS 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 MPLS echo request has been forwarded from the correct path. The LSP ping/traceroute response is carried in a MPLS echo reply.

First, we verified that a SR sender can initiate an LSP ping/traceroute request to a target SR responder, which responded with an LSP ping/traceroute reply to the SR sender.

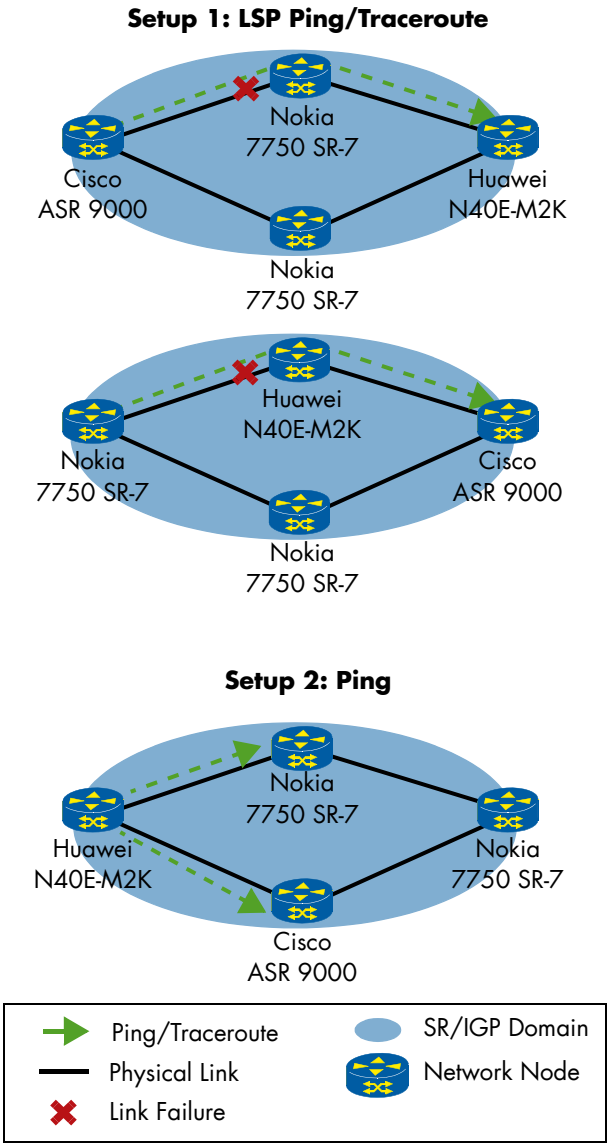


Figure 12: LSP Ping/Traceroute

Vendors participating in the LSP ping/traceroute tests were:

- Ingress Node: Cisco ASR 9000, Huawei N40E-M2K, Nokia 7750 SR
- P Node: Cisco ASR 9000, Huawei N40E-M2K, Nokia 7750 SR
- Egress PE: Cisco ASR 9000, Huawei N40E-M2K, Nokia 7750 SR

During the test, we observed that vendors have different interpretations regarding the sub-TLV type/length in FEC stack TLV for SR LSPs. Some vendors claimed that the standard (RFC 8287) does not say clearly whether to consider the reserved octets to be part of the Length or not. After the interop event, a technical errata request was raised at IETF by one of the participating vendors to clarify the length of the Sub-TLVs.

Ethernet Virtual Private Network

Multi-Vendor A/A Site for an EVPN MPLS VLAN-Based Service

Ethernet Virtual Private Network (EVPN) provides separation between the data plane and control plane, which allows the use of different encapsulation mechanisms in the data plane such as MPLS and Virtual Extensible LAN (VXLAN).

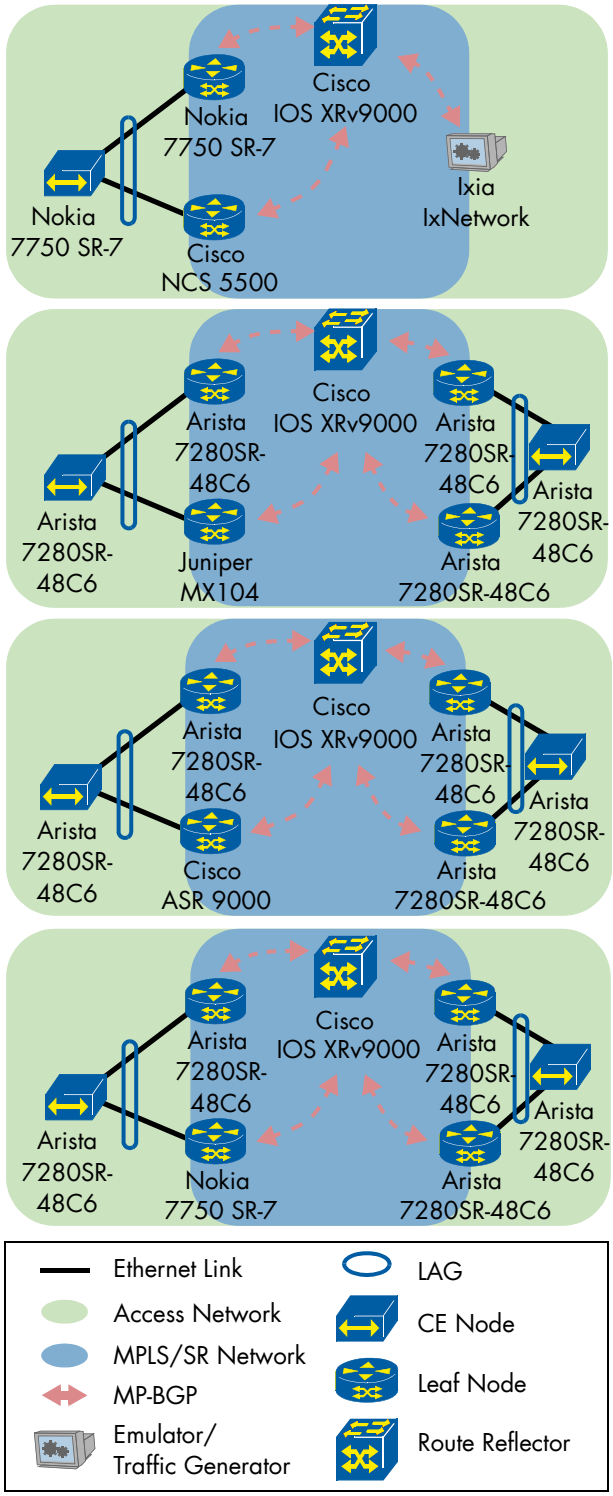


Figure 13: Active-Active EVPN with MPLS/Segment Routing Transport

This test verifies the EVPN functionality over MPLS data plane which is based on Segment Routing.

In the SR domain, IS-IS with SR extensions were used as the IGP protocol.

In this scenario, we run 4 different vendor combinations. In all of them, Cisco IOS XRv9000 was acting as Route Reflector.

After checking the IGP and the Segment Routing information, we verified that the Routes Type 1, 2, 3, and 4 were correctly imported.

Finally, we sent unicast end-to-end traffic using Ixia IxNetwork and Spirent TestCenter and observed no packet loss.

For multi-homed sites an additional CE device was used to setup a Link Aggregation Group (LAG) to the PE nodes. For this role either Nokia's virtual switch, within 7750 SR-7 chassis, or Arista's 7280SR-48C6 was used.

We verified the aliasing functionality in Active-Active multi-homed sites and that the Non-Designated Forwarder (NDF) was blocking remote's site BUM traffic.

For the PE roles we tested the following devices: Arista 7280SR-48C6, Cisco NCS 5500, Cisco ASR 9000, Juniper MX104 and Nokia 7750 SR-7 in multi-homed site configuration; and Ixia IxNetwork performing PE emulation in a single-homed configuration. The detailed vendor combinations are depicted in Figure 13.

Ethernet Line (E-Line)

The MEF defines a point-to-point Ethernet service as Ethernet Line (E-Line). The IETF now proposes a solution framework in order to support this service in MPLS networks using EVPN. The IETF discussed the features in the "VPWS support of the EVPN" draft and requires the use of VPWS to meet the E-Line requirements. In addition, EVPN also supports inherited functions to make the VPWS implementation more effective.

In the test, we setup a point-to-point connection between two given PEs as depicted in Figure 14. We configured an EVPN instance between each pair and enabled VPWS inside EVPN instances.

Due to time constraint issues, some participating vendors only tested in single-homing mode. During such test, we verified that the ESI field was set to zero and that the Ethernet Tag field mapped to the VPWS identifier, both of which were carried in the EVPN AD per EVI route.

The following vendors participated in this scenario: Cisco NCS 5500 (dual-homed), Huawei NE9000-8 (single-homed), Juniper MX104 (single-homed) and Nokia 7750 SR-7 (dual-homed) PE routers. Additionally, Ixia IxNetwork and Spirent TestCenter acted as emulated PEs and traffic generators.

Cisco IOS XRv9000 was used as route reflector, and a virtual-switch running in Nokia's 7750 SR-7 router was used as CE device.

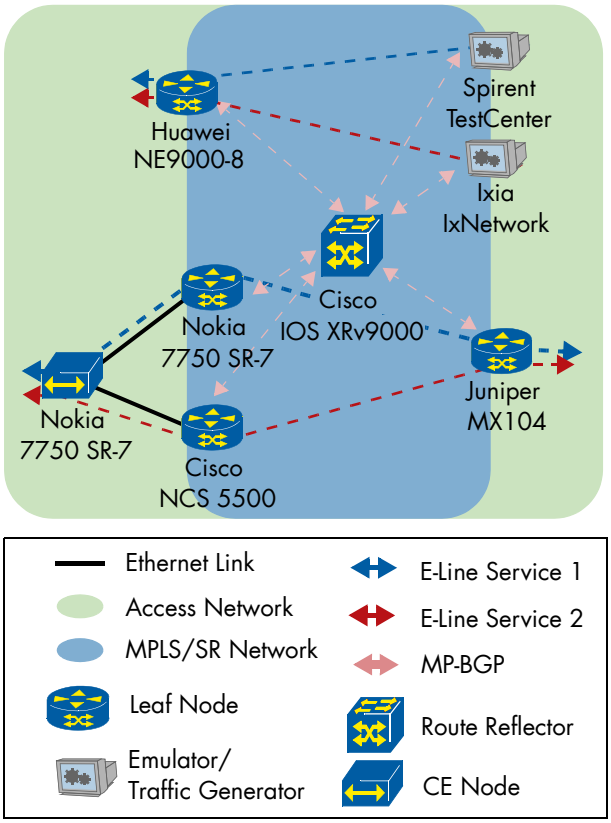


Figure 14: E-Line Services with SR-MPLS Transport

Ethernet Tree (E-Tree)

The MEF defines a rooted-multipoint Ethernet service as Ethernet Line (E-Line). Again, the IETF proposes a solution for supporting this service in MPLS networks by using EVPN.

In the setup we verified that the EVPN technology can support E-Tree functional requirements. Based on the current IETF standard (RFC 8317), we tested a root/leaf per AC (attached circuit) scenario.

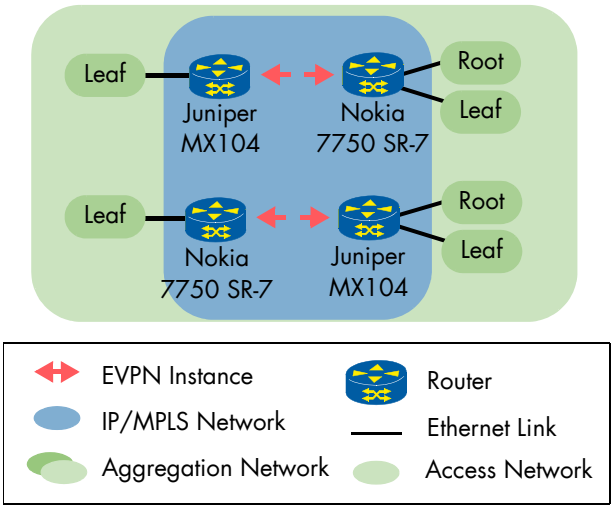


Figure 15: E-Tree Service -
Leaf or Root Sites per AC

We verified that the MAC addresses learned on a leaf Attachment Circuit where advertised with the expected leaf flag and installed in the remote PE as leaf MAC addresses.

We also verified that the two PEs exchanged the ESI Leaf label, used to identify BUM traffic generated from a leaf, as per the RFC8317.

The following vendors joined this test scenario: Juniper MX104 and Nokia 7750 SR-7.

EVPN Enhancements

ARP Proxy

Within EVPN, PEs advertise MAC/IP addresses, along with an MPLS label, to other PEs using Multi-Protocol BGP (MP-BGP).

ARP proxy functionality of EVPN eliminates ARP flooding within the transport network by advertising MAC addresses along with their corresponding IP addresses in the MAC/IP advertisement route, type-2. When a PE receives ARP request from its directly attached hosts, it intercepts the ARP Request and performs an IP/MAC lookup for the requested IPs. If the lookup is successful, the PE will send an ARP Reply in behalf of the requested IP endpoint. The ARP Request will not be flooded through the EVPN network or any other attachment circuit. If the lookup is not successful, the PE will flood the ARP Request in the EVPN network and the other local CEs.

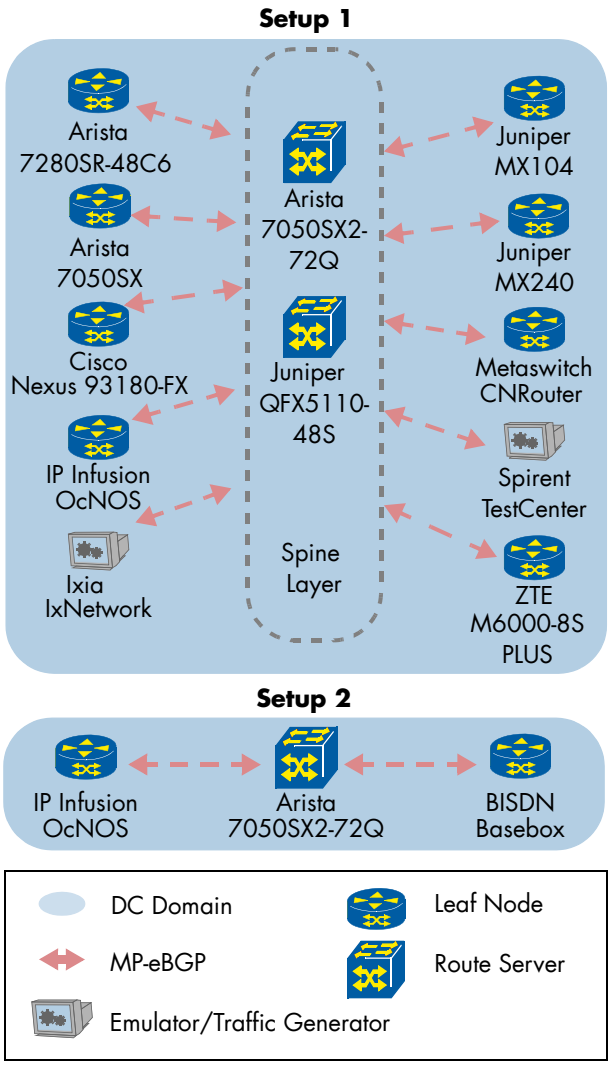


Figure 16: EVPN with ARP Proxy

We performed the ARP proxy tests with two different setups. In the first setup, only symmetric IRB forwarding was used. In setup 2 on the other hand, pure L2 traffic was forwarded, and routing was not involved in the setup.

Thanks to the broad vendor support for ARP Proxy, this year we could setup a typical CLOS topology, with the spine layer serving as Route-Server for eBGP sessions for both the underlay routing (IPv4), and the overlay (EVPN). Arista 7050SX2-72Q and Juniper QFX5110-48S acted as Route Servers for setup 1.

In the Leaf role for setup 1, as depicted in Figure 16 the following vendors participated: Arista 7050SX2-72Q, Arista 7280SR-48C6, Cisco Nexus 93180-FX, IP Infusion OcNOS (AS7712-32X), Juniper MX104, Juniper MX240, Metaswitch CNRouter, ZTE ZXR10 M6000-8S PLUS.

For setup 2, the leaf role was fulfilled by BISDN Basebox and IP Infusion OcNOS (AS7712-32X). Arista 7050SX2-72Q acted as Route Server.

Unfortunately there was not enough vendor support to test ND Proxy this year.

IGMP Proxy

The goal of IGMP proxy mechanism is to reduce the flood of IGMP messages (both Queries and Reports) in EVPN instances among PE Routers, just like ARP/ND suppression mechanism in EVPN reduces the flooding of ARP messages over EVPN.

Hosts in a VXLAN domain express their interests in multicast groups on a given subnet/VLAN by sending IGMP membership reports (Joins) for their interested multicast group(s). Furthermore, 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.

Furthermore, if there is no physical/virtual multicast router attached to the EVPN network for a given multicast group (*,G), or multicast sender (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.

In this test Cisco Nexus 93180-FX and Nokia 7750 SR-7 participated as PE routers. We emulated a host behind Cisco's single-homed PE sending IGMP reports and we observed how this were propagated as Route-Type-6 (SMET route) over the EVPN overlay.

The Nexus 9000 Leaf originated the EVPN Route-Type6 based on the Join received from the Emulated Host. The Multi-Site Border Gateway (BGW) relayed this message and forwarded the SMET-Route to the external EVPN Speaker (Nokia) seamlessly.

On the other hand, the current behavior in one vendor implementation is to send encapsulated IGMP messages to feed IGMP Proxy to avoid unnecessary unknown multicast flooding.

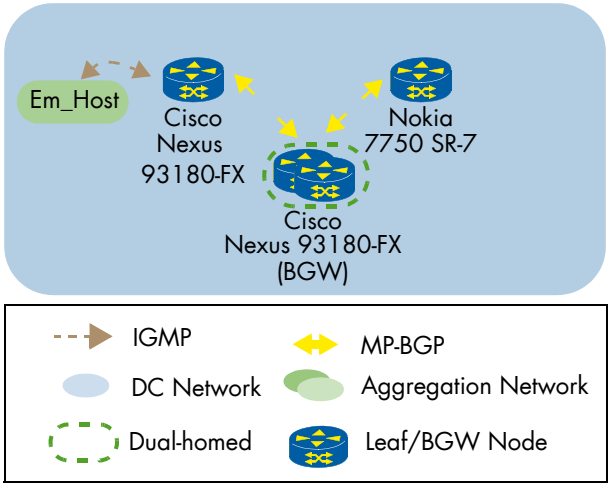


Figure 17: IGMP Proxy with EVPN

EVPN Routing

EVPN - Integrated Routing and Bridging

Ethernet VPN Integrated Routing and Bridging (IRB) status is currently "work in progress" at the IETF and provides solution for inter-subnet forwarding in data center environments with EVPN. MP-BGP EVPN enables communication between hosts in different VXLAN overlay networks by distributing Layer 3 reachability information in the form of either a host IP address route(route type-2) or an IP prefix (route type-5). Depending on the required lookup at the ingress or/and egress Network Virtualization Edge (NVE), the draft defines two different semantics for IRB: Asymmetric and symmetric IRB model.

While the asymmetric IRB semantic requires both IP and MAC lookups at the ingress NVE with only MAC lookup at the egress NVE, in the symmetric IRB semantic, both IP and MAC lookup are required at both ingress and egress NVEs.

We tested using a three-stage Clos topology for all profiles, also referred to as a "Leaf and Spine" network (as discussed in RFC 7938). The route servers acted as spine switches and aggregated a set of horizontal EVPN PE devices as leaves. A fixed number of IPv4 subnets was connected to every leaf device. Then we configured eBGP in the physical network between spine and leaf devices (as underlay). External-BGP was used between the spine and leaf to advertise the overlay EVPN routes with a VXLAN forwarding plane.

In this setup, the EVPN-VXLAN was accessed by both IPv4 subnets connected to the leaf device whereas the subnets were emulated by the traffic generators.

In all test combinations we verified full-mesh connectivity between all leaves with traffic generators.

EVPN - Symmetric Integrated Routing and Bridging

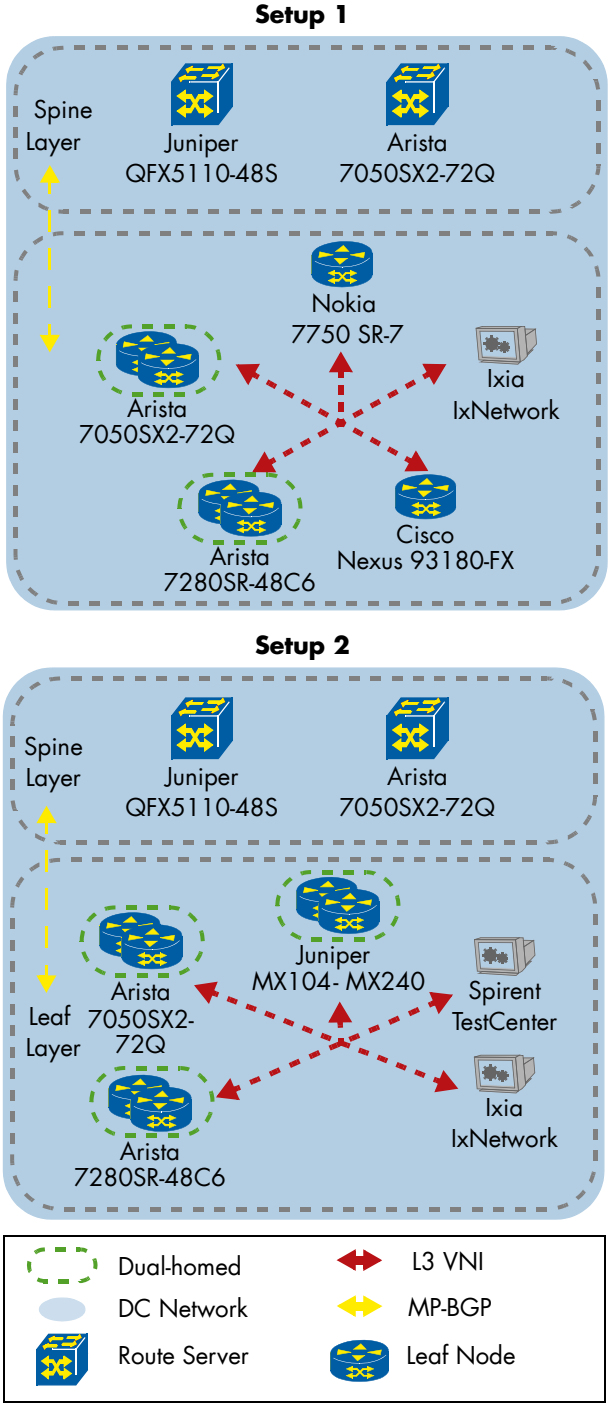


Figure 18: EVPN Symmetric IRB

As depicted in Figure 18, the following vendors successfully participated in the test as EVPN PE routers:

- Setup 1: Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Cisco Nexus 93180-FX, Ixia IxNetwork, Nokia 7750 SR-7
- Setup 2: Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Ixia IxNetwork, Juniper MX104, Juniper MX240 (dual-homed), Spirent TestCenter (STC)

Juniper QFX5110-48S and Arista 7050SX-72Q acted as Spine switches and BGP route servers.

EVPN - Asymmetric Integrated Routing and Bridging

The following vendors successfully participated in the test as EVPN PE:

- Setup 1: Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Metaswitch CNRouter, Spirent TestCenter (STC),
- Setup 2: Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Ixia IxNetwork, Nokia 7750 SR-7

Arista 7050SX2-72Q was placed as Spine node acting as Route Server. The setup is depicted in Figure 19.

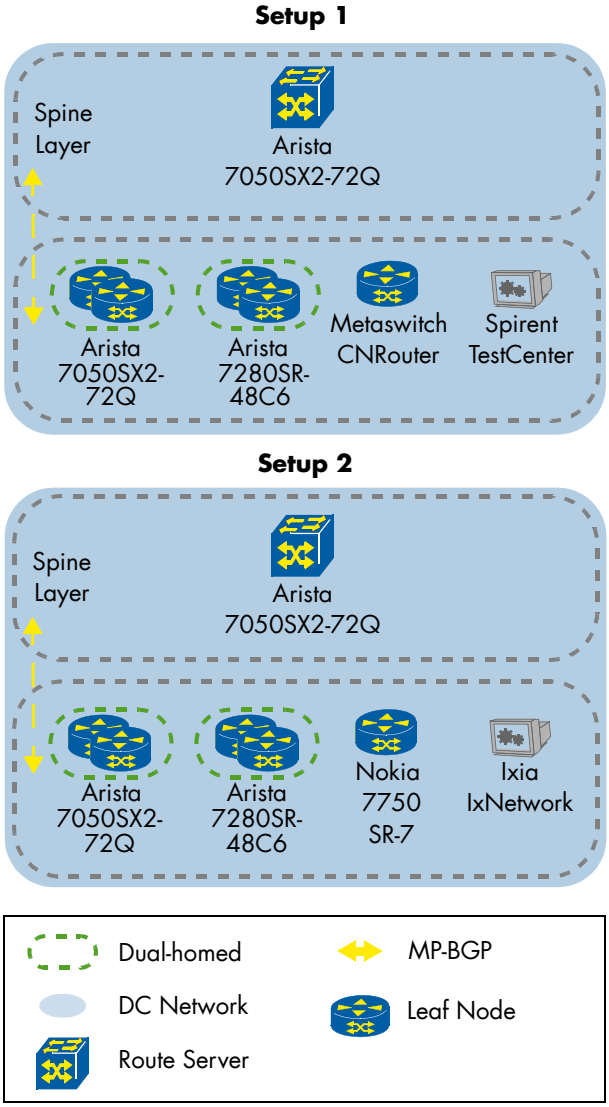


Figure 19: EVPN Asymmetric IRB

EVPN IP-VRF-to-IP-VRF

Ethernet VPN Prefix Advertisement draft status is "work in progress" at the IETF and provides a solution for efficient handling of inter-subnet forwarding in a data center environment with EVPN. In an EVPN network environment, there is a requirement for IP prefix advertisement for subnets and IPs residing behind an IRB interface. This scenario is referred to as EVPN IP-VRF-to-IP-VRF.

The EVPN prefix advertisement draft provides different implementation options for the IP-VRF-to-IP-VRF model:

- Interface-less model, where no Supplementary Broadcast Domain (SBD) and overlay index are required
- Interface-full with unnumbered SBD IRB model, where SBD is required as well as MAC addresses as overlay indexes
- Interface-full with SBD IRB model, where SBD is required as well as Gateway IP addresses as overlay indexes

In the test we focused on EVPN-VXLAN for VXLAN data plane provisioned within the data center fabric. External BGP (eBGP) was used for both underlay and overlay NLRI exchanges. The eBGP configuration on the Spine nodes was modified so that the next-hop attribute was not changed during BGP update propagation between Leaf nodes.

For all tests, we verified that the VXLAN virtual network identifier (VNI) was directly mapped to the EVPN EVI. We confirmed that the RT-5 (IP Prefix advertisement route) carried the correct IP Prefix and length, as well as the corresponding Gateway IP address (zero in case of the Interface-less model). The route-tables were verified via CLI. Additionally, a RT-2 was used in interface-full mode. It carried MAC address length and MAC address. The IP length was set to 0. Following, we sent IPv4 test traffic from all IPv4 subnets to any other IPv4 subnets, and expected to receive traffic on all IPv4 subnets without any packet loss.

Interface-full with Unnumbered SBD IRB Model

This test is depicted in Figure 20 and the following vendors participated as PE routers:

- Cisco Nexus 93180-FX, Nokia 7750 SR-7

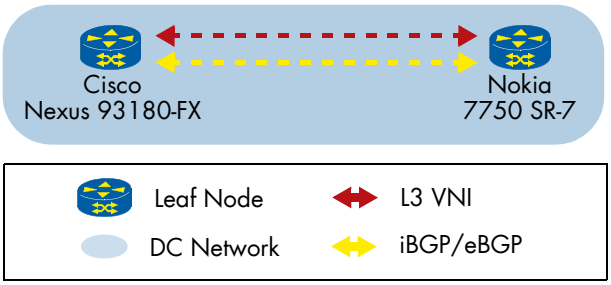


Figure 20: Interface-full with Unnumbered SBD IRB Model

Interface-full with SBD IRB Model

We observed some issues in this test case, as some vendors do not handle the data plane properly when different VNI value used in Type 2 and Type 5 Routes, though the VNI of type 5 route is irrelevant in this case.

For this reason, IxNetwork and Spirent TestCenter in setup 1 and 2 respectively could only successfully generate traffic to Nokia 7750 SR-7 node.

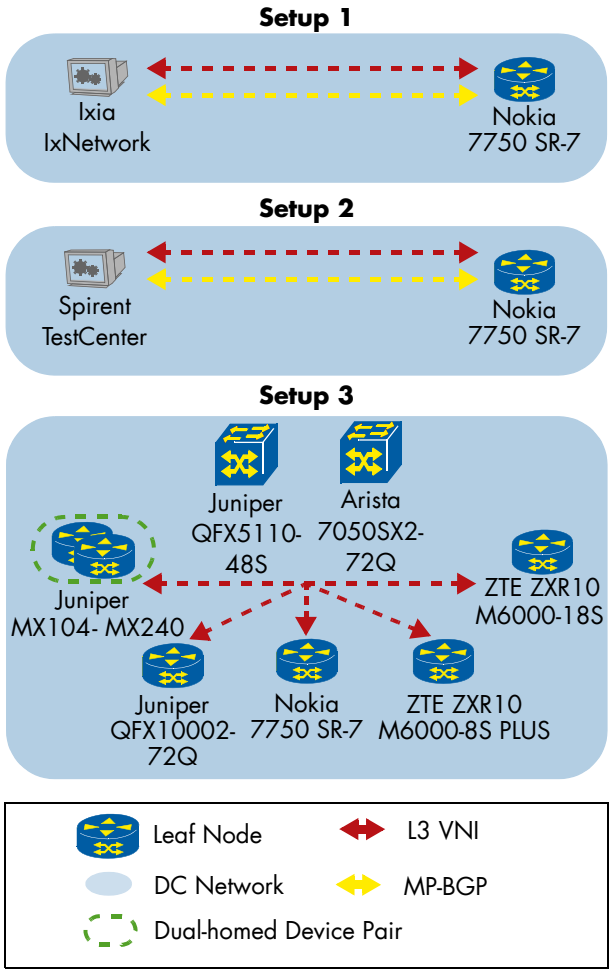


Figure 21: Interface-full with SBD IRB Model

This test is depicted in Figure 21 and the following vendors participated as PE routers:

- Setup 1: Ixia IxNetwork, Nokia 7750 SR-7
- Setup 2: Nokia 7750 SR-7, Spirent TestCenter
- Setup 3: Juniper MX104-MX240 (multi-homed), Juniper QFX10002-72Q, Nokia 7750 SR-7, ZTE ZXR10 M6000-18S, ZTE ZXR10 M6000-8S PLUS

Additionally in setup 3, Arista 7050SX2-72Q and Juniper QFX5110-48S acted as Spines/BGP Route servers in this scenario.

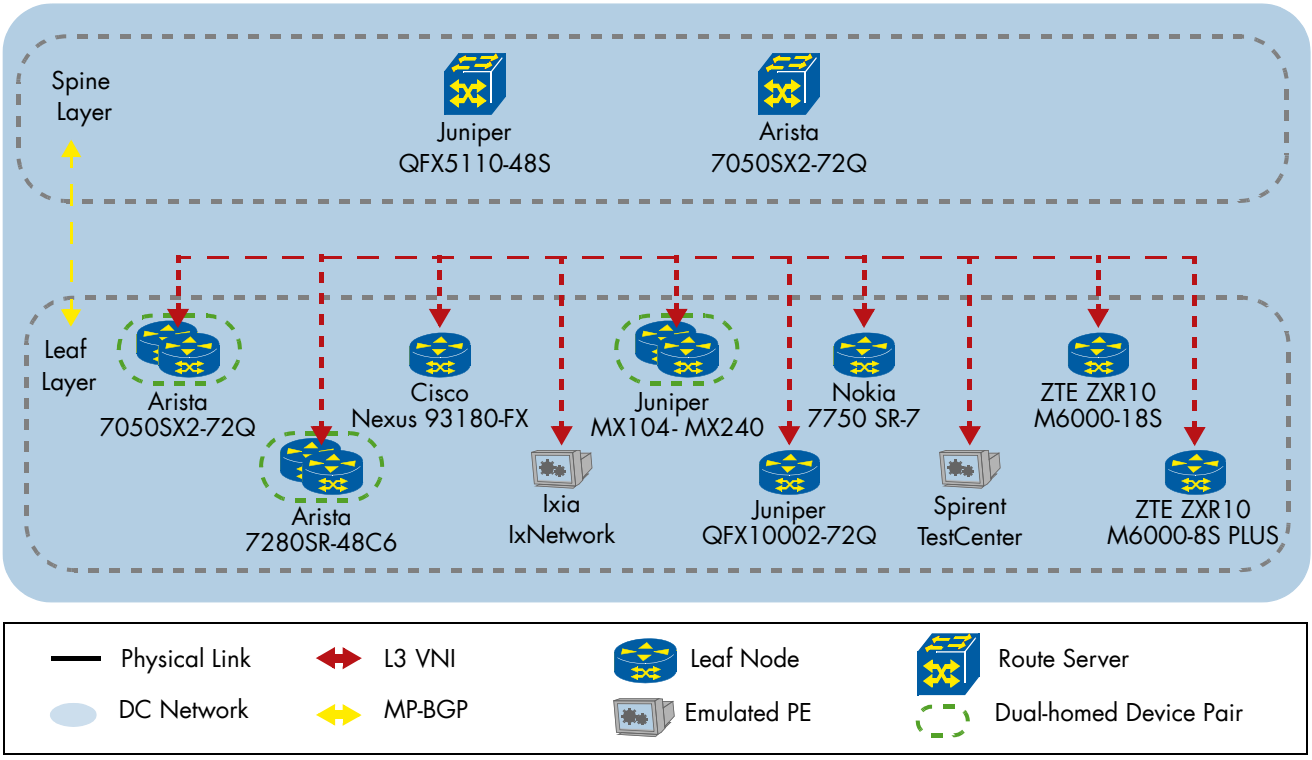


Figure 22: Interface-less Model EVPN IP-VRF-to-IP-VRF

Interface-less Model

Interface-less model was the most widely supported model to do IP routing with Prefix-advertisement (route type-5) in EVPN overlays.

This test is depicted in Figure 22 and the following vendors participated as PE routers:

- Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Cisco Nexus 93180-FX, Ixia IxNetwork, Juniper MX104, Juniper MX240 (dual-homed), Juniper QFX10002-72Q, Nokia 7750 SR-7, Spirent TestCenter, ZTE ZXR10 M6000-8S PLUS, ZTE ZXR10 M6000-18S

Arista 7050SX2-72Q and Juniper QFX5110-48S acted as route servers.

We did not observe any issues which show mature vendor implementations and clear definition in the standard procedures.

EVPN Interworking

EVPN and IP-VPN Interworking

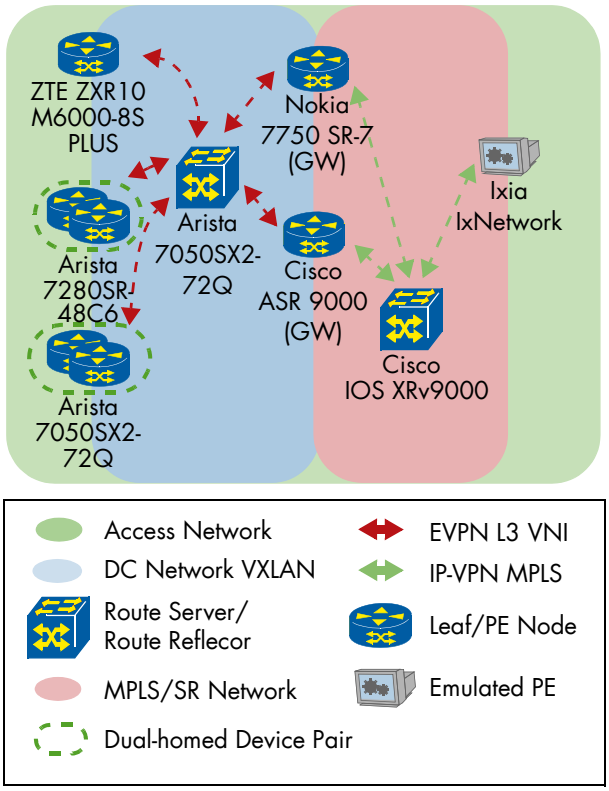


Figure 23: EVPN and IP-VPN Interworking - Setup 1

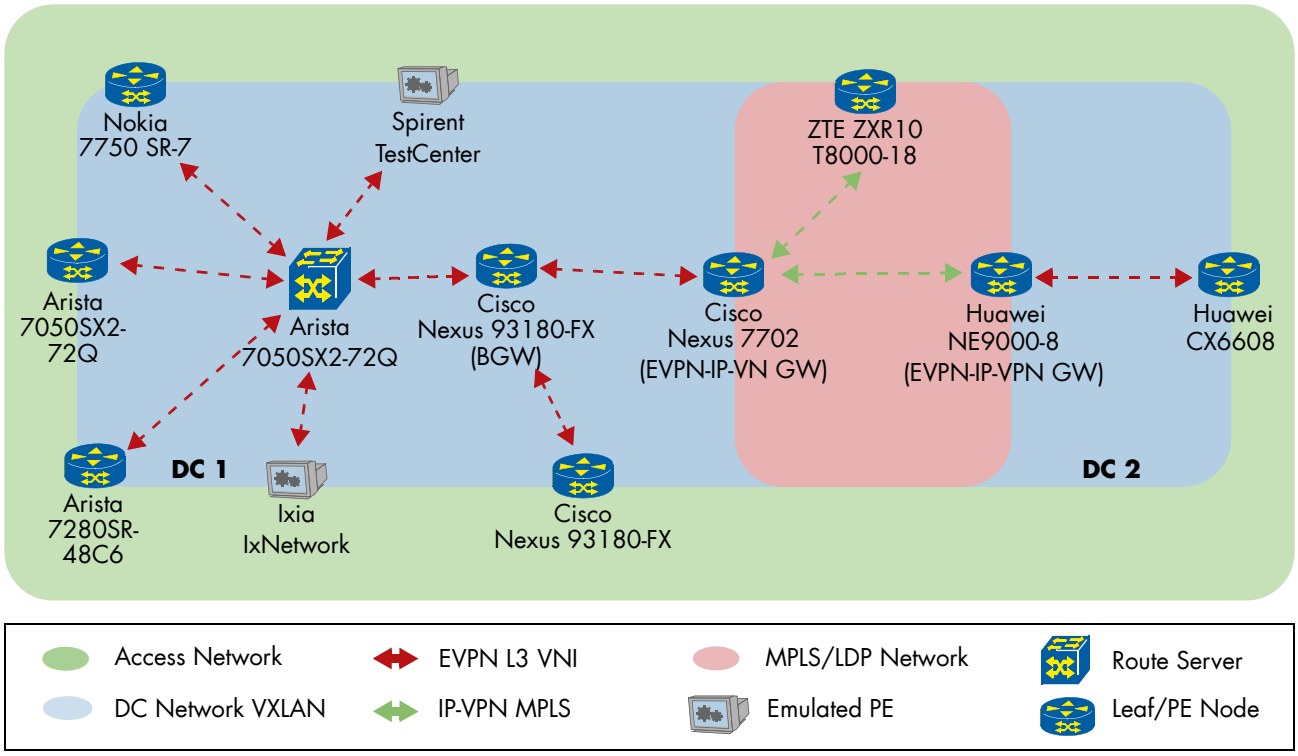


Figure 24: EVPN and IP-VPN Interworking - Setup 2

One of the most important use cases of EVPN is interconnection of data centers across an IP/MPLS core. The goal of this test is to verify an interworking use case between a WAN network that is based on IP-VPN and data center networks that are based on EVPN.

The system must provide control plane and data plane interworking between the EVPN network and the IPVPN technology supported.

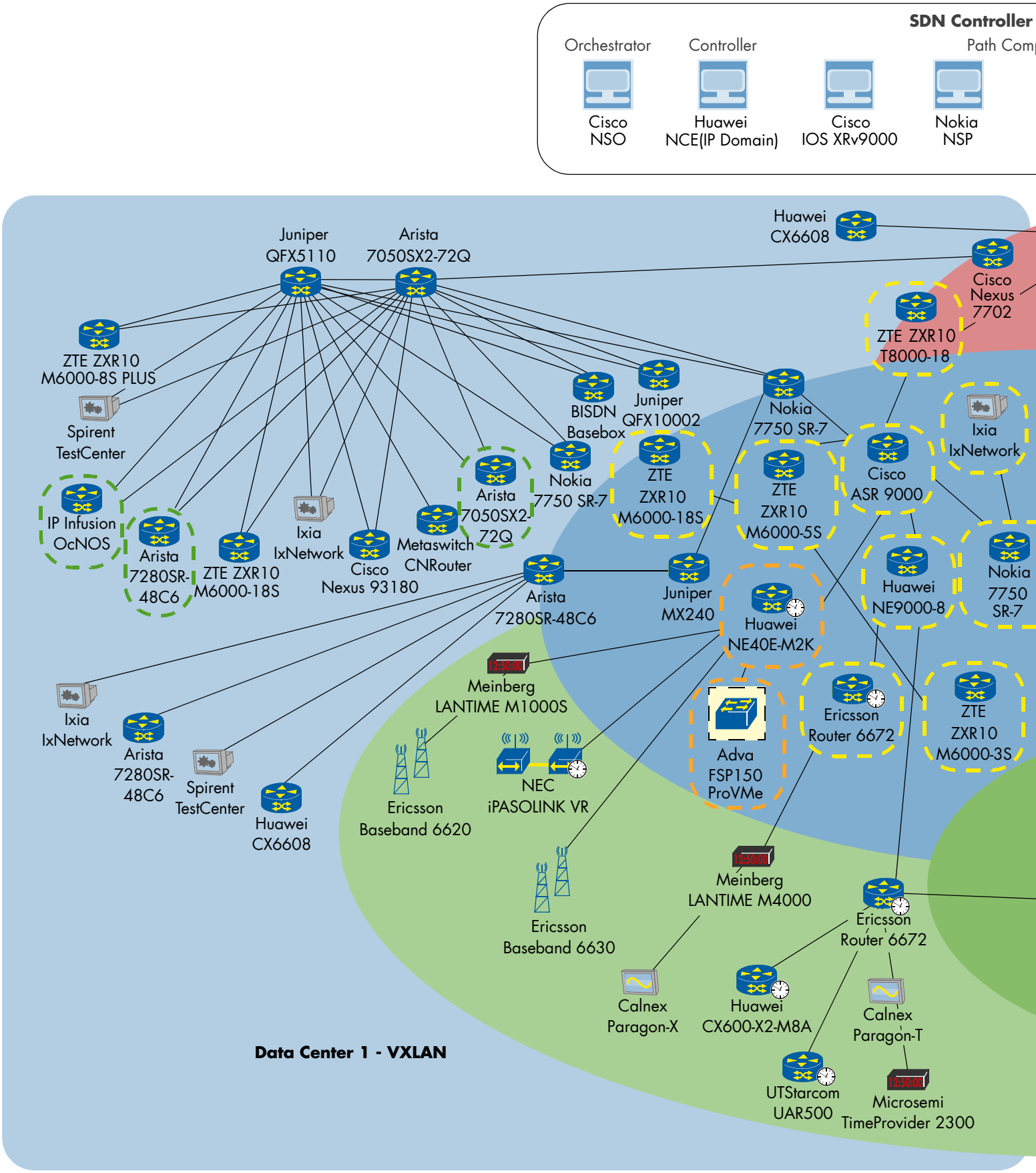
For setup 1 (Figure 23), we tested the following vendors:

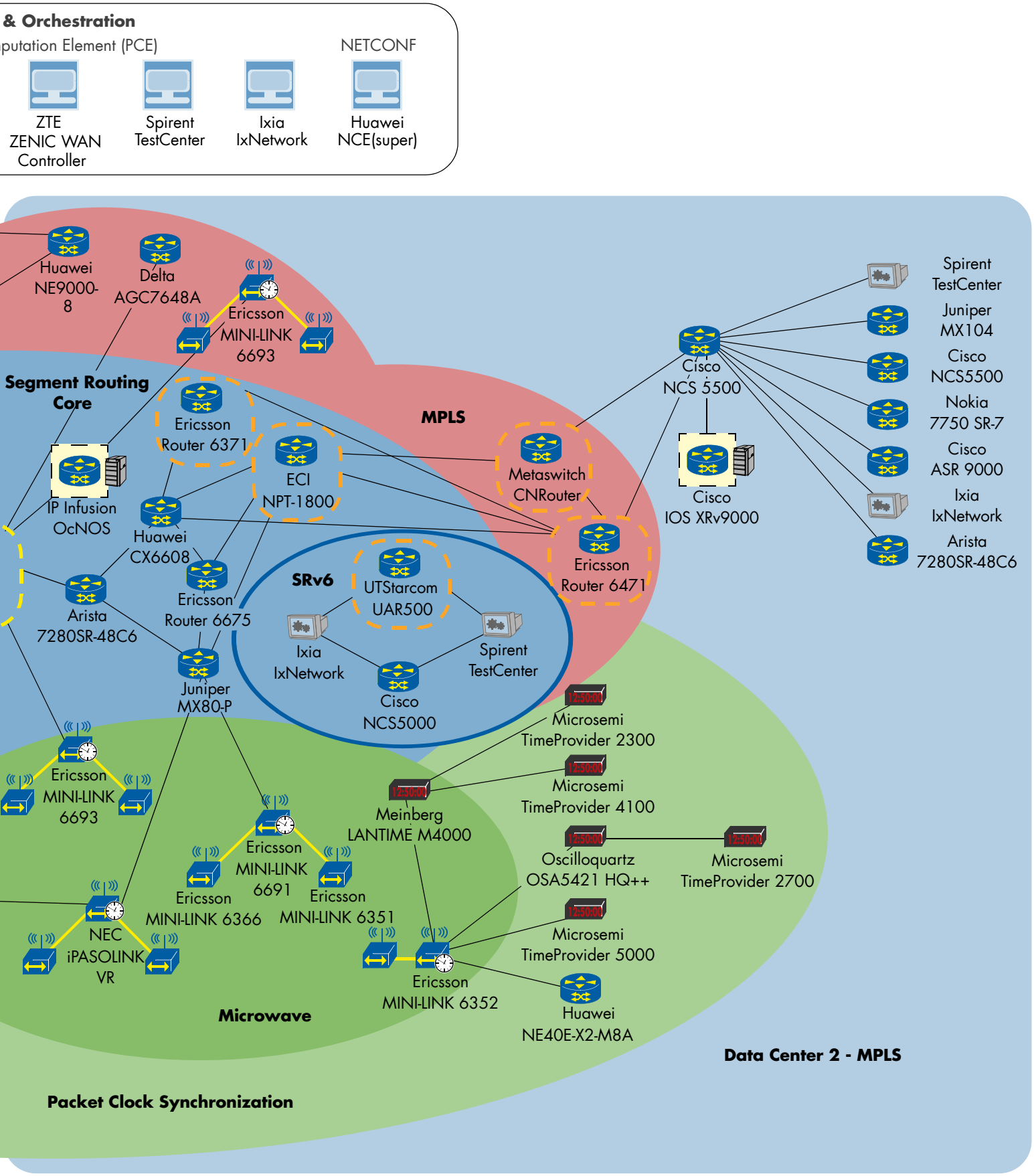
- EVPN-VXLAN PE Role: Arista 7050SX2-72Q, Arista 7280SR-48C6, ZTE ZXR10 M6000-8S PLUS
- IP-VPN/MPLS PE Role: Ixia IxNetwork
- IP-VPN - EVPN Gateway Role: Cisco ASR 9000, Nokia 7750 SR-7
- BGP Route Reflector Role: Cisco IOS XRv9000
- BGP Router Server/Spine Role: Arista 7050SX2-72Q

Setup 2 was performed to allow vendors to participated performing a different network function or demonstrate the same capabilities in different product lines.

For setup 2 (Figure 24), we tested the following vendors:

- EVPN-VXLAN PE Role: Arista 7050SX2-72Q, Arista 7280SR-48C6, Cisco Nexus 93180-FX, Ixia IxNetwork, Huawei CX6608, Nokia 7750 SR-7, Spirent TestCenter
- IP-VPN/MPLS PE Role: ZTE ZXR10 T8000-18
- IP-VPN - EVPN Gateway Role: Cisco Nexus 7702, Huawei NE9000-8
- Border Gateway Role: Cisco Nexus 93180-FX
- BGP Route Server Role: Arista 7050SX2-72Q


SDN Controller
Orchestrator
Controller
Path Com
Cisco NSO
Huawei NCE(IP Domain)
Cisco IOS XRv9000
Nokia NSP
Juniper QFX5110
Arista 7050SX2-72Q
Huawei CX6608
Cisco Nexus 7702
ZTE ZXR10 T8000-18
ZTE ZXR10 M6000-8S PLUS
Spirent TestCenter
BISDN Basebox
Juniper QFX10002
Nokia 7750 SR-7
Ixia IxNetwork
IP Infusion OcNOS
Arista 7280SR-48C6
ZTE ZXR10 M6000-18S
Ixia IxNetwork
Cisco Nexus 93180
Metaswitch CNRouter
Arista 7050SX2-72Q
Nokia 7750 SR-7
ZTE ZXR10 M6000-18S
ZTE ZXR10 M6000-5S
Cisco ASR 9000
Arista 7280SR-48C6
Juniper MX240
Huawei NE40E-M2K
Huawei NE9000-8
Nokia 7750 SR-7
Meinberg LANTIME M1000S
Ixia IxNetwork
Arista 7280SR-48C6
Spirent TestCenter
Huawei CX6608
Ericsson Baseband 6620
NEC iPASOLINK VR
Adva FSP150 ProVMe
Ericsson Router 6672
ZTE ZXR10 M6000-3S
Ericsson Baseband 6630
Meinberg LANTIME M4000
Ericsson Router 6672
Calnex Paragon-X
Huawei CX600-X2-M8A
Calnex Paragon-T
UTStarcom UAR500
Microsemi TimeProvider 2300
Data Center 1 - VXLAN
Segment Routing
Data Center
Physical Link
Microwave Link
Clocking
Microwave
MPLS/LDP
SDN Controller & Orchestration
Core Router
Access Device
Route Reflector/ Route Server


& Orchestration
putation Element (PCE)
NETCONF
ZTE ZENIC WAN Controller
Spirent TestCenter
Ixia IxNetwork
Huawei NCE(super)
Huawei NE9000-8
Delta AGC7648A
Ericsson MINI-LINK 6693
Segment Routing Core
Ericsson Router 6371
ECI NPT-1800
MPLS
Metaswitch CNRouter
Cisco NCS 5500
Cisco IOS XRv9000
Spirent TestCenter
Juniper MX104
Cisco NCS5500
Nokia 7750 SR-7
Cisco ASR 9000
Ixia IxNetwork
Arista 7280SR-48C6
IP Infusion OcNOS
Huawei CX6608
Ericsson Router 6471
SRv6
UTStarcom UAR500
Ixia IxNetwork
Spirent TestCenter
Cisco NCS5000
Arista 7280SR-48C6
Ericsson Router 6675
Juniper MX80-P
Microsemi TimeProvider 2300
Microsemi TimeProvider 4100
Meinberg LANTIME M4000
Oscilloquartz OSA5421 HQ++
Microsemi TimeProvider 2700
Microsemi TimeProvider 5000
Huawei NE40E-X2-M8A
Ericsson MINI-LINK 6693
Ericsson MINI-LINK 6691
Ericsson MINI-LINK 6366
Ericsson MINI-LINK 6351
Ericsson MINI-LINK 6352
NEC iPASOLINK VR
Microwave
Data Center 2 - MPLS
Packet Clock Synchronization
Emulator
Synchronous Node
Controller
Devices Managed by PCE
Phase/Frequency Analyzer
Clock Node
Data Center Server
Devices Managed by NETCONF
Synchronous Microwave Node
Microwave Node
Virtual Network Function
Dual-homed Leaf

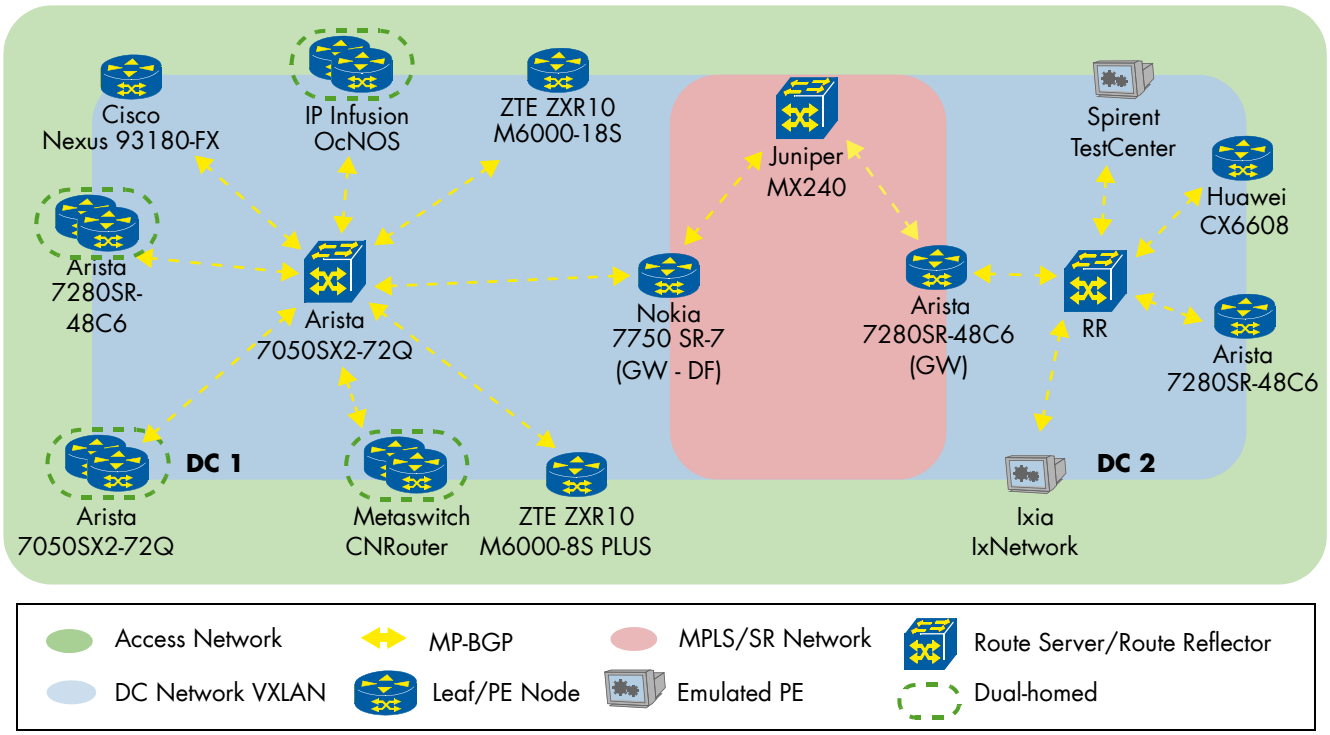


Figure 25: EVPN VXLAN - SR-MPLS Interworking

EVPN-VXLAN and EVPN-MPLS Interworking

In data center networks, it is more common to run a pure IP network without MPLS support. In those cases, VXLAN is one of the data plane options for multi-tenancy. The goal of this test is to verify an interworking use case between a WAN network that is based on EVPN-MPLS and data center networks that are based on EVPN-VXLAN.

This test focuses on "Integrated Interconnect solution" which means the NVO Gateway (GW) and the WAN Edge functions are integrated into the same system. This system must provide control plane and data plane interworking between the EVPN-VXLAN network and the EVPN-MPLS technology supported in the WAN.

This year we had the opportunity to test the MPLS section exclusively with Segment Routing, using ISIS SR extensions.

As shown in Figure 25, the devices in the data center network provided EVPN-VXLAN connections. The IP/MPLS edge routers provided the interconnection between the EVPN-VXLAN network and EVPN-MPLS for the control plane and data plane interoperability.

Once we tested the BGP sessions in the underlay and overlay, we checked the BGP routing table on each of the IP/MPLS edge devices (PE). Each of the PE devices received route type-3 from each remote EVPN PE.

We generated unicast Ethernet traffic between the sites and started to verify the MAC/IP advertisement (route type-2) routes in each of the network nodes. The MAC/IP advertisement routes in the IP/MPLS network segment were carrying the RD of common EVI, the MAC addresses, and the MPLS label associated with the MAC.

The MAC/IP advertisement routes in the VXLAN network segments were carrying the RD of common EVI, the MAC addresses, the VNI associated with the MAC, and VXLAN encapsulation as extended community.

In this scenario, we tested the following vendors:

- EVPN-VXLAN PE Role: Arista 7050SX2-72Q (dual-homed), Arista 7280SR-48C6 (dual-homed), Cisco Nexus 93180-FX, Huawei CX6608, Ixia IxNetwork, Spirent TestCenter, ZTE ZXR10 M6000-18S, ZTE ZXR10 M6000-8S PLUS, Metaswitch CNRouter (dual-homed), IP Infusion OcNOS (AS7712-32X) (dual-homed),
- VXLAN-MPLS Gateway Role: Arista 7050SX2-72Q, Nokia 7750 SR-7
- Route-reflector/server Role: Arista 7050SX2-72Q, Juniper MX240

Software Defined Networking

Service providers aim to increase the agility of their networks. Adopting centralized network management protocols and application-specific service orchestration architecture can be instrumental to help service providers achieve this goal. The following two sections describe our Path Computation Element Protocol (PCEP) and NETCONF/YANG interoperability test descriptions, results and overall interoperability findings.

Path Computation Element Protocol

PCEP is defined by the IETF in RFC 5440 as a mechanism to communicate between a Path Computation Client (PCC) and a Path Computation Element (PCE). PCEP sessions run over TCP and enable PCC and PCE nodes to exchange path computation requests, responses, session status and reports. Traffic engineering paths are computed by the PCE and pushed towards the PCC. Both the PCE and PCC can trigger this computation and provide the path requirements. PCEP can also be used to re-optimize and update existing paths.

In order to demonstrate all interoperable PCEP combinations, we designed the tests with a unidirectional LSP per combination. This meant that a single PCC head-end was sufficient to showcase PCC-to-PCE interoperability.

This year, we noticed a growing interest in PCEP's usage to govern SR-TE paths. Most participating vendors favored testing PCEP with SR-TE over RSVP-TE, which we tested thoroughly in 2017.

PCE-initiated SR-TE Paths in a Stateful PCE Model

In environments where the LSP placement needs to change in response to application demands, it is useful to support the dynamic creation and tear down of LSPs.

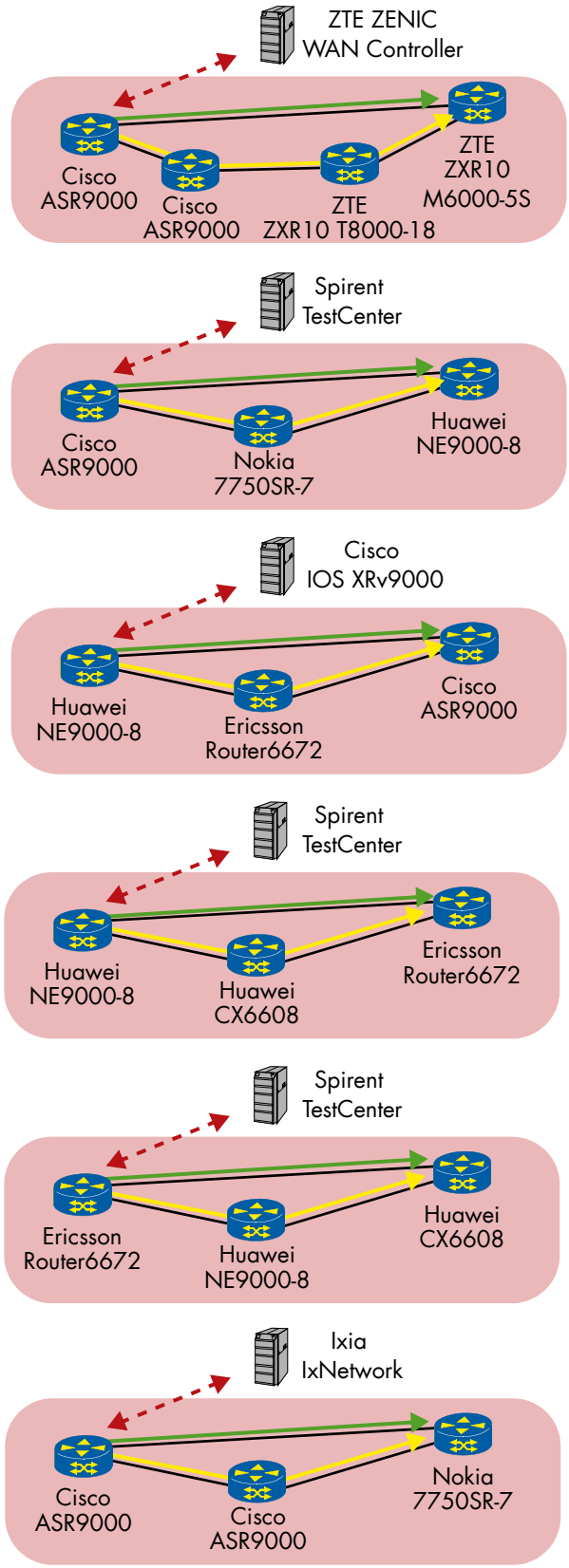
In this test, we verified the creation of an SR-TE path within a single IGP domain when initiated by the PCE. We also verified the state synchronization and deletion of LSPs.

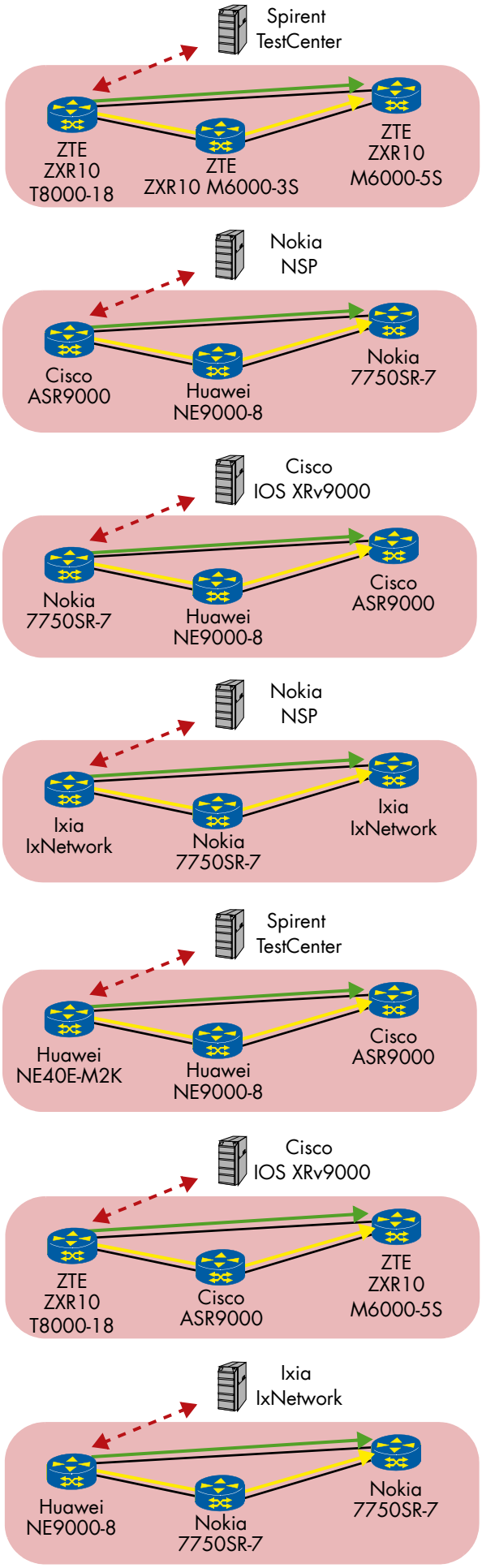
The test topology included three network nodes, one of which acted as PCC. Participating vendors chose IS-IS-TE to synchronize the TED information. The LSP was restricted to a suboptimal path to ensure that our test traffic did not follow the IGP shortest path.

We started the test by verifying the stateful PCEP session state and TED information on the network nodes. After initiating the LSP from the PCE, we checked the LSP database on the PCE and ensured a single transport path entry on each PCC.

In order to verify state synchronization, we asked the participating vendors to terminate the PCEP session (by the PCE or PCC), clear the LSP database on the PCE, then re-establish the PCEP session. We verified that existing LSPs on the PCC were synchronized with the PCE after the session's restoration and that test traffic was not affected by the PCEP session interruption.

Finally, the PCE deleted the LSP and we verified that the LSP information on the PCC were deleted. Following the LSP's deletion, test traffic followed the IGP shortest path as expected. Figure 26 depicts successful participant combinations.





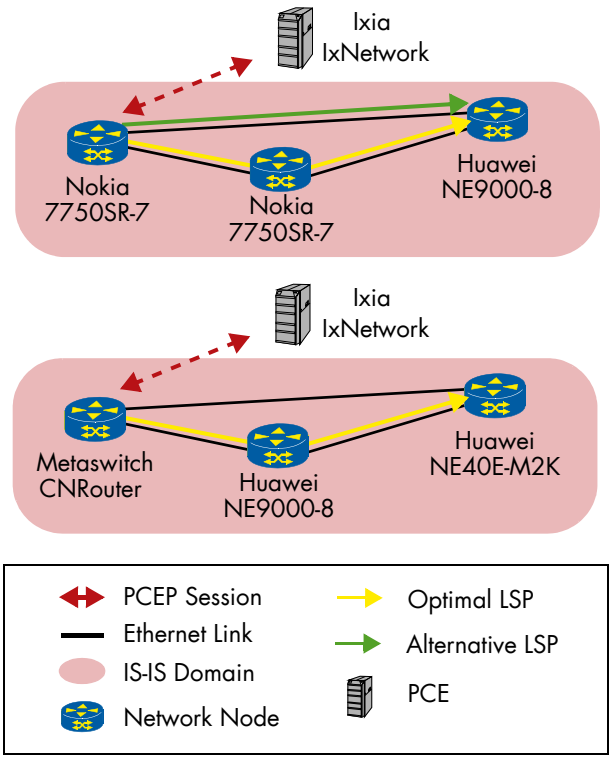


Figure 26: PCE-initiated SR-TE Path Creation and Deletion

It is worth mentioning that current vendor support for bandwidth constraints was limited, hence we excluded that requirement from the test. Also one vendor implementation reported an abnormal LSP state synchronization process that triggered the deletion of the LSP after the PCEP session had been restored. Some test combinations were prevented by a Maximum SID Depth (MSD) mismatch.

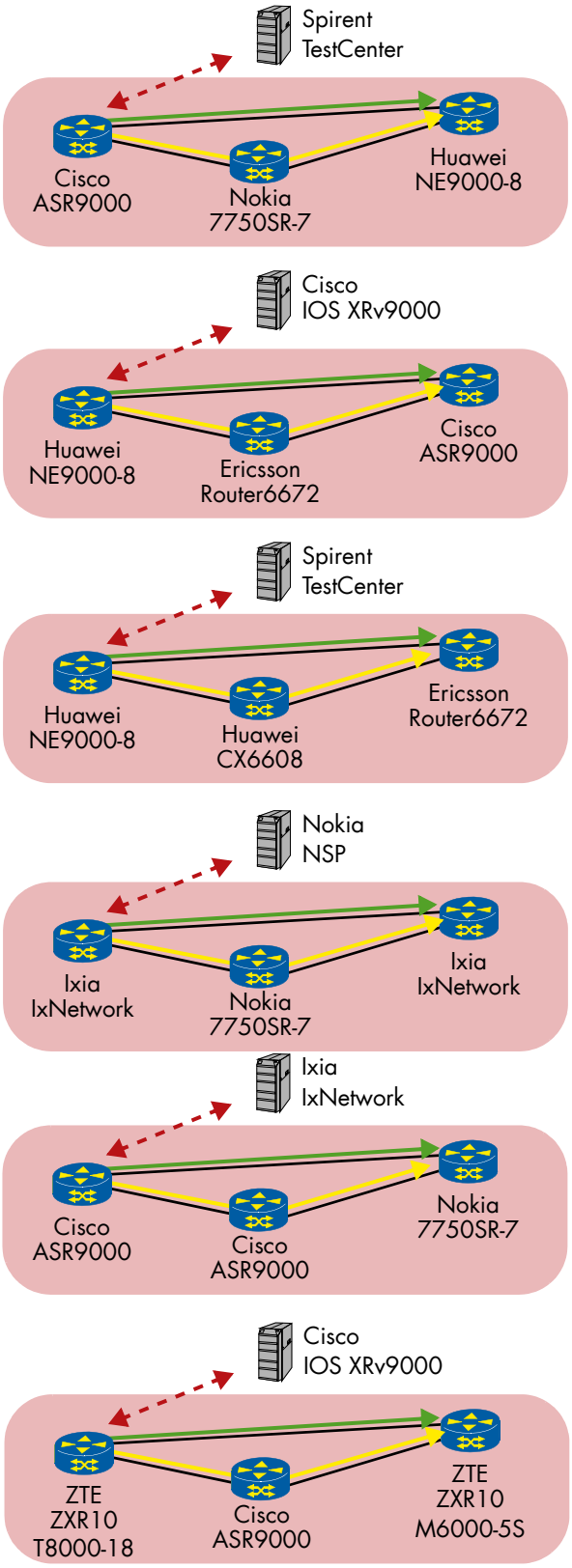
Finally, we observed that a single vendor implemented the TED state synchronization differently which created interoperability challenges and generated error reports during the LSP database re-synchronization process.

PCC-initiated SR-TE Paths in a Stateful PCE Model

According to definition, once a PCC has successfully established a PCEP session with a selected PCE node, it can send a path computation request to that PCE (PCReq message) containing the path attributes and constraints. After receiving the path request, the PCE pushes the computed LSP down towards the requesting PCC. This can be helpful for applications that demand certain network conditions to operate.

In this test, we verified the SR-TE LSP initiation and creation process by checking the TED and LSP information on the PCEP peers and sending test traffic over the created tunnel. The traffic was generated from the PCC side towards a secondary network node via a suboptimal path – defying the IGP shortest path route. Similar to the PCE-initiated test, vendors chose IS-IS-TE to synchronize the TED information.

After the creation of the PCC-initiated LSP, we verified the delegation of the LSP from the PCC to the PCE by checking the LSP's "Delegate" flag on the PCEP peers. Finally, we tested the LSP termination and confirmed that the TED and LSP database were cleared. After the termination, as expected our test traffic followed the IGP shortest path – as opposed to the suboptimal path. Figure 27 depicts successful combinations.



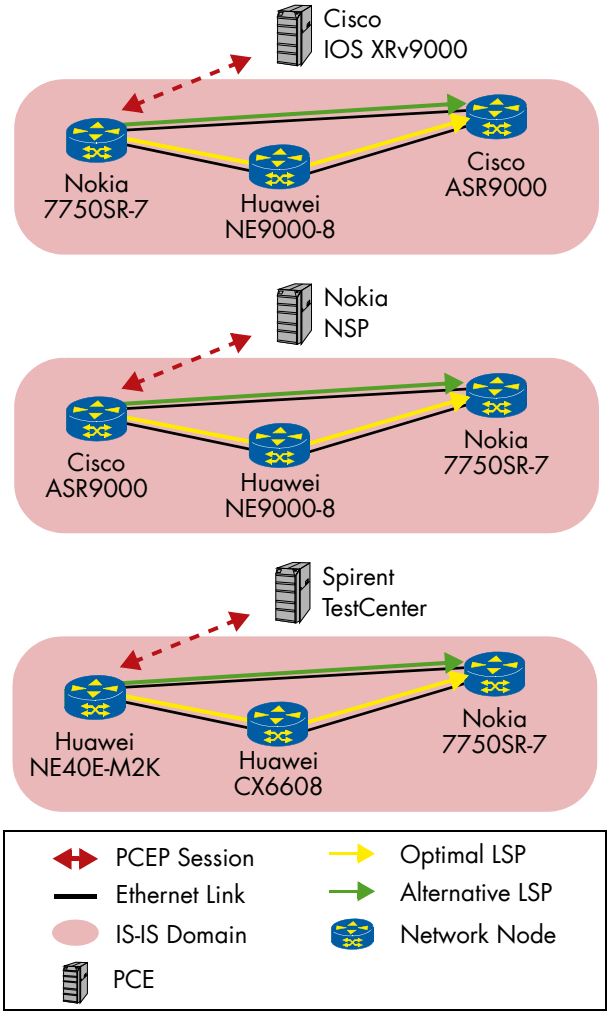


Figure 27: PCC-initiated SR-TE Path Creation and Deletion

We noticed a few differences in implementing PCC initiation between vendors. Some vendors used Path Computation Report – with an empty Explicit Route Object (ERO) versus using PCReq to initiate the LSP by the PCC. Also some implementations delegated the LSP to the PCE by default or left no option for revoking the delegation.

SR-TE Path Update and Re-optimization in a PCEP Network

The SDN architecture centralizes network management decisions. The PCE nodes should be able to make localized and holistic service updates according to a variety of conditions.

In this test we verified the ability of the PCEP peers to compute and install re-optimized paths when the state of the network changes. Given the wide range of possible triggers, and differences in supporting those triggers by vendors, we asked participants to use one of three options to trigger path recomputation:

- Increase a link cost on the original path
- Interrupt a link on the primary path
- Manually trigger path updates on the PCE in case of PCE emulation solutions (i.e. Ixia IxNetwork, Spirent TestCenter)

Since this test shares most of the preliminary steps with PCE-initiated SR-TE Paths in a Stateful PCE Model 23 and PCC-initiated SR-TE Paths in a Stateful PCE Model 24, we ran the LSP update process as an intermediate step within those tests whenever possible. With the exception of a single test combination, all vendor combinations that are listed in Figure 26 and Figure 27 also took part in this test. In those combinations, we reused the test topology where the updated LSPs took the direct path instead of suboptimal one as highlighted in the figures. Other successful vendor combinations that participated in this test but were not previously listed are depicted in Figure 28.

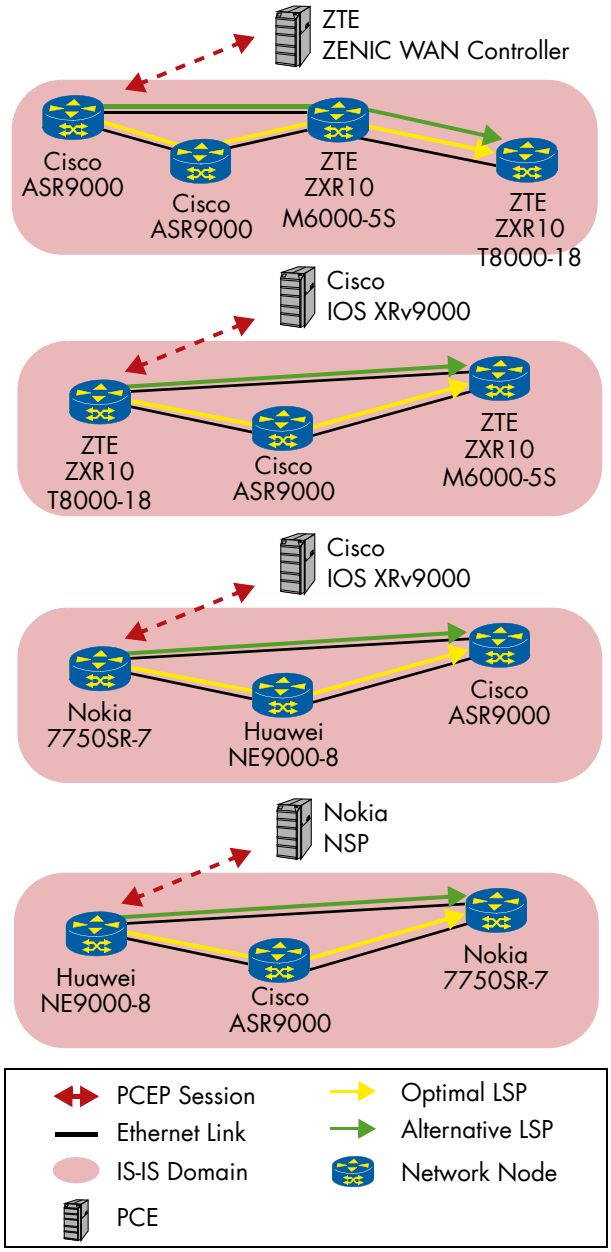


Figure 28: Additional SR-TE Path Update and Re-optimization

Inter-Domain Segment Routing Traffic Engineering - BGP-LS and PCE Integration

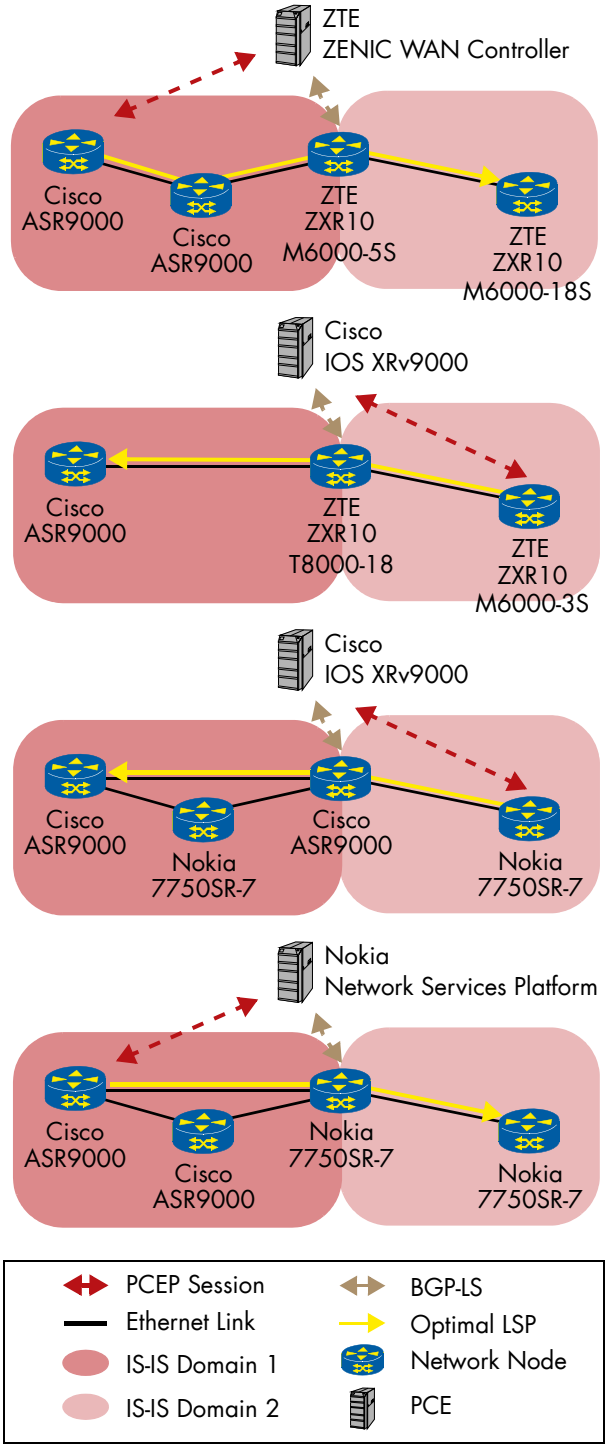


Figure 29: BGP-LS and PCE Integration

The visibility of topology, Traffic Engineering Database (TED) and Link State Database (LSDB) in an inter-domain network remains local to each domain. To create a TE LSP that transits through two or more network autonomous systems, a PCE can utilize BGP Link State (BGP-LS).

In this test, we asked participating vendors to set up an optimal inter-IGP-domain LSP. We verified PCEP sessions between the PCC and PCE, and that the PCE established BGP-LS with a multi-domain router.

We also confirmed that the PCE had the full network topology for both domains. Afterwards, we asked the PCE to create an end-to-end LSP across both network domains and verified the LSP on the PCC nodes. The test topology and vendor roles are depicted in Figure 29.

Cisco IOS XRv9000, Nokia Network Services Platform and ZTE ZENIC WAN Controller took the PCE role while Cisco-ASR 9000, Nokia 7750 SR-7 and ZTE ZXR10 M6000-3S participated as PCC. The topology differed slightly due to the number of involved vendors and physical ports availability on the network nodes.

Inter-AS Segment Routing Traffic Engineering

Another use case of PCEP is to program SR-TE paths across autonomous systems. This can be a valuable tool when application-specific traffic engineering can make use of the AS topology and signal the most optimal path towards the PCC head-end.

In this test, the PCE learned the multi-domain topology via BGP-LS. In addition, the inter-AS link was modeled using SR Egress Peering Engineering (EPE) SIDs. The PCE then used this information to compute the optimal path and push it towards the PCC head-end. After we verified the creation of the inter-AS LSP we shut down one link on the path in order to trigger the recomputation of the path. An alternative path was computed and pushed to the PCC successfully. We verified the label stack on the PCC and asked the PCE vendor to terminate the LSP. The termination was successful and the LSP information was cleared on the PCEP peers. ZTE ZXR10 T8000-18 participated as PCC head-end, while Cisco IOS XRv9000 took the role of PCE. The full topology and the LSPs are displayed in Figure 30.

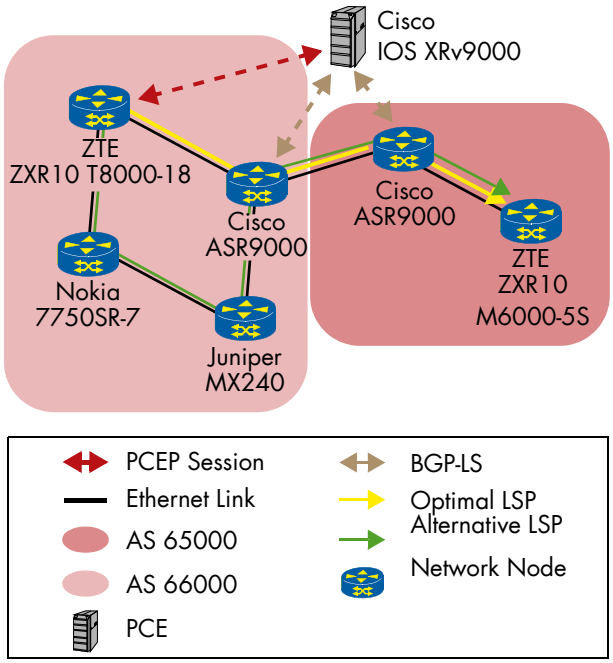


Figure 30: Inter-AS SR-TE

NETCONF/YANG

Many service providers are investigating zero-touch networks to reduce their service orchestration and operations overhead. Part of the convergence process is to successfully define and implement a multi-vendor network service.

The combination of network configuration protocols – such as NETCONF and RESTCONF – with YANG modeling language are potential instruments. Standardization bodies attempt to define service catalogues that can be used by different vendors to define the same network service, such as Layer 3 and Layer 2 VPNs. The following two tests cover both use cases.

L3VPN Service Creation and Termination

In this test, we defined the parameters for a L3VPN service. IETF RFC 8299 defines this service model in YANG. The test expected that a controller will translate the service parameters from YANG to vendor-specific NETCONF/YANG configuration parameters on the network nodes.

The service models specified the interface parameters, virtual routing and forwarding (VRF) instance and CE-PE routes. The service was created and terminated using the orchestrator's northbound interface. Vendors chose MPLS or SRv6 as their preferred transport data plane.

To verify the service creation and deletion we checked the running configuration on the network nodes. Following the service creation, we sent traffic through the network expected no traffic loss. We also confirmed that the service orchestration was non-intrusive by comparing the configuration on network nodes before and after the test. The device configurations matched with the exception of one case where the device configuration had extra unreadable characters. The responsible vendor explained that the added characters do not interfere with the device operations.

Six vendors successfully participated in this test. Both, Cisco NSO and Huawei NCE acted as NETCONF/YANG orchestrators. ECI NPT-1800, Ericsson Router 6471, Metaswitch CNRouter and UTStarcom UAR500 acted as provider edge. Cisco NSO exposed its northbound interface using NETCONF/YANG, while Huawei NCE exposed its northbound API using RESTCONF/YANG.

Successful combinations are depicted in Figure 31.

Finally, it is worth mentioning that the YANG data model for L3VPN over SRv6 data plane has not yet been defined in the standard by IETF (indicated as work in progress in draft-raza-spring-srv6-yang-01), so NETCONF behavior for setting up L3VPN service over SRv6 data plane had proprietary implementation at this stage with provisioning performed by a controller in several steps (transactions).

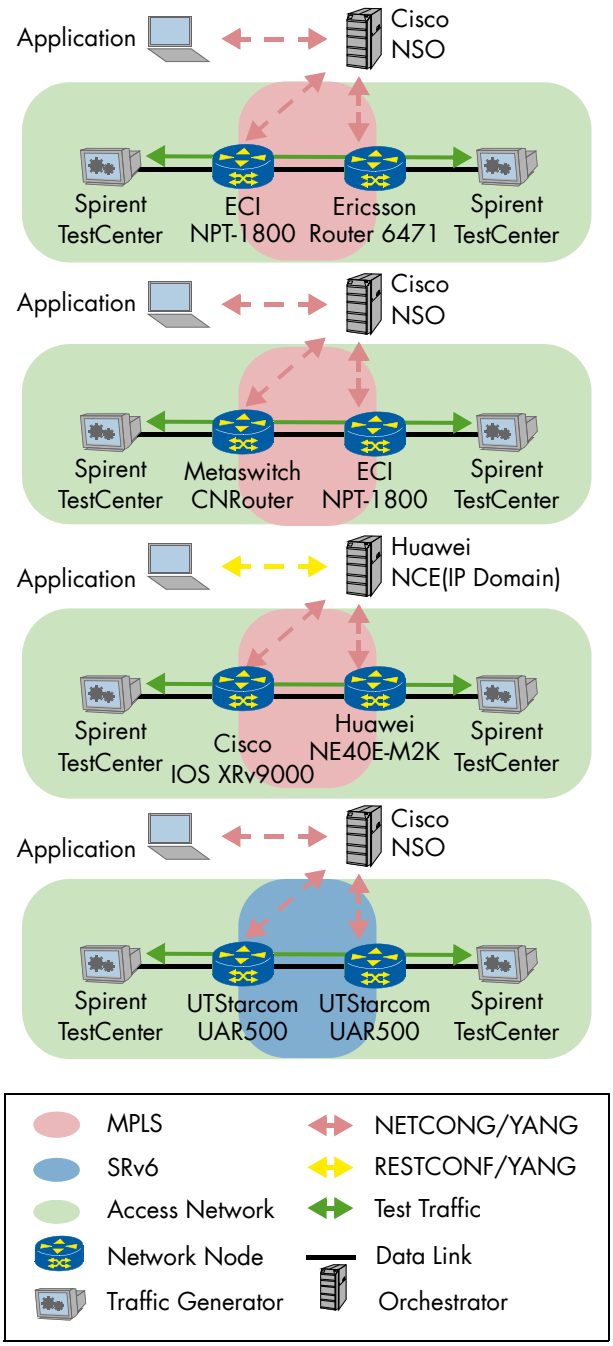


Figure 31: L3VPN Service Creation and Termination

L2VPN Service Creation and Termination

Another use case of NETCONF/YANG is L2VPN service modeling and deployment. This test setup was similar to the previous L3VPN one with a slight difference in the service parameters – such as the Service VLAN. In this test, the service was directly modelled on the orchestrator due to the lack support for draft YANG models such as the draft-ietf-l2sm-l2vpn-service-model by the IETF.

The orchestrator initiated and terminated the L2VPN service on the PEs successfully. We verified the network service by sending traffic between the two customer sites. Two vendors took part in this test: Cisco NSO acted as the orchestrator while UTStarcom UAR500 acted as PE. SRv6 was chosen by UTStarcom as the transport data plane.

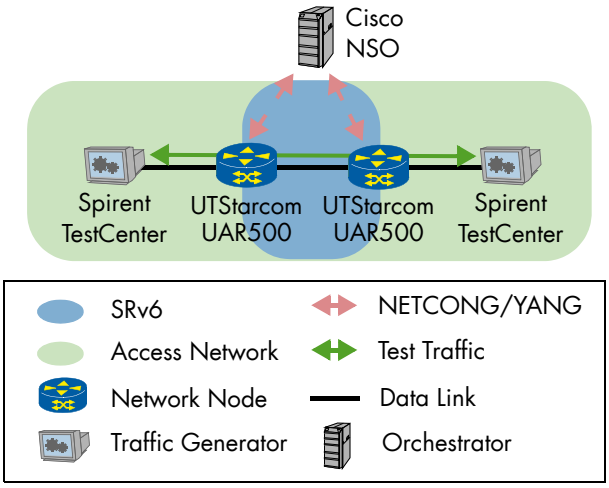


Figure 32: L2VPN Service Creation and Termination

Multi-Vendor/Multi-Domain Controllers Orchestration

The SDN architecture can span multiple network domains. Typically, each domain is controlled by a domain controller. Inter-domain services can be orchestrated by a multi-domain controller that has a full view of its subdomains.

In this test, we experimented with three management protocols: NETCONF, RESTCONF and PCEP. The three combined managed two network domains. The domains were connected directly via two provider edges. We asked the vendors to create and delete an end-to-end L3VPN service modeled in RFC 8299.

Four vendors took part in this test. Cisco NSO managed two Ericsson provider edges via NETCONF/YANG. The other domain contained two Huawei provider edges managed using PCEP by Huawei's NCE(IP domain). A single Huawei NCE(Super) instance managed the domain controllers using NETCONF/YANG and RESTCONF/YANG.

Figure 33 describes the test topology and connectivity between the participating components.

At the beginning of this test we checked management sessions and configuration information on the network nodes and controllers. We then asked the multi-domain controller to trigger the service creation and monitored the north-bound interface on both domain controllers. The multi-domain controller re-modelled the service to two different services using RFC 8299 YANG and pushed them into the two domain controllers. Each domain controller translated the service into device configuration and pushed them towards the network nodes. To verify the service creation, we checked the device configuration and sent traffic via the service path. The multi-domain controller deleted the service successfully and the devices configuration was cleared successfully.

Test traffic was dropped after the service's deletion as expected.

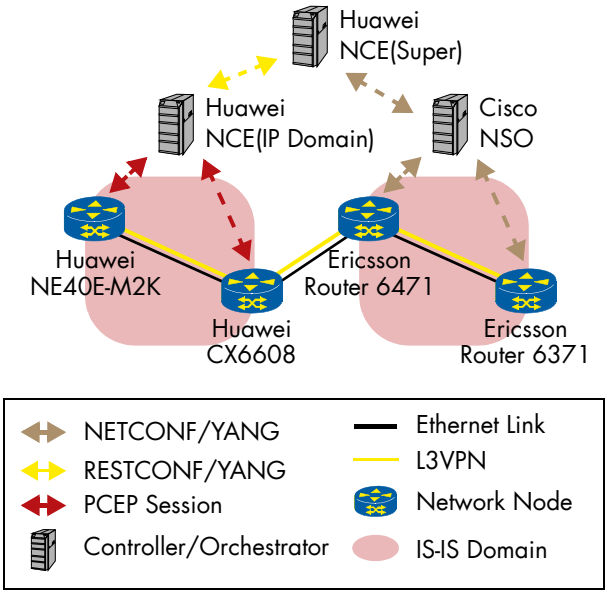


Figure 33: Multi-Vendor/Multi-Domain Controller Orchestration

Virtual CPE Integration with NETCONF

NETCONF can be extended to manage virtualized resources such as the virtual customer premises equipment (vCPE).

Adva took part in this test with two FSP150 ProVMe devices, connected back-to-back, to illustrate the WAN integration of vCPE components. On each of the two Adva units, a virtual router from a third party vendor was instantiated and configured manually.

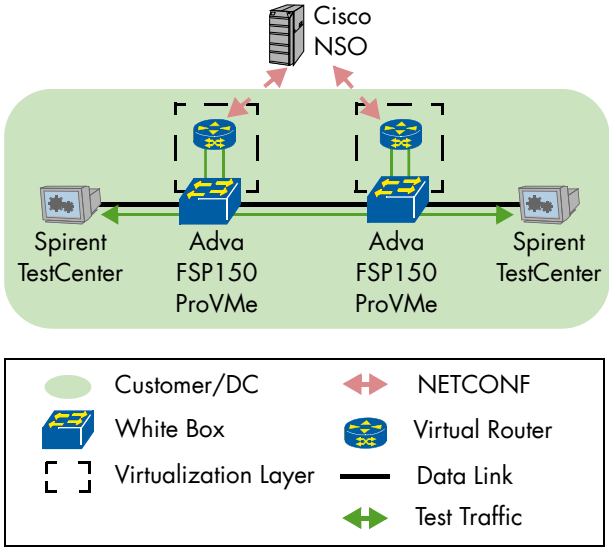


Figure 34: Virtual CPE Integration with NETCONF

The successful configuration was verified by sending user data across the two hybrid CPEs and their virtual routers using Spirent TestCenter. Subsequently, NETCONF sessions were established between each vCPE virtual router and the Cisco NSO. Due to lack of time for test execution, further application of NETCONF/YANG was not covered.

Microwave

As operators upgrade their networks with LTE-A in preparation for a 5G future, questions are being asked around the role Microwave transport will play. We see a trend of integrating the more specialised microwave devices into the standard IP/MPLS router domain, and thus looked into two particular aspects of this in the following tests.

Bandwidth Notification

Today Mobile Backhaul networks are often built as an overlay with routers sitting on top of microwave devices. In the past there was limited communication between these two domains, but with the bandwidth notification messages (ETH-BN) defined by ITU-T Y.1731, it is now possible for the microwave systems to signal a change in bandwidth to the routers.
This enables a router to apply service policies to the traffic it sends on to the microwave system based on the bandwidth information within the ETH-BN packets.
At the beginning of this test, the Microwave nodes were using the maximum modulation possible, as depicted in Table and sent end-to-end traffic. In the next step we emulated severe weather conditions in the link between the microwave nodes by using a RF attenuator and verified that the aggregation router could process the bandwidth notification messages (ETH-BN) and accordingly apply service policies to the traffic sent to the microwave system, based on the bandwidth information within the ETH-BN.

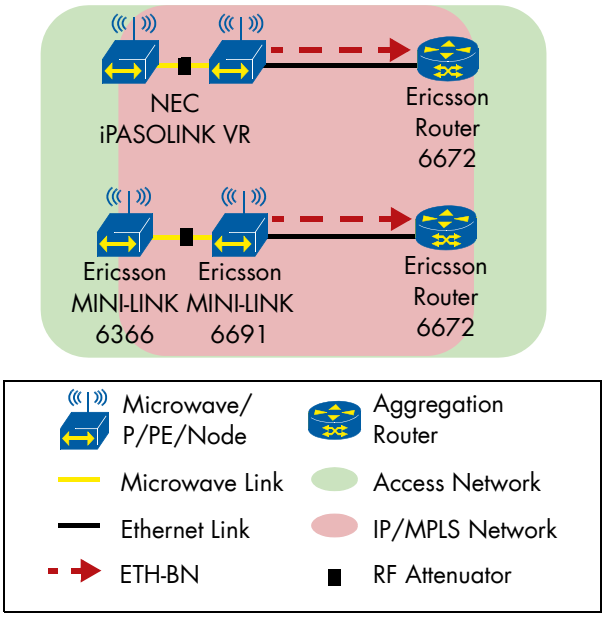


Figure 35: Bandwidth Notification

We successfully tested the following combinations: Ericsson Router 6672 acted as the aggregation router in both combinations. The microwave link was established between two NEC iPASOLINK VR and an Ericsson MINI-LINK 6366 and MINI-LINK 6691.

Participants	Modulation Scheme	Channel Spacing
Ericsson MINI-LINK 6691	4096 QAM	112 MHz
Ericsson MINI-LINK 6366	4096 QAM	112 MHz
Ericsson MINI-LINK 6651	4096 QAM	112 MHz
NEC iPASOLINK-VR	2048 QAM	56 MHz

Table 2: Modulation and Channel Spacing Used

Layer 3 Microwave MPLS-based Services

The aim of this test was to confirm the capability to establish IP/MPLS service on a microwave platform crossing or terminating on existing infrastructure.
We tested two different combinations relying on different transport profiles, and verified that a L3VPN service can be set up between IP/MPLS capable microwave systems and IP/MPLS aggregation routers in multi-vendor scenario.
In the first scenario we used IS-IS as the IGP protocol and LDP for the MPLS label allocation/distribution. In the second we changed the IGP to OSPF with LDP.
We created both end-to-end services between two different microwave vendors with standalone routers participating as aggregation router, as well as directly between two microwave vendors.
In the tests, Ericsson MINI-LINK 6366, Ericsson MINI-LINK 6691 and NEC iPASOLINK VR microwave nodes acted as PE/P nodes. Juniper MX80 participated as P node in the combination which was using L3VPN with ISIS and LDP. In another combination with only Ericsson MINI-LINK 6691 and NEC iPASOLINK VR, we tested the L3VPN with OSPF and LDP.

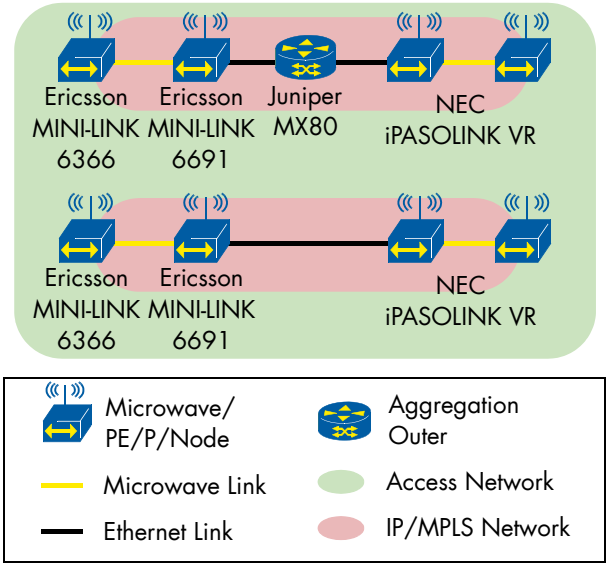


Figure 36: Layer 3 Microwave MPLS Based Services

Layer 3 Microwave Transport Resiliency

Bringing IP/MPLS to the microwave network provides additional resiliency options in the access network and increases the end-to-end service availability.

The goal of this test was to show that a microwave node can react to degradation of the radio link by re-routing the traffic via a different path.

We used Spirent TestCenter to act as CE and send bidirectional traffic across the network. We verified that the microwave nodes were using the main path with the maximum modulation scheme available as depicted in Table and that no packets were lost. We then emulated severe weather conditions by reducing the available bandwidth of the channel with a RF attenuator.

In this test, Ericsson MINI-LINK 6691 and NEC iPASOLINK VR acted as P nodes, Ericsson MINI-LINK 6366 and NEC iPASOLINK VR acted as the microwave stations and PE nodes. Ericsson MINI-LINK 6651 and NEC iPASOLINK VR participated as resiliency nodes. Juniper MX80 acted as PE router.

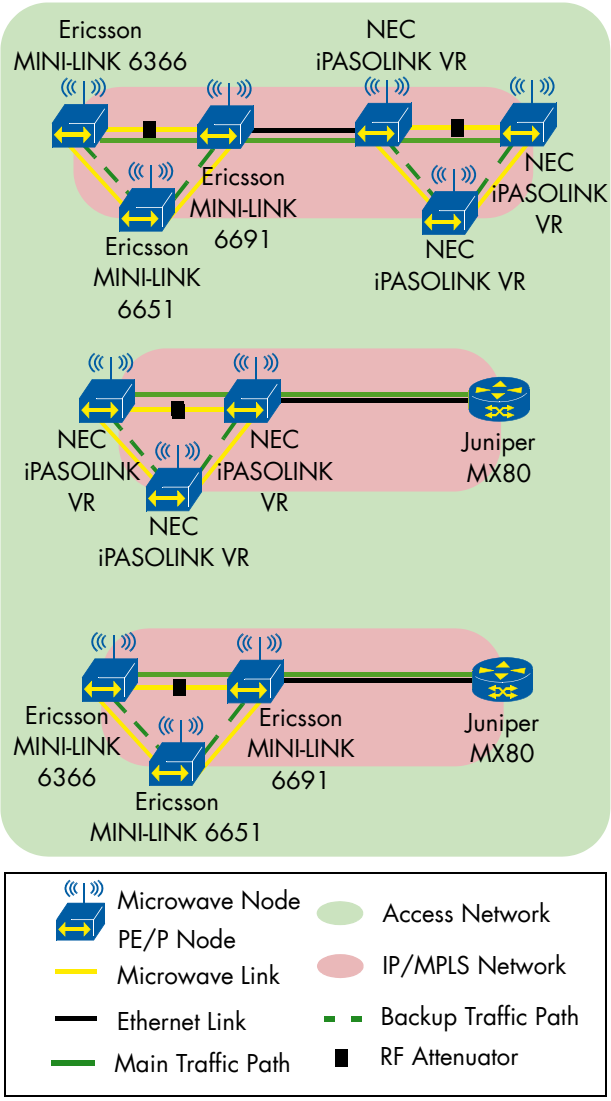


Figure 37: Layer 3 Microwave Transport Resiliency

Clock Synchronization

In this year's event we focused on time/phase delivery, including a lot of resiliency scenarios, using full and assisted partial timing support setups, as well as PTP deployments at network edges for legacy networks supporting only Sync-E.

We tested the behavior of the time signal delivery in optimal and suboptimal conditions: network delay asymmetry, hold-over performances, source failover between two grandmaster clocks.

We defined the accuracy level of ±1.5 µs (ITU-T recommendation G.8271 accuracy level 4) as our end-application goal, with 0.4 µs 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 ±1.1 µs.

The primary reference time clock (PRTC) was GPS using an L1 antenna located on the roof of our lab.

The synchronization test team was really prolific this year, with above 40 successful combinations, willing to test brand new software versions, products and interface types, including PTP over 100 GbE.

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

Phase/Time Partial Timing Support

This test was performed using only the ITU-T G.8275.2 profile (PTP telecom profile for Phase/Time-of-day synchronization with partial timing support from the network), without any physical frequency reference – such as SyncE.

In this setup the grandmaster clock was provided with GPS input, while the slave and boundary clock started from a free running condition.

In the first step, we enabled PTP on the boundary clocks while the Calnex Paragon-X was emulating a PDV according to the profile defined in G.8261 test case 12. In the combination including Meinberg LANTIME M4000 as T-GM and Ericsson Router 6675 as T-BC (as well as all the other occurrences of this combination in the current document), the boundary clock did not manage to lock to the grandmaster clock using this attenuation profile. For this reason, we agreed to use an attenuated version of this impairment, where all parameters were reduced by 50%.

After the boundary locked to the grandmaster clock, we let the slave clock also lock to the boundary clock via PTP and verified that the phase accuracy and frequency of the ITU-T G.823 SEC mask requirements were satisfied.

We successfully tested the following combinations: Meinberg LANTIME M4000 and Oscilloquartz OSA5421 HQ++ participated as grandmaster clock, Ericsson-Router 6675 and Meinberg LANTIME M1000S as boundary clock and Microsemi TimeProvider 2300, Microsemi TimeProvider 4100, and Oscilloquartz OSA5421 HQ++ as slave clock.

The Calnex Paragon-X was used to provide the PTP impairments. The measurements for phase output of the devices under test was done either with the Calnex Paragon-X or the Calnex Paragon-T.

In one additional setup of this test case we planned to use a transparent clock instead of the boundary clock, but in this case the slave did not manage to lock to the grandmaster while using the impairment.

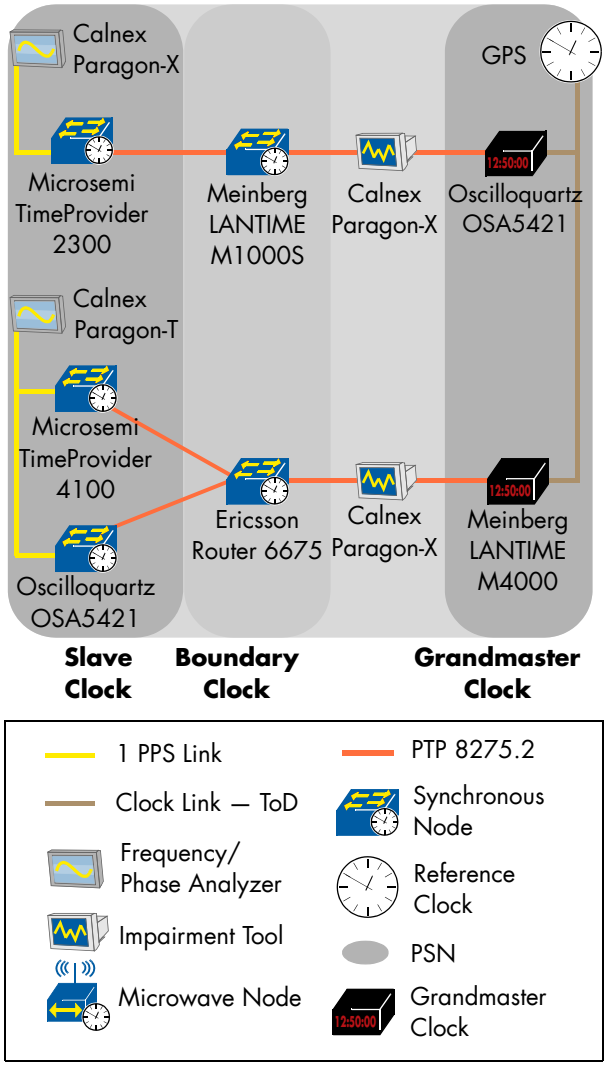


Figure 38: Phase/Time Partial Timing Support

Phase/Time Assisted Partial Timing Support

This test was performed using the ITU-T G.8275.2 profile between the grandmaster and boundary clock, with the participants having the choice of running G.8275.1 or G.8275.2 between boundary and slave clocks.

Both the grandmaster and boundary clocks were connected to GPS and the slave clock locked via PTP to the boundary clock. We then started the impairment which emulated a PDV according to the profile defined in G.8261 test case 12. Some combinations used an attenuated version of this impairment, where all parameters were reduced by 50%.

Upon disconnecting the GPS from the T-BC-P we verified that it switched to PTP and confirmed that the

output met the phase accuracy requirement of ±1.1 µs and frequency requirements. In the last step we reconnected the GPS antenna to the T-BC-P and repeated the measurement.

We successfully tested the following combinations: Ericsson Router 6471, Meinberg LANTIME M4000, Microsemi TimeProvider 4100 and Oscilloquartz OSA5421 HQ++ participated as grandmaster clock, Ericsson-Router 6675, Ericsson MINI-LINK 6651, Ericsson MINI-LINK 6366, Meinberg LANTIME M1000S and Oscilloquartz OSA5421 HQ++ participated as boundary clock, Oscilloquartz OSA5421 HQ++, Microsemi Time Provider 2300, Microsemi TimeProvider 4100 Huawei NE40E-M2K, Huawei NE40E-X2-M8A and Ericsson Baseband 6630 participated as slave clock.

The Calnex Paragon-X was used to provide PTP impairments. The measurements for phase output of the devices under test was done either with the Calnex Paragon-X or the Calnex Paragon-T.

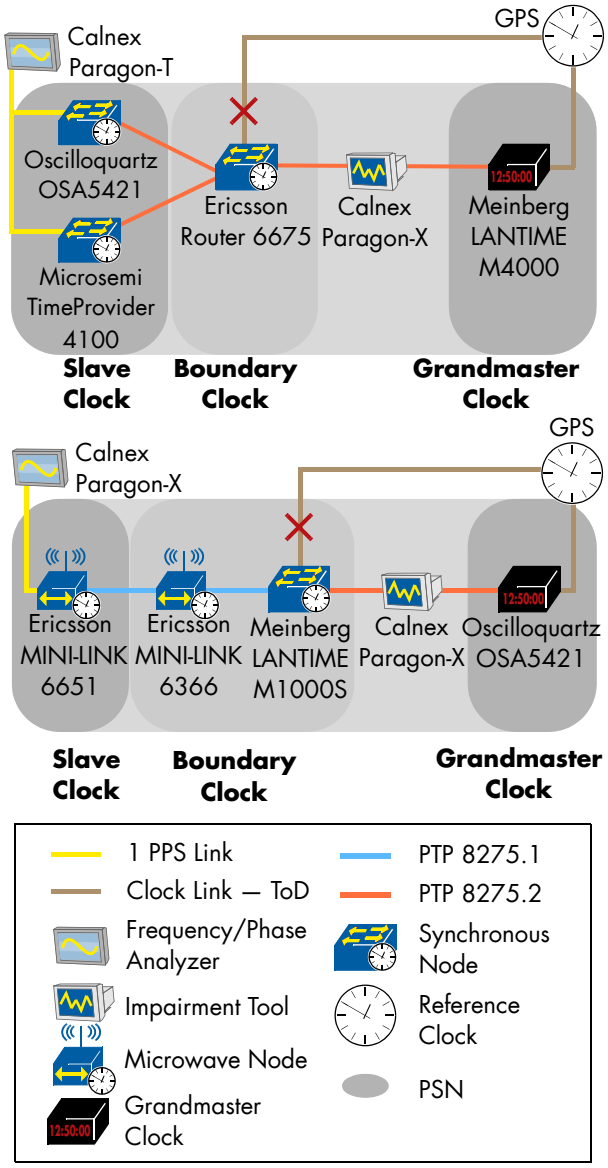


Figure 39: Phase/Time Assisted Partial Timing Support

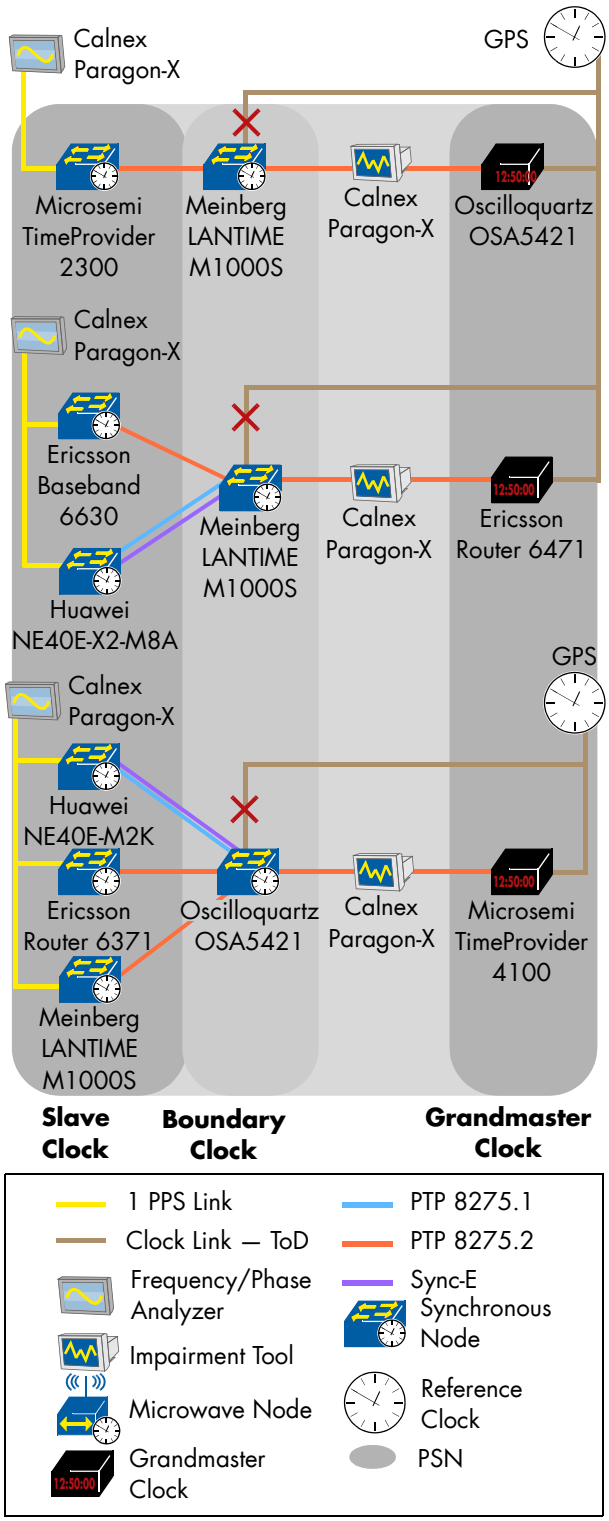


Figure 40: Phase/Time Assisted Partial Timing Support

In one combination we observed a problem with the boundary clock: the APTS backup path did not reach the ready/locked state while using both ITU G.8275.1 and 8275.2 profiles at the same time downstream towards different clients. In another combination we observed that the boundary clock was acting as T-GM and thus did not increment the steps. Removed value and did not show the T-GM parent ID to the slave clocks. The vendor explained that since the T-BC standardization process is still ongoing, this feature is not implemented yet.

Phase/Time Assisted Partial Timing Support: Delay Asymmetry

This test was performed using the ITU-T G.8275.2 profile between the grandmaster and boundary clock, with the participants having the choice of running G.8275.1 or G.8275.2 between boundary and slave clocks.

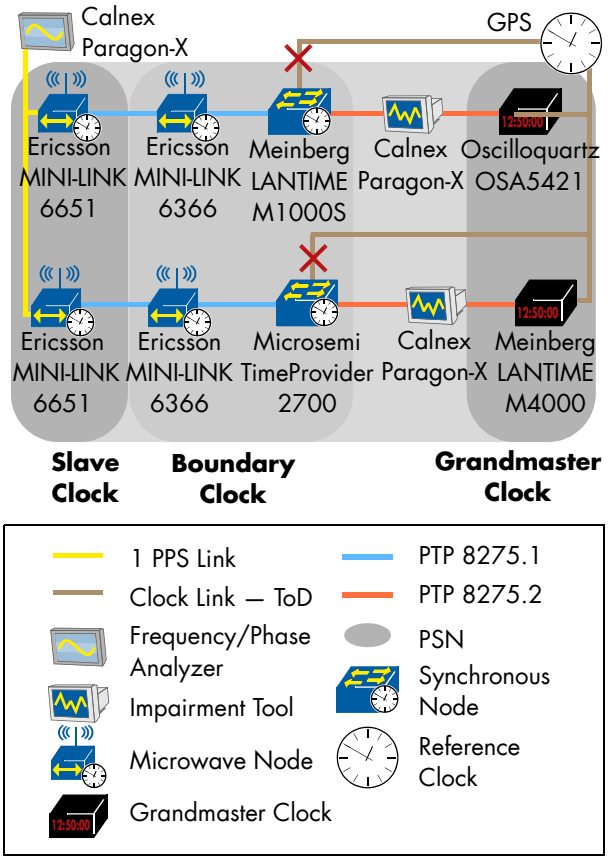


Figure 41: Phase/Time Assisted Partial Timing Support: Delay Asymmetry

Both the grandmaster and boundary clocks were connected to GPS. We started the impairment emulating a PDV according to the profile defined in G.8261 test case 12. Some combinations used an attenuated version of this impairment, where all parameters were reduced by 50%.

After disconnecting the GPS from the boundary clock, we used the Calnex Paragon-X to introduce an additional delay asymmetry of 125 µs and verified that the boundary could calculate and compensate the asymmetry introduced.

We successfully tested these combinations: Ericsson Router 6471, Meinberg LANTIME M4000, Microsemi TimeProvider 4100 and Oscilloquartz OSA5421 HQ++ participated as grandmaster clock, Ericsson Router 6675, Meinberg LANTIME M1000S, Microsemi TimeProvider 2700 and Oscilloquartz OSA5421 HQ++ participated as boundary clock, Ericsson MINI-LINK 6651, Ericsson MINI-LINK 6366, Ericsson Baseband 6630, Huawei NE40E-X2-M8A, Huawei NE40E-M2K, Microsemi TimeProvider 2300, NEC iPASOLINK VR and Oscilloquartz OSA5421 HQ++ participated as slave clock.

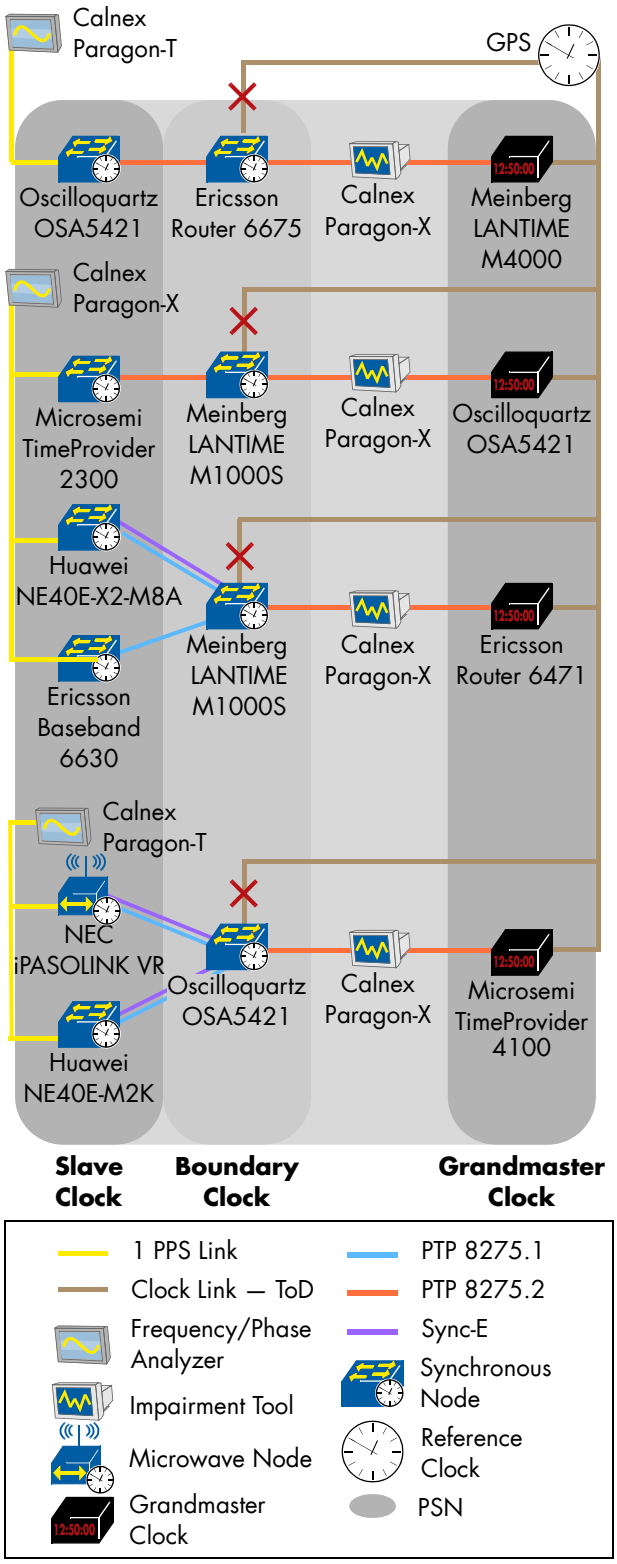


Figure 42: Phase/Time Assisted Partial Timing Support: Delay Asymmetry

The Calnex Paragon-X was used to provide PTP impairments. The measurements for phase output of the devices under test was done either with the Calnex Paragon-X or the Calnex Paragon-T.

In this test we observed different behaviors of the devices acting as boundary clock, depending on the internal oscillator quality. Upon inserting the delay asymmetry, some T-BCs kept the lock to the grandmaster clock, while others went in holdover mode (ClockClass 160) while measuring the

asymmetry introduced, preferring the internal oscillator time. As soon as the calibration was performed, they reverted to ClockClass 6.

Phase/Time Synchronization: Source Failover

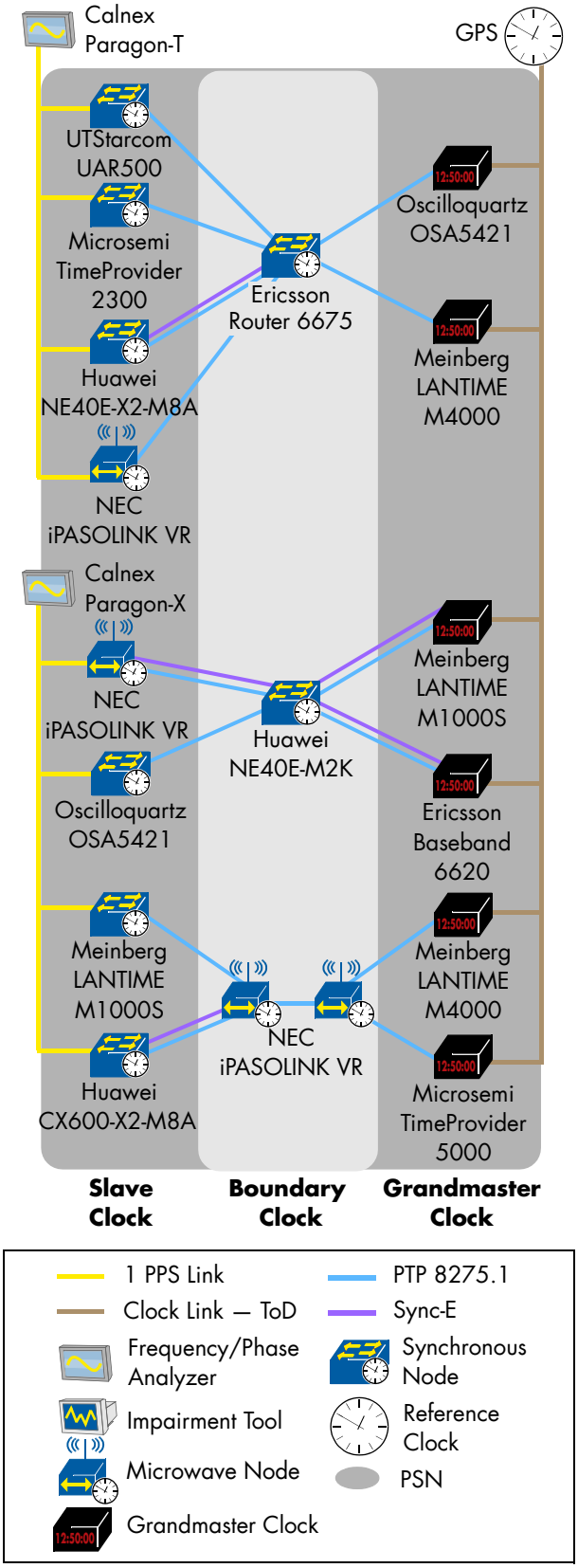


Figure 43: Phase/Time Synchronization: Source Failover

In this setup, both grandmasters were provided with a GPS signal from a common GPS antenna. We allowed the boundary clock to lock to the primary grandmaster and then degraded the primary grandmaster's quality by disconnecting its GPS input. We verified that the boundary clock switched over to the secondary grandmaster and measured the slave clock's transient response. We also tested if the correct clockClass values are being signalled by the grandmasters according to the telecom profiles, which allows the alternate best master clock algorithm running on the boundary clock to correctly select the best grandmaster during each step of the tests.

We used the priority2 field as tie-break parameter.

We successfully tested the following combinations: Ericsson Baseband 6620, Meinberg LANTIME M4000, Meinberg LANTIME M1000S, Microsemi TimeProvider 5000 and Oscilloquartz OSA5421 HQ++ participated as grandmaster, Ericsson Router 6675, Huawei NE40-M2K and NEC iPASOLINK VR participated as boundary clock, Huawei NE40E-X2-M8A, Meinberg M1000S, Microsemi TimeProvider 2300, NEC iPASOLINK VR, Oscilloquartz OSA5421 HQ++ and UTStarcom UAR500 participated as slave clock.

Furthermore two combinations were using 100 GbE links: in the connection between Ericsson Router 6675 acting as boundary clock and Huawei NE40E-X2-M8A and UTStarcom UAR500 acting as slave clocks.

In one failed combination (not shown in this report) we observed an issue related to the ptpTimeScale flag set to TRUE, while the currentUTCOffset was set to 36s and the currentUTCOffsetValid was FALSE.

The PTP standard section 8.2.4.2 a) states that in case of ptpTimescale TRUE the currentUTCOffset shall be obtained from the primary reference (GPS in this case). Therefore we would have expected the UTC offset to be 37s and the valid flag set to TRUE. This issue led the T-BC not to lock to the grandmaster clock.

Phase/Time Synchronization with full Timing Support: Microwave Transport

We started the test with the slave clock in free running mode and generated a constant bit rate at the maximum line rate for the maximum modulation scheme (10% of 576 byte packets, 30% of 64 byte packets, 60% of 1518 byte packets) and expected no traffic loss.

After the slave clock locked, we performed baseline measurements using the Calnex Paragon-T device. To emulate severe weather conditions, we reduced the bandwidth between the two nodes of the microwave network using an RF attenuator. As expected the nodes reacted by changing the modulation used.

We then verified that the PTP traffic was unaffected by the change of modulation, as it was prioritized over other data traffic and the slave clock output retains the required quality level. Since the bandwidth decreased accordingly, we saw that data packets were dropped according to the available bandwidth.

In the first setup, the microwave stations acted as boundary clocks, while in the second setup they were acting as transparent clocks.

We successfully tested the following combinations: Meinberg LANTIME M4000 and Microsemi TimeProvider 5000 participated as grandmaster clock, Ericsson MINI-LINK 6352, Ericsson MINI-LINK 6651 and NEC iPASOLINK VR participated as boundary clock, NEC iPASOLINK VR participated as a transparent clock, Ericsson Router 6371, Huawei CX600-X2-M8A, Huawei NE40E-X2-M8A, Meinberg LANTIME M1000S and Microsemi TimeProvider 4100 participated as slave clock.

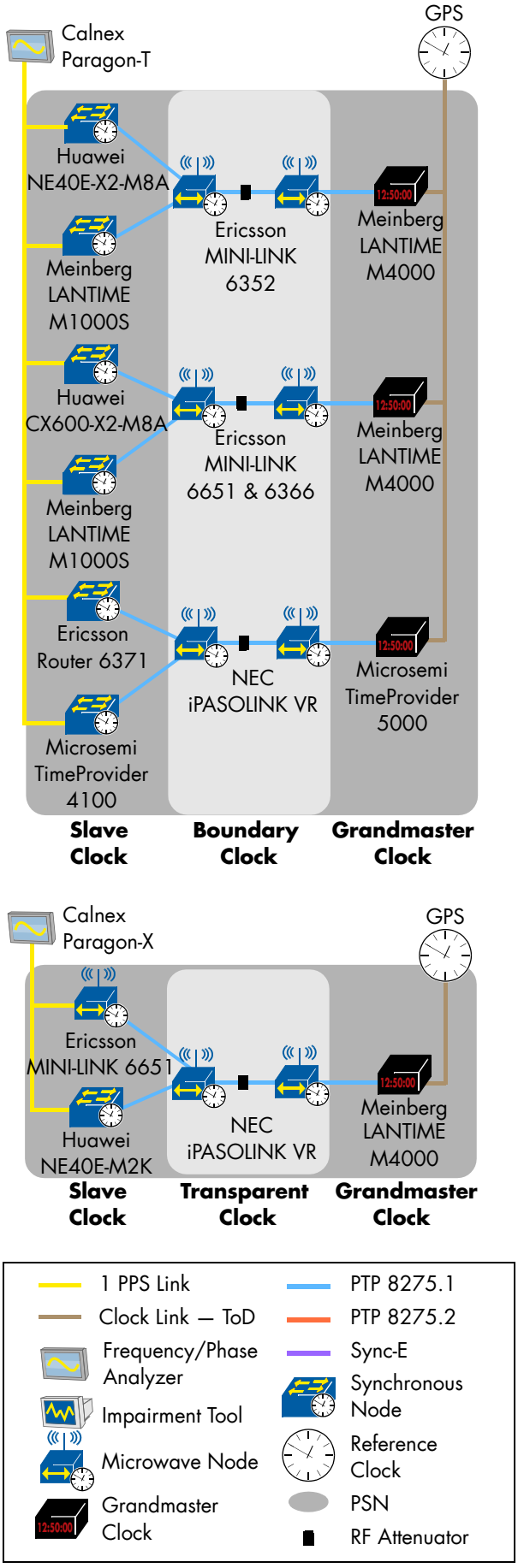


Figure 44: Phase/Time Synchronization with full Timing Support: Microwave Transport

Phase/Time Synchronization: Degradation of Primary Source

According to the architecture defined in ITU-T G.8275 a boundary clock can become a grand-master and can also be slaved to another PTP clock. The goal of this test was to check the capability to swap the role of a boundary clock's port from master to slave and vice-versa.

This test was performed using the ITU-T G.8275.1 profile.

Both the grandmaster and one of the boundary clocks (BC-A) were provided with a GPS signal. We allowed the grandmaster and the boundary clock A to lock to GPS input. The boundary clock A acted as primary grandmaster for the upstream boundary clock (BC-B).

We then disconnected the antenna of the boundary clock A to emulate a GPS failure and verified that both boundary clocks locked via PTP to the central grandmaster. In the last step, we recovered the GPS of the boundary clock A and verified that the boundary clock B locked again to the downstream boundary clock A.

We successfully tested the following combinations: Ericsson Router 6675 and Microsemi TimeProvider 4100 participated as grandmaster clock, Ericsson Router 6371, Huawei CX600-X2-M8A and NEC iPASOLINK VR and Oscilloquartz OSA5421 HQ++ participated as boundary clock B, Meinberg LANTIME M1000S participated as boundary clock A. The link between Ericsson Router 6675 and Huawei NE40E-X2-M8A was based on 100 GbE interface.

In other failed combinations (not shown in the picture), the devices acting as BC-A were not able to swap the operation mode of the port from Master to Slave, as this function was not implemented.

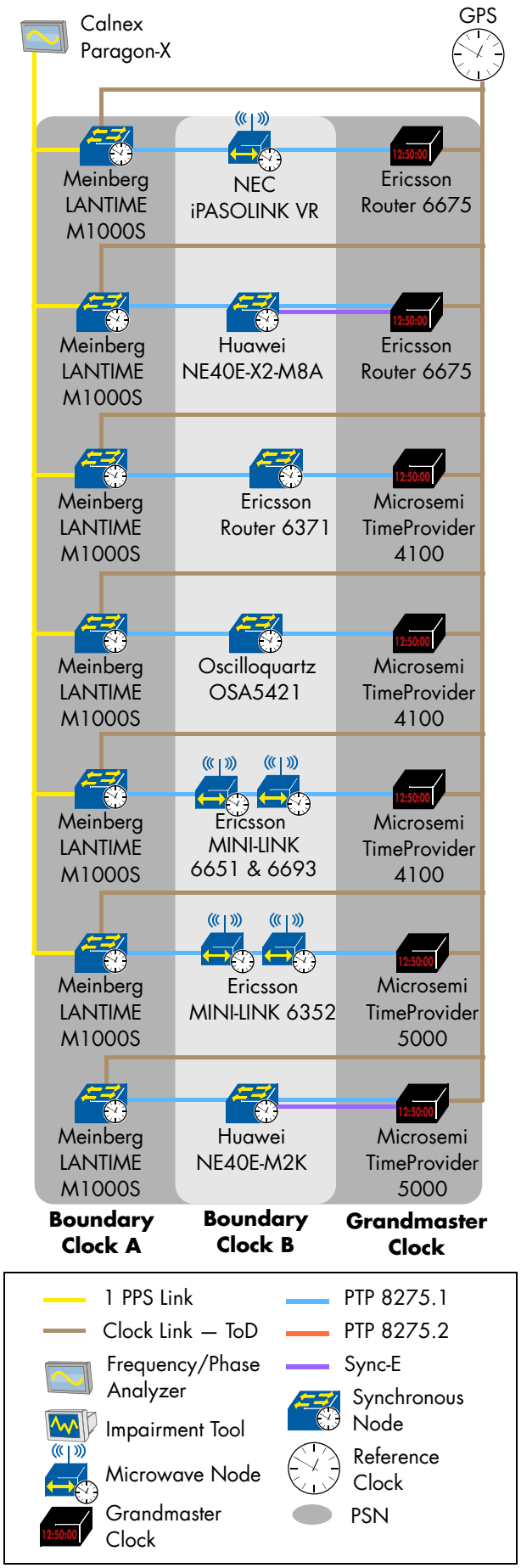


Figure 45: Phase/Time Synchronization: Degradation of Primary Source

Time/Phase holdover with Sync-E support in the core

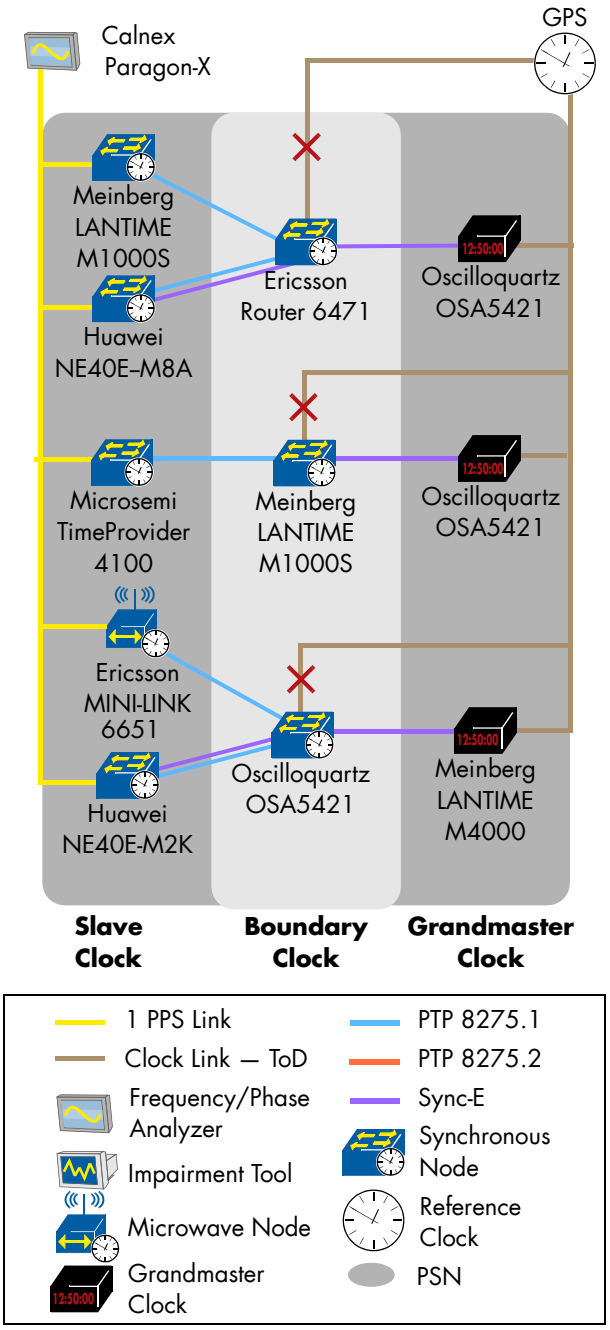


Figure 46: Time/Phase Holdover with Sync-E Support in the Core

In scenarios where it is not possible to deploy PTP in the core network, it is common to have one or more "distributed grandmaster clocks" provided with GPS at the edges of the network.

The goal of this test case was to verify the holdover performance of a boundary clock in relation to phase/time stability while GPS is unavailable and a Sync-E is used as backup to GPS.

In this setup, both the grandmaster and the boundary clocks were provided with a GPS signal from a common GPS antenna.

In the upstream between the boundary and the grandmaster clock only Sync-E was available, while the ITU-T G.8275.1 was used in the downstream link between boundary and slave clock.

In the first step we allowed the slave clock to gain a stable lock, then we emulated a GPS failure of the antenna connected to the boundary clock.

We then measured that the phase accuracy requirement of ±1.1µs is fulfilled for at least 30 minutes while in hold-over.

We successfully tested the following combinations: Meinberg LANTIME M4000 and Oscilloquartz OSA5421 HQ++ participated as grandmaster clock, Ericsson Router 6471, Meinberg LANTIME M1000S and Oscilloquartz OSA5421 HQ++ participated as boundary clock, Huawei NE40E-X2-M8A, Meinberg LANTIME M1000S and Microsemi TimeProvider 4100 participated as slave clock.

In one additional combination (not shown in the picture) we observed that after GPS has been unplugged, the T-BC went into holdover-within-spec, clockClass 135 as expected, but after only 7 minutes changed to holdover-out-of-spec, clockClass 165. In this situation the slave clock lost it PTP lock to the boundary clock.

Summary

It was a pleasure for the whole EANTC team to have 21 dedicated vendors with us in the lab in Berlin. During two intensive weeks, they made great progress and thus jointly advanced the industry. Kudos to all of them and we are looking forward to see what 2019 will bring!

EANTC AG European Advanced Networking Test Center	Upperside Conferences
Salzufer 14 10587 Berlin, Germany Tel: +49 30 3180595-0 Fax: +49 30 3180595-10 info@eantc.de http://www.eantc.com	54 rue du Faubourg Saint Antoine 75012 Paris - France Tel: +33 1 53 46 63 80 Fax: + 33 1 53 46 63 85 info@upperside.fr http://www.upperside.fr

This report is copyright © 2018 EANTC AG.

While every reasonable effort has been made to ensure accuracy and completeness of this publication, the authors assume no responsibility for the use of any information contained herein.

All brand names and logos mentioned here are registered trademarks of their respective companies in the United States and other countries.

2018 v5