MPLS+SDN
+NFVWORLD
→PARIS 2016

Interoperability Showcase 2016 White Paper

EANTC

Editor's Note

Carsten Rossenhövel
Managing Director, EANTC

Welcome to the 2016 edition of the EANTC showcase! This year's white paper reports on our multi-vendor testing of MPLS, SDN, NFV and Clock Synchronization.

The integration of legacy packet transport and software defined networks has advanced to a point where the two technology families begin to integrate.

We witnessed much more industry support for Segment Routing and EVPN this time. Both technologies are at the beginning of their life-cycle; this EANTC test helped to get to the next level of multi-vendor interoperability. We will continue to monitor progress next year.

Interestingly, not all participating transport equipment vendors supplied Segment Routing solutions to our test. This is indicative of the widening range of choices available to service providers, which creates challenges in itself for the industry to address. Likewise, EVPN interoperability tests were limited due to incompatible implementations supporting only MPLS or VxLAN or other transport protocols.

Since these limitations may be tied to hardware, we recommend service providers may want to take particular care to detail their needs accurately when issuing RFPs and ensure they understand how a specific approach taken today may limit or widen their future network design options.

From our SDN testing it is apparent that MPLS and SDN are evolving in the same direction and becoming part of a single topic. Service providers no longer face an either/or scenario. MPLS can be regarded as one of the ways to implement SDN.

The testing of NETCONF/YANG models was of particular interest this year. All test cases went very well. Even so, there's a lot more work to be done as only 25 % of the participating vendors brought NETCONF implementations.

Based on initial, very positive vendor feedback, the EANTC test plan had included an NFV use case for virtual CPEs across SDN into the data center. A number of early ETSI proof of concept (PoC) demonstrations had already focused on this area so we guessed it would not be a major deal to get it working. Seemingly, we were wrong: SDN/NFV integration is still cumbersome in larger, realistic network scenarios. Additionally, some vendors who market NFV and SDN integration the most did not participate in our test; they seem to focus on single-vendor or selective partner solutions — a design approach that service providers will likely help to correct in the future.

EANTC will continue to focus SDN/NFV integration testing, both from the standardization aspect (within ETSI, supporting the TST004 work item) and as part of our *New IP Agency (NIA)* test program.

Across the transport network in our lab, we again tested advanced packet clock synchronization interoperability. As if to highlight the relevance of conducting reliability testing, the worldwide GPS service failed during the first day of hot staging. In general, all test cases were successfully completed with a growing number of participating implementations. We have tested phase synchronization for years; the industry uses it extensively now, with TDD-LTE becoming more widely deployed and time division networks being utilized for digital video broadcast.

Our team and I hope that this year's white paper will provide useful insights into the state of interoperability today. EANTC is available for any questions and welcomes feedback to our test scenarios and results.

Table of Contents

Introduction

We collected fantastic results in two weeks of hot staging and we can split the tests into three distinct areas:

Although we have tested **MPLS and Ethernet Transport technologies** on many occasions, developments continue that make testing valuable. The emergence of Ethernet VPN is a key driver for service providers but interoperability challenges exist and there is scope for vendors to interpret elements of the IETF standardization differently. We sought to clarify inconsistencies in EVPN for both single-homing and multi-homing applications for reporting to the standards body and helping vendors to see where further tuning is needed. In addition we extended the scope of the Segment Routing testing after last year's initial activity, in order to assess interoperability progress as service provider appetite for it has increased. We also revisited IP Fast Reroute to test performance and validate vendors' achievement of service providers' sub 50 millisecond re-routing goal.

Software Defined Networking (SDN), along with complimentary technology network functions virtualization, was among the most discussed topics in the networking industry in 2015 - and that discussion will continue throughout 2016 as service providers roll out the technologies across their networks. Those roll-outs will be at varying paces and in different areas of service provider

businesses according to their strategies and business needs. This year we're looking in more depth at the systems that will manage the virtualized environment and will test virtual components such as CPE as well as orchestration and management elements. We've continued to test OpenFlow performance but are increasingly focusing on the performance of NETCONF/YANG models, which service providers are also embracing.

Progressing beyond our previous successful validation for **Clock Synchronization**, this year we tested interoperability of the two published telecom profiles for frequency and phase/time, and also looked into the latest status of assisted partial timing support before that is standardized. We focused on delay asymmetry, multi-vendor grandmasters, hold over performance and microwave transport. All of these are fundamental areas for service providers and interest has been catalyzed by the deployment of LTE-Advanced which is reliant on phase synchronization. A further stimulus is virtualization which will rely on accurate timing to support all the complex handovers involved.

Participants and Devices

Vendor	Devices
ADVA	FSP150-GE114Pro OSA 5401 OSA 5421
Calnex	Paragon-t Paragon-X Sentinel
Cisco	Nexus 5672UP Nexus 7702 Nexus 9396 ISR 4000 IOS XRv
ECI	NPT 1800
Ericsson	MINI-LINK 6691 MINI-LINK TN RBS 6501 Router 6672 Router 8801 SSR 8004 Virtual Router
Huawei	Agile Controller NetMatrix NE40E-X8A NE40E-X3 NE40E-X2-M8 NE40E-M2E NE40E-M2F ATN 980B ATN 950B ATN 910B
Ixia	IxNetwork
Juniper Networks	CSE2000 MX80 MX240 QFX10002
Meinberg	LANTIME M1000S LANTIME M4000
Metaswitch	vRouter
Microsemi	EdgeAssure 1000 IGM-1100 TimeProvider 2700 TimeProvider 5000
Nokia	7750 SR
Omnitron	iConverter XM5
Spirent Communications	TestCenter N4U
Tail-f Systems	Cisco NSO enabled by Tail-f

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. Sometimes vendors ended up with demonstrations because there was no partner to test with, or because multi-vendor combinations failed so that they could not be reported.

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 Solutions, Ixia and Spirent Communications for their test equipment and support throughout the hot-staging.

MPLS, Ethernet & Data Center Interconnect

MPLS based Ethernet VPNs (EVPN) represents the next generation solution for Ethernet multipoint services by addressing the requirements of Carrier Ethernet and Data Center Interconnect (DCI) market segments. EVPN can also be used as a control plane for an overlay within a data center. Currently defined in RFC 7432, this technology introduces a fundamental shift to control plane-based MAC learning, replacing the data plane-based MAC learning approach employed by VPLS, and relying on Multipoint-to-Point (MP2P) LSPs used instead of pseudowires.

EVPN provides separation between the data plane and control plane. That allows the use of different encapsulation mechanisms in the data plane such as MPLS and Virtual Extensible LAN (VXLAN), which provides a means to interconnect a layer 2 Ethernet segment or layer 3 Services over a layer 3 network. This is important for service providers looking to deploy scalable, dynamic networks with NFV technologies.

However, challenges are abundant in this complex topic and vendors may need to make some adjustments to their systems in order to achieve interoperability.

Ethernet VPNs

Ethernet VPN continues to develop and already counts some active customer deployments, but technical interoperability continues to be among the technology's greatest challenges.

While the base high-level Control-Plane specification for EVPN, RFC7432, is well understood, there is still a need for vendors to clarify the wording of some of the other work-in-progress IETF documents. Furthermore, the drafts offer a lot of options which appear to lead to different vendor implementations.

This year we tested both MPLS and VXLAN based EVPN. The first goal of the tests was to verify that some of the issues identified last year were solved. We also tested a lot of extensions and additional features, like multi-homing, MAC mobility, ARP proxy, Provider Backbone Bridging, Inter-Subnet forwarding and a new Route Type (RT-5).

Single-Homing. VXLAN allows interconnection of a layer 2 Ethernet segment or layer 3 Services over a layer 3 network where the original layer 2 frame has a VXLAN header added and then placed in a UDP packet (MAC-in-UDP encapsulation). A 24 bit identifier, the VXLAN Network Identifier (VNI), is used to designate individual interconnections.

When VXLAN is used as an overlay for EVPN, the VNI directly maps to EVPN EVI in the case of VLAN-based services. In VLAN-aware bundle services, the VNI maps to a bridge table within the EVI. The BGP MAC update will contain the MPLS label field which will be used to carry the VNIs.

During the preparation phase vendors recognized that they were using different options carried in the BGP MAC Advertisement Route.

The Ethernet Tag ID of this route is set to zero for VLAN-based mode, where there is one-to-one mapping between a VNI and an EVI. In such cases, there is no need to carry the VNI in the MAC advertisement route because Broadcast Domain ID can be derived from the RT associated with this route (Ethernet Tag ID must be set to zero). However, for VLAN-aware bundle mode, where multiple VNIs can be mapped to the same EVI, the Ethernet Tag ID must be set to the VNI.

The participating vendors implement either mode, but not both. Since the two modes of operation are not interoperable as per RFC7432, some vendors were able to make VLAN-aware bundle mode work with their VLAN-based implementations by ignoring the Ethernet Tag value when it was other than zero. In other cases we split the test combination to solve this difference.

The vendors interconnected all participating PEs using overlay tunnels (with VXLAN or MPLS Encapsulation). They then configured a common EVPN instance on each PE. In this setup, Ixia IxNetwork was used to emulate Customer Edge (CEs) devices, each attached to a single Provider Edge (PE) in a single-homing set up. In the core network, vendors agreed to use OSPF or IS-IS as the IGP protocol.

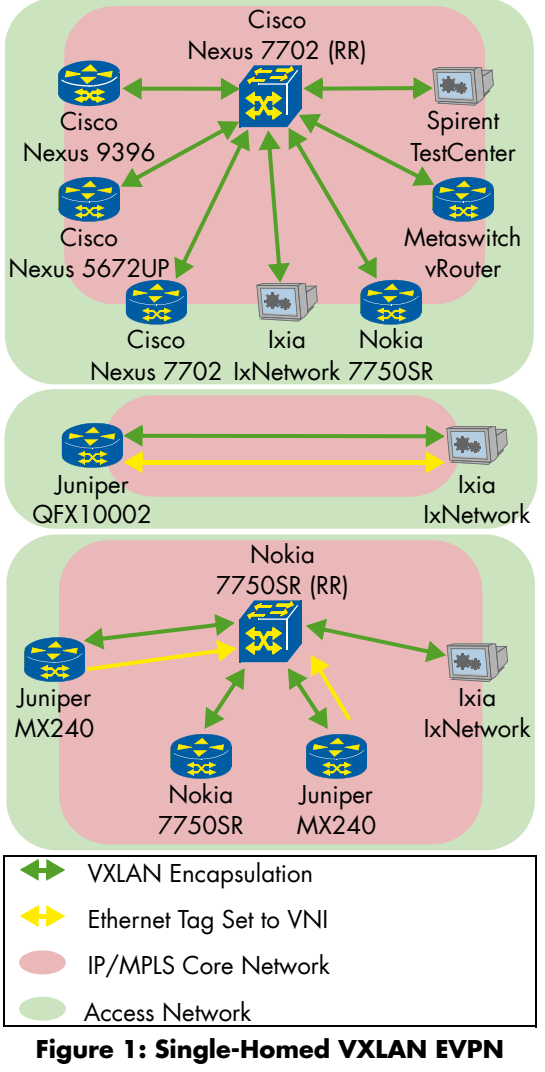


Figure 1: Single-Homed VXLAN EVPN

Once the IGP was up and running in the core network, we enabled MP-BGP between all PEs and route reflector (RR), where used. We first verified that BGP EVPN NLRI was properly negotiated between all BGP peers. For the VXLAN data plane tests, we also verified the use of the BGP Encapsulation Extended Community with encapsulation type 8 (VXLAN). The next step was to assure that each EVPN PE node received Inclusive Multicast Ethernet Tag routes (BGP route type 3) from all other PEs. We then started generating traffic between all emulated CEs and verified that EVPN PEs learned the Customer MAC (C-MAC) on the local segment in the data plane according to the normal bridging operation.

Furthermore, we checked that the previously learned MAC addresses were received on the remote EVPN PE through BGP NLRI using BGP MAC Advertisement route (BGP route type 2). In the last step of this extensive test, we generated bidirectional known traffic between all CEs using Ixia IxNetwork. We did not observe traffic loss for the configured services.

The following devices participated in this test using VLAN-based VXLAN encapsulation: Cisco Nexus 9396, Cisco Nexus 5672UP, Cisco Nexus 7702, Ixia IxNetwork, Metaswitch vRouter, Nokia 7750SR, and Spirent TestCenter. Another Cisco Nexus 7702 was used as a Route Reflector. We verified that all devices could send any-to-any unicast traffic. We used bidirectional Protocol Independent Multicast (BiDir PIM) to verify multicast traffic in the overlay. Devices supporting Ingress-Replication (Cisco Nexus 9396, Ixia IxNetwork, Metaswitch vRouter, Nokia 7750SR, Spirent TestCenter) showed the ability to exchange Multicast Traffic.

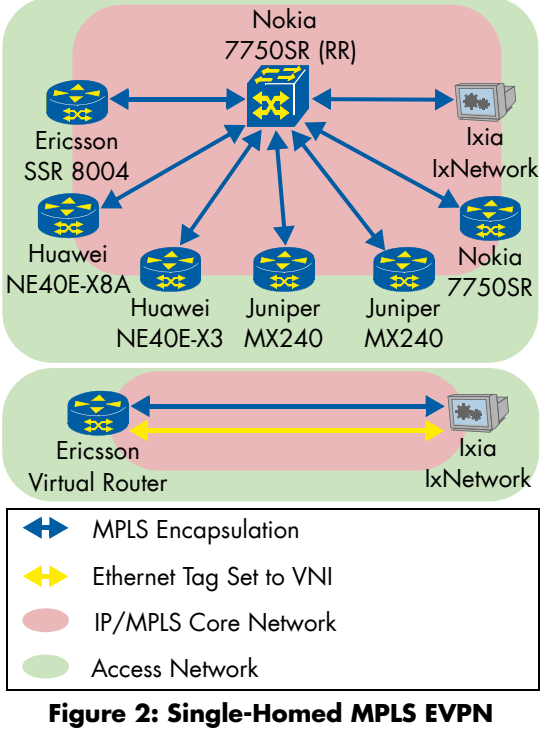


Figure 2: Single-Homed MPLS EVPN

The following devices participated in the test using VLAN-aware VXLAN encapsulation in two different setups (please see Figure 1 for more details): Ixia IxNetwork, Juniper QFX10002, two Juniper MX240, two Nokia 7750SR (one acting as PE, one as Route Reflector).

Moving to the MPLS-based EVPN, we successfully tested the VLAN-based EVPN between Ericsson SSR 8004, Huawei NE40E-X8A, Ixia IxNetwork, Juniper MX240 and Nokia 7750SR.

In addition, another test pair successfully participated in the VLAN-aware EVPN using MPLS encapsulation: Ericsson Virtual Router and Ixia IxNetwork.

We also observed an issue with the Ethernet Tag value carried in the Route-Type 2 when using a Route Reflector (RR).

One of the RRs used to test EVPN VXLAN, programmed to work only in VLAN-based mode could not distribute routes between two PEs working in VLAN-aware bundle mode. The RR dropped the packets with Ethernet Tag value set to non-zero values, judging them as malformed. In other words, the main reason for this interoperability issue was that one vendor was expecting a transparent advertisement through the RR of the other vendor which supported only non-transparent distribution. The routes were not able to be forwarded to other PE devices supporting the same operating mode.

Although the vendors had different opinions about the behavior of the RR in this specific situation, the common understanding is that the RR is a service-unaware device, so as long as the EVPN route is valid, it should be reflected.

MAC-Mobility. EVPN provides greater flexibility and control over the MAC mobility process. MAC mobility allows flexibility for a given host or end-station (as defined by its MAC address) to move from one Ethernet segment to another. EVPN introduces sequence numbering in its Type 2 routes which prevent race conditions which might exist with multiple rapid moves (RFC 7432 section 7.7).

A PE receiving a MAC/IP Advertisement route for a MAC address with a different Ethernet segment identifier and a higher sequence number than it had previously advertised withdraws its MAC/IP Advertisement route.

Based on the topologies created previously (see paragraph Single-Homing.), we tested the Mac-Mobility feature by replicating an existing MAC Address on the traffic generator (Ixia IxNetwork) port connected to another PE. As expected the new PE realized that the MAC was already present at another site and announced it with a higher sequence number. The first PE sent a withdrawal for that C-MAC. The following devices showed to support the Mac-Mobility feature in the VXLAN based scenario: Cisco Nexus 9396, Cisco Nexus 7702, Ixia IxNetwork, Metaswitch virtualized vRouter, Nokia 7750SR. For the MPLS encapsulation scenario we tested the following devices: Huawei NE40E-X8A, Huawei NE40E-X3 (acting both as PE and RR), Ixia IxNetwork.

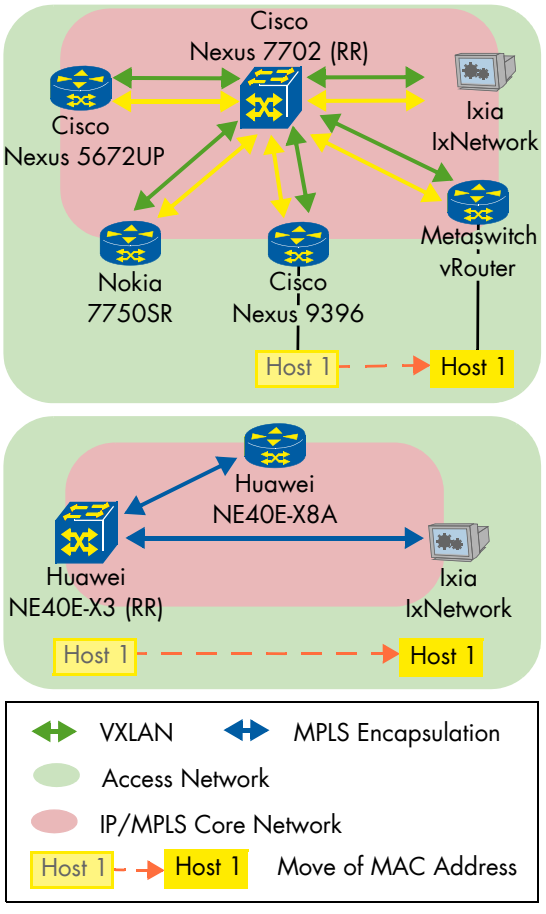


Figure 3: MAC Mobility

ARP Proxy. The 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. If a PE receives an ARP request from CEs, it intercepts the ARP request and performs a Proxy-ARP lookup for the requested IPs. If the lookup is successful, the PE will send an ARP reply without flooding the ARP request to the EVPN network or any other local CEs.

We tested this additional feature using the EVPN test setups previously created (see Figure 1). We sent an ARP request using our test equipment and verified that the PEs intercepted any ARP request for the MAC Addresses already present in their local table without flooding it to the other PEs.

The following devices showed the capability to suppress ARP flooding within the EVPN by activating the ARP-Proxy feature: Cisco Nexus 5672UP, Cisco Nexus 7702, Cisco Nexus 9396, Ericsson Virtual Router, Metaswitch vRouter, Nokia 7750SR.

Multi-Homing. One key feature of EVPN is multi-homing. A multi-homed customer site attached to two or more PEs can increase the site's availability as well as enable load balancing between the links.

In this setup Ixia IxNetwork emulated a CE and was dual-homed to two PEs, Juniper MX240 and Nokia 7750SR, using single-active redundancy mode and manually configured DF. One additional Juniper MX240 acted as remote PE.

We first verified that the VXLAN EVPN was correctly set up and the PEs were learning C-MAC destination addresses. Ixia's emulated CE started any-to-any unicast and multicast traffic and we could verify that the unicast traffic was correctly sent on the active link and the multicast traffic was not replicated on the secondary link.

We inserted a link failure by shutting down the interface between the DF (Juniper MX240) and the CE and measured the outage by counting the packets lost while the Ixia test equipment was sending packets at a fixed rate of 1000 pkts/s (1 packet lost = 1 ms outage).

As expected, we observed that the port failure triggered the BDF (Nokia 7750SR) to take over and that the traffic was swapped to the secondary link. The measured outage for this failover was around 3 seconds.

According to the Nokia engineer, the reason for this delay is that the BDF waited for the DF election timer before taking over.

The RFC7432 says that the failure should trigger again the DF election but it doesn't specify whether this election should use the same timers it uses when a new PE joins the Ethernet Segment or happen immediately.

Consistent with this explanation, we observed that the Nokia 7750SR was implemented to wait for the default election timer (set by default to 3 seconds, which can be reduced up to zero) before taking over, while other vendors (see the demonstration scenarios for more single-vendor MH setups) prefer to switch immediately to the backup DF. Despite this difference the test pair showed to interoperate.

We want to highlight that the failover time optimization was not the goal of this test, but rather making sure both control and data plane interop by working through all these events as described above.

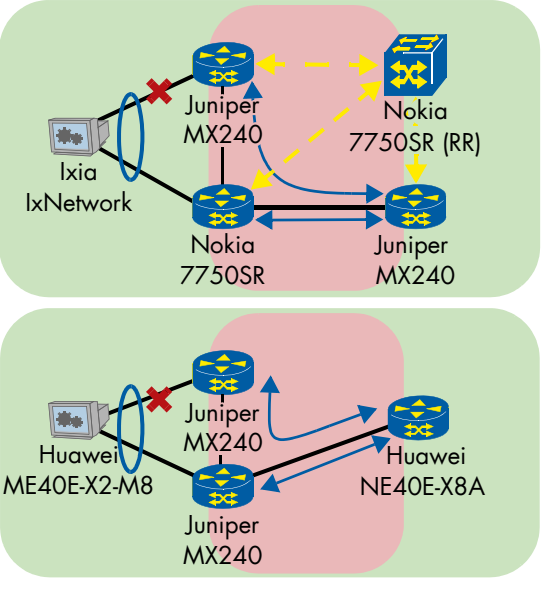


Figure 4: EVPN Multi-Homing

Upon restoring the link, the DF started again forwarding the traffic with an outage of 185 ms.

Huawei and Juniper successfully participated in a Multi-Homing test using two Juniper MX240 as multi homing pair and Huawei NE40E-X8A and Huawei NE40E-X2-M8 as remote PEs. When we introduced the link failure, the PEs switched the traffic to the active links upon receiving the route withdrawal for failed link and we measured a maximum of 125 ms outage. Restoring the link caused a maximum of 197 ms of outage.

Provider Backbone Bridging EVPN. The development of Provider Backbone Bridging EVPN (PBB-EVPN) is specified in RFC 7623. PBB-EVPN extends EVPN by employing the PBB mechanism to reduce the number of BGP MAC advertisements via aggregation, provide backbone MAC (B-MAC) subnetting and client MAC (C-MAC) mobility. We tested this feature both in a single-homing and in a multi-homing setup.

In this test we first verified that each PE established a BGP session with the RR and the EVPN correctly set up. We started a bidirectional traffic flow between the CEs and observed the local C-MAC learning and the bindings between the learned C-MAC to the corresponding B-MAC. We verified that the PEs advertise the local B-MAC reachability information in BGP to all remote PEs in the same EVPN instance through BGP NLRI using MAC Advertisement Route (type 2).

We then repeated this test with a Multi-Homing setup. In this case the vendors agreed to use All-Active multi-homing, so we also verified that the PEs were able to load-balance between the links. Upon a link failure (between DF and CE) the DF sent a withdrawal of the B-MAC route and the backup DF took over the traffic transmission.

We tested two combinations for the Single-Homed scenario. The first one included Huawei NE40E-X8A, Huawei NE40E-X3 and Nokia 7750SR. The second one included Ixia IxNetwork, two Juniper MX240 and two Nokia 7750SR. One of the Nokia devices was acting as RR.

In the Multi-Homing scenario Ixia IxNetwork emulated a CE behind the Multi-Homing pair, which was composed of one Juniper MX240 and one Nokia 7750SR. The Remote PE was also a Juniper MX240 and another Nokia 7750SR acted as RR.

Huawei, Ixia IxNetwork and Nokia successfully participated the PBB Multi-Homing feature in two different setups. The first one included Ixia IxNetwork as emulated CE behind one Huawei NE40E-X8A and Huawei NE40E-X3. Ixia IxNetwork was also acting as Remote PE.

In the second setup a Huawei NE40E-X2-M8 acted as CE behind Huawei NE40E-X8A and Huawei NE40E-X3. Two Nokia 7750SR were acting as Remote PEs.

While testing another combination we observed one issue due a control word (as based on the Generic PW MPLS Control Word as defined in RFC 4448) included into the MPLS packet (inserted between MPLS label and ethernet header). This was not expected by another vendor. The RFC 7623 (PBB-EVPN standard) does not clearly specify the use of PW control word. However it reuses the data path specified in RFC7432 which states that the implementation should use control word.

In addition, one vendor sent VLAN-tagged Ethernet-based payload, which was not expected by the other vendor. For these reasons both vendors could not inter-operate with each other.

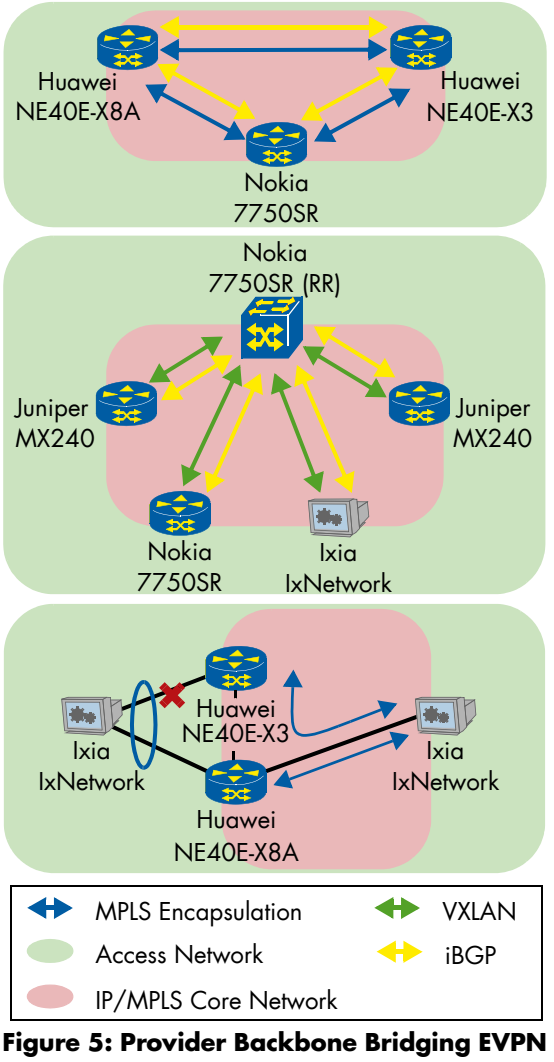


Figure 5: Provider Backbone Bridging EVPN

Inter-Subnet Forwarding. Ethernet VPN (EVPN) Integrated Routing and Bridging (IRB) provides a solution for both intra- and inter-subnet forwarding. Both types of forwarding are useful in a data center environment where there is a need for both layer 2 and layer 3 forwarding to enable interworking with tenant layer 3 VPNs. EVPN IRB is still work in progress at the IETF.

Traditionally, when two Tenant Systems belonging to different subnets connected to the same PE wanted to communicate, their traffic needed to be back hauled to centralized layer 3 Gateway (L3GW) nodes where inter-subnet switching is performed and then back to the PE node where both subnets were attached. In an EVPN network environment an EVPN IRB provides the solution to overcome the traffic hair-pinning issue at a central L3GW by routing and bridging traffic locally at the EVPN PE (NVE).

There are two broad approaches for IRB described in the draft document: Asymmetric and Symmetric. As their names indicate, these modes have different congruency behaviors for bidirectional flows as well as different host's MAC/IP learning requirements on VTEP switches. In the asymmetric IRB scenario both layer 2 and layer 3 lookup associated with the inter–subnet forwarding are performed in the ingress PE, whereas the egress PE performs layer 2 lookup only. In the symmetric IRB scenario both ingress and egress PEs perform layer 2 and layer 3 lookup. Since both forwarding modes are not interoperable we created two setups for the tests.

In our test setup we observed the exchange of the Route-Type 2 and verified that the PEs imported and installed the correct information into the MAC-VRF table (C-MAC along with the BGP next-hop address as tunnel destination address and the VXLAN VNI corresponding to MAC-VRF) and into the IP-VRF table (Customer IP-Addresses along with the corresponding EVPN PE's MAC address from MAC-VRF along with the BGP next-hop address as tunnel destination address and the VXLAN VNI corresponding to IP-VRF).

The following devices successfully participated in symmetric inter-subnet forwarding: Ixia IxNetwork, Cisco Nexus 9396, Cisco Nexus 5672UP and two Cisco 7702 (one as PE and one as route reflector).

One test pair also successfully participated in the asymmetric inter-subnet forwarding: Ixia IxNetwork and Juniper MX240.

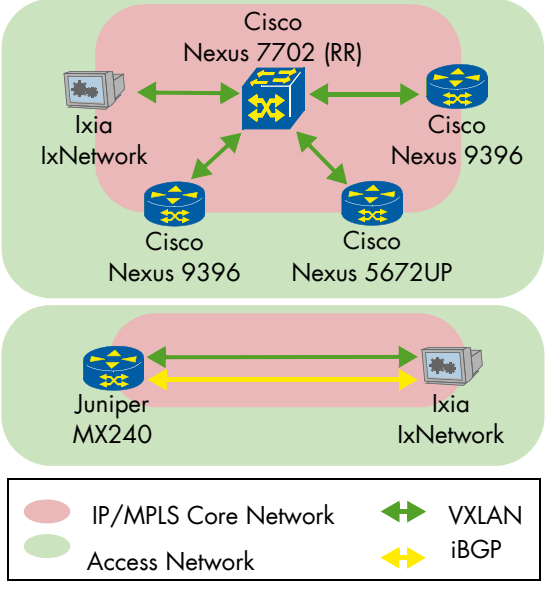


Figure 6: EVPN Inter-Subnet Forwarding

Route Type 5. Extension to RFC7432, RT-5 defines how EVPN may be used to advertise IP Prefixes.

The route type 5, i.e. IP Prefix Advertisement route, decouples the advertisement of IP prefixes from the advertisement of any MAC address related to it. This brings some major benefits required by certain inter-subnetwork forwarding scenarios. During the hotstaging we tested the IRB forwarding on NVE for subnet (IP-VRF-to-IP-VRF) scenario.

We verified the format of the Type 5 routes, which should carry the RD of the related EVI (IP-VRF RT), the IP Address (with IP address length) of the next hop, the Gateway IP Address set to 0.0.0.0 and the BGP Encapsulation Extended Community indicating the tunnel encapsulation (VXLAN). We then exchanged unicast traffic between the traffic generators behind the PEs and verified that no packets were lost.

We carried out the test with two different combinations (see Figure 7) and verified that the following devices correctly exchanged the RT-5 and were able to exchange intra- and inter-subnet traffic: Cisco Nexus 9396, Cisco Nexus 7702 (route reflector), Cisco Nexus 5672UP, Ixia IxNetwork, Juniper QFX10002, Nokia 7750SR and Spirent TestCenter. Devices supporting ingress replication (Cisco Nexus 9396, Ixia IxNetwork, Juniper QFX10002, Nokia 7750SR and Spirent TestCenter) could also exchange intra-net multicast traffic.

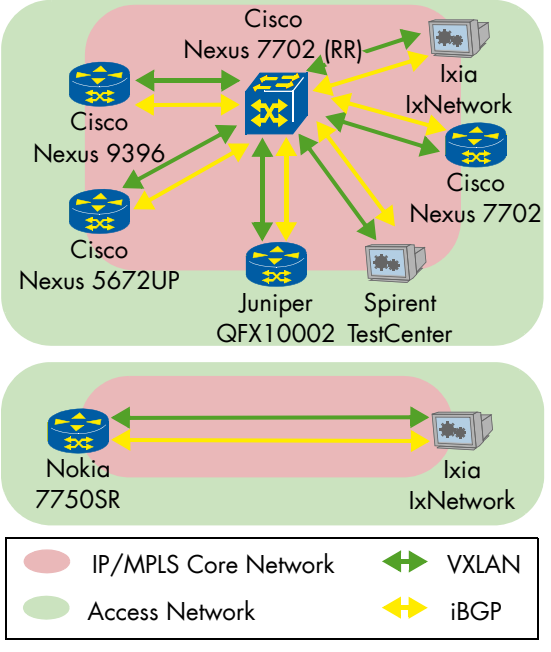


Figure 7: Route-Type 5 Exchange

Segment Routing

Following on from last year's first test of Segment Routing, we sought to verify functionality by setting up an end-to-end path within one domain and using an Interior Gateway Protocol (IGP) to distribute segment information. This year's test focused on two key areas. First, verifying the distribution of the forwarding state, node and adjacency segments, using extension to the existing link-state routing protocols for Segment Routing, and second, verifying that the edge network nodes do a proper encoding of the data path as a stack of segments. There is significant service provider enthusiasm for Segment Routing because of its capability to enable simplification as networks become larger and more complex.

Segment Routing with IGP. In this test, we created an IS-IS network and verified the capability

of the device under test to exchange node segment information based on the IS-IS routing protocol. We configured the ingress edge router (PE located at the edge of core network, also participating in the access network) to establish two segment routing paths: the dynamic shortest path and the explicit path. The latter path should request a next hop located in a longer path towards remote egress edge router.

We observed that in line with the path selection criteria all paths were successfully established on the edge router. These included a dynamic path consisting of a single hop starting after the edge router and another explicit path with two hops towards the egress edge router. Finally, we sent IPv4 test traffic from the traffic generator to the ingress edge router and, as expected, the edge router - by introducing test traffic into configured paths to carry stacked MPLS labels - successfully sent segment packets over MPLS data plane. We also verified that each intermediate node switched the stacked MPLS packets without any lost packet.

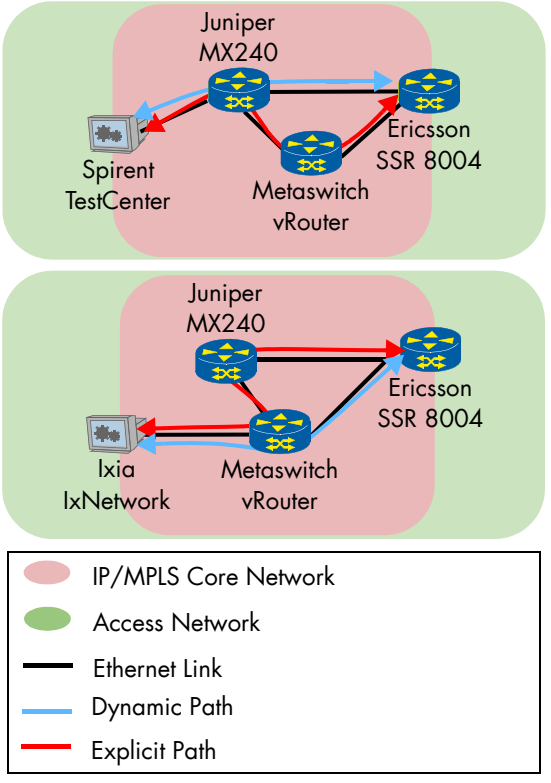


Figure 8: Segment Routing with IGP

The following devices successfully participated in this test: Ericsson SSR 8004, Ixia IxNetwork, Juniper MX240, Metaswitch vRouter and Spirent TestCenter.

A combination initially showed packet loss because a vendor's device only supported PHP (Penultimate-hop Popping), which was not expected by another vendor's device. This issue was soon solved, because the latter vendor successfully implemented the PHP function to the device during the hot staging, so both parties tested successfully.

Segment Routing with IGP-based MPLS Tunneling. The test verified also that MPLS-enabled services were transported over an SR domain without any modification to the service operation (both control or data plane). We used the IS-IS protocol to build the SR domain and verified that on the edge router both the dynamic shortest path and the explicit path (latter case including one additional hop) were established as expected. We used IS-IS as control plane protocol without any requirement of LDP and RSVP-TE signaling protocols. We then configured VPWS services based on the pre-determined paths between the edge routers.

As expected, we observed that VPWS network services were established successfully over the SR paths between the edge routers. All VPWS data traffic was forwarded without any frame loss.

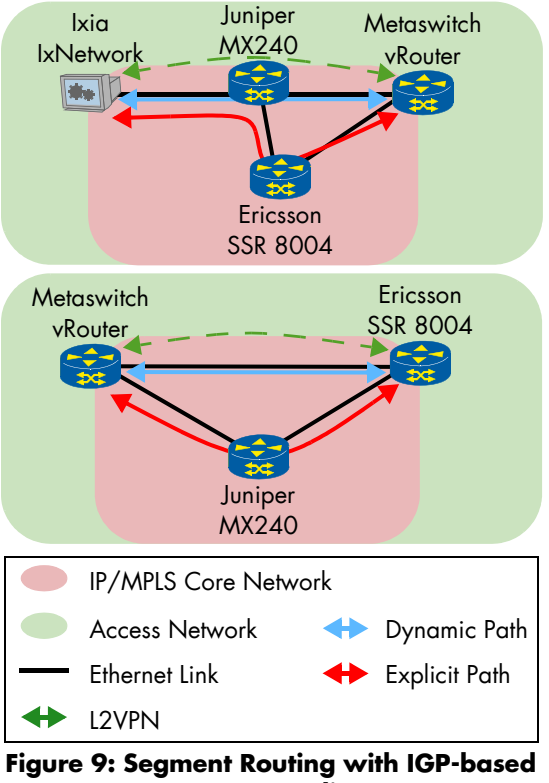


Figure 9: Segment Routing with IGP-based MPLS Tunneling

The following devices successfully participated in this test: Ericsson SSR 8004, Ixia IxNetwork, Juniper MX240, Metaswitch vRouter and Spirent TestCenter.

Resiliency

IP Fast Reroute — Loop Free Alternates. IP FRR is the calculation and usage of Loop Free Alternates to provide local protection for unicast traffic in pure IP and LDP-based MPLS networks. This enables networks to support sub-50ms recovery times without the overhead of complex soft-state protocols that may have trouble scaling to the larger networks of the future. Backup routes are repaired locally by the routers detecting the failure without the immediate need to inform other routers of the failure. The disruption time is limited and taken to detect the adjacent failure and invoke the backup routes. IP FRR consists of two features: a mechanism for the routers adjacent to the failure to

rapidly invoke a repair path which is unaffected by any subsequent re-convergence and a micro-loop control mechanism in topologies that are susceptible to micro-loops.

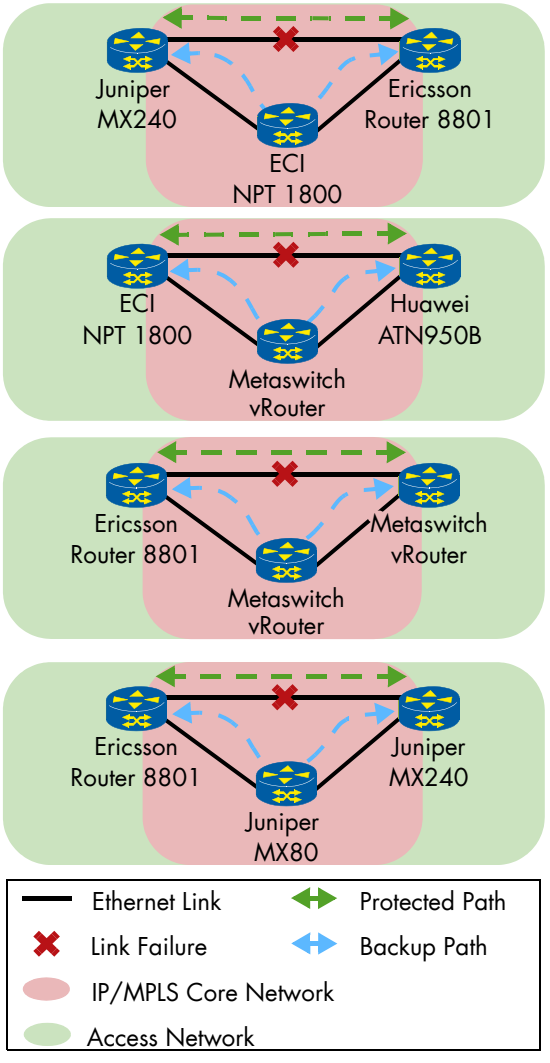


Figure 10: IP Fast Reroute

We tested the vendors in four groups, each consisting of three devices and validated the data path using L2VPN. IS-IS was used as an IGP and a targeted LDP session was established between each pair of PEs for the L2VPN traffic.

We first verified that the Loop-free criteria was satisfied (advertised metrics must ensure that backup path does not use the broken link) and that the ingress PE installed both the primary next-hop MPLS label and the alternate next-hop in its forwarding table.

We introduced the link failure by physically disconnecting the cable between ingress and egress PEs. We expected the LDP (MPLS) traffic to follow the loop-free alternates indicated by the IGP in case of failure.

The test equipment was generating packets at a fixed rate of 1,000 packets/s and we measured the failover and recovery time by counting the packets lost (1 packet = 1 ms).

At the end we verified that no backup path was installed if the Loop-Free criterion was not satisfied by increasing the metrics between remote PE and backup P Router.

Five vendors participated in this setup using the following devices: ECI NPT1800, Ericsson Router 8801, Huawei ATN950B, Juniper MX240, Juniper MX80 and Metaswitch vRouter. We tested them in different combinations, rotating the positions and the roles of each device.

Table 1 shows the maximum, minimum and average outage time in each direction (referring to Figure 10) upon breaking and restoring the protected path. The traffic redirection always took less than 50 ms. When the link was restored the operations were even faster but still showed some packet loss.

Table 1: Measured Outage Time for IP LFA

Event	Max.	Min.	Average
Link broken	48 ms	4 ms	32.7 ms
Link reconnected	29 ms	0 ms	9.1 ms

Remote Loop Free Alternates FRR. Remote LFA FRR is a resiliency approach that extends the basic IP FRR mechanism by using tunnels to provide additional logical links which can then be used as loop free alternates. Remote LFA FRR addresses the limitation of the basic loop-free alternate (LFA) mechanism in a ring based topology.

The basic idea behind the Remote LFA FRR is to find a set of nodes that can be reached by the Point of Local Repair (PLR) immediate neighbors without traversing the primary next-hop (extended P-space) and nodes that can reach the destination by normal forwarding, without traversing the failed link (Q-space). A set of nodes in extended P-space and Q-space are termed PQ nodes. Tunnel technologies, such as Multiprotocol Label Switching (MPLS), IP in IP, GRE or Segment Routing can then be used to encapsulate traffic from the PLR to the PQ-nodes.

We tested four combinations of vendors, each consisting of four devices and validated the data path using L2VPN or L3VPN. We first verified that the IGP (OSPF or IS-IS) and LDP session were correctly set up between all the devices. We then verified that the repair tunnel was established between the PLR and the PQ-node. While generating traffic at a fixed rate of 1000 Pkts/s we introduced a link failure by disconnecting the cable and measured the time the PLR needed to swap the link by counting the packets lost (1 packet lost = 1 ms outage).

Following devices participated in the test as PLR: Ericsson SSR 8004, Huawei NE-40E-X2-M8, Huawei NE40E-M2F, Metaswitch vRouter. The following devices acted as PQ: Huawei NE-40E-M2E, Juniper MX240, Metaswitch vRouter. In addition we used the ECI NPT 1800 and the Huawei ATN980B as P Router.

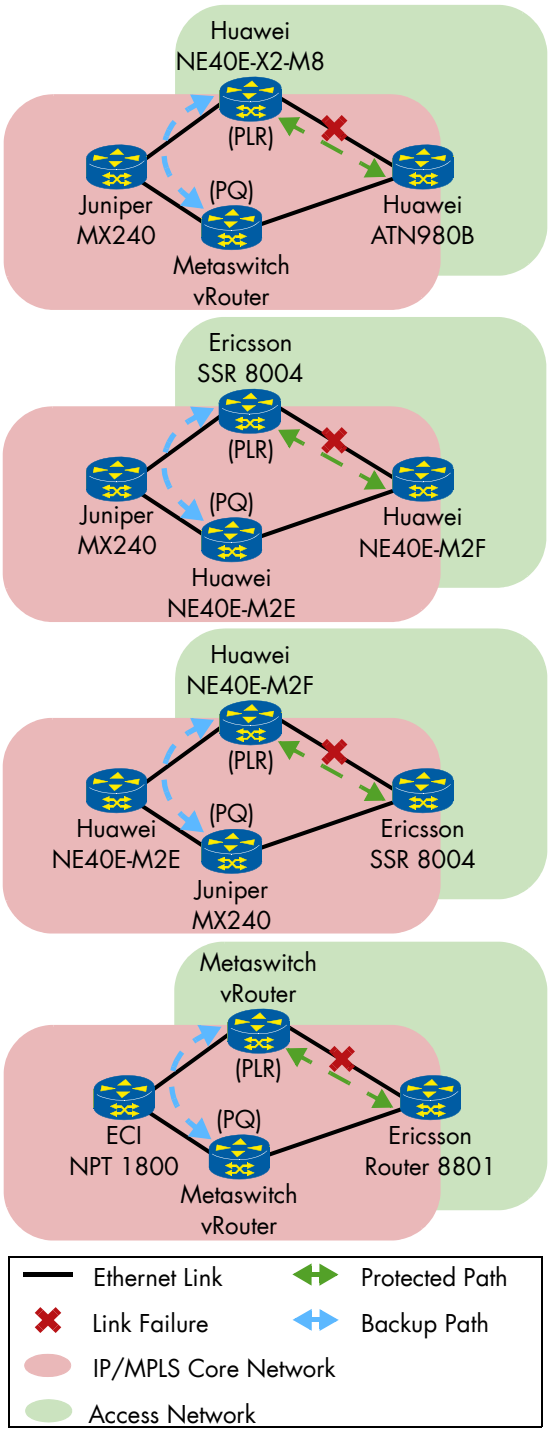


Figure 11: Remote LFA FRR

Upon Link failure the devices reacted with very different speeds. We measured a wide span of results, as shown in Table 2. After restoring the link, we measured no packet loss in all the combinations.

Table 2: Measured Outage Time

Event	Min.	Max.	Average
Link broken	3 ms	126 ms	64.6 ms
Link reconnected	0 ms	0 ms	0 ms

BGP Fast Reroute. The BGP Fast Reroute (FRR) feature improves BGP convergence after a failure in the network. BGP FRR creates and stores a backup path in the forwarding information base. When a failure on the primary path is detected, the backup path can immediately take over. The convergence does not depends on the number of prefixes, thus enables fast failover times. BGP FRR can be categorized into either Core or Edge. BGP FRR Core describes the scenario where a link or node on the path to the BGP Next-hop fails, but the next-hop remains reachable. BGP FRR Edge describes a scenario where an edge link or node fails, resulting in a change of the next-hop.

For time reasons we only tested the Core feature.

During this test we first verified that the IGP, LDP and BGP session were correctly established. We then verified that the PE installed the primary and the secondary path in its forwarding table.

We used the traffic generator to send bidirectional traffic at a fixed rate of 10,000 packets/s. We introduced a link failure by disconnecting one interface along the protected path and measured the outage time for the failover by measuring the packet loss (1 packet lost = 0.1 ms outage).

We tested one combination of devices, consisting of Ericsson SSR 8004 (P Router), Huawei ATN980B (PE), Huawei NE40E-M2F (P Router) and Spirent TestCenter (Traffic Generator and emulation of remote BGP peers). The time needed to switch from primary to backup path was of 2.3 ms.

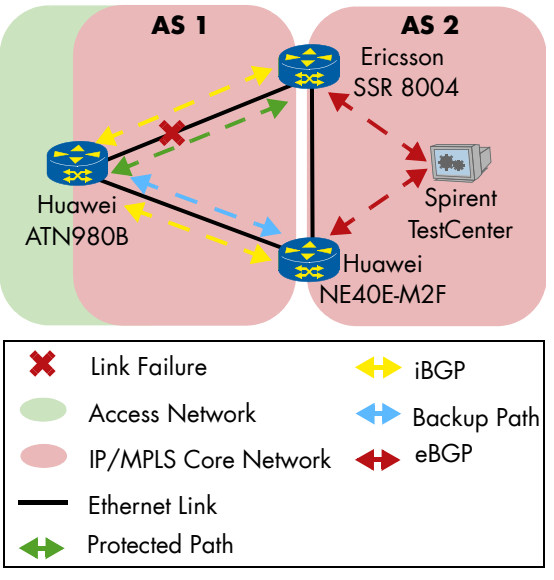
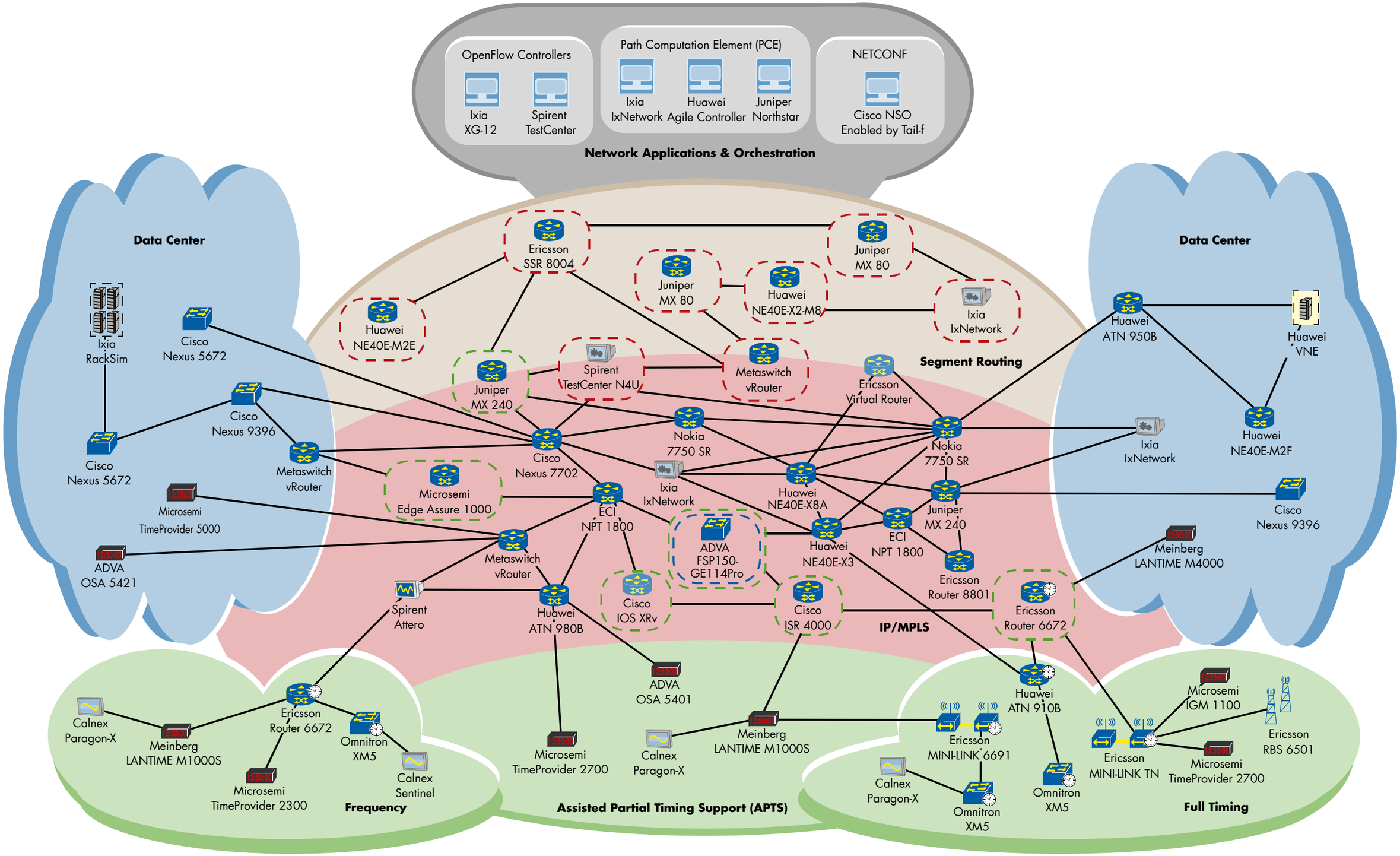


Figure 12: BGP Fast Reroute

BGP Error Handling. The BGP error handling as revised in the IETF RFC7606 allows a BGP speaker that receives an UPDATE message containing a malformed attribute (existing attributes as nominated in the RFC 7606) to avoid resetting the BGP session. The improved error handling is aimed to minimize the impact on routing upon receiving a malformed UPDATE message while maintaining protocol correctness to the extent possible.

In this test we emulated the advertisement of one IPv4 and one IPv6 prefix to the device under test.


OpenFlow Controllers
Ixia XG-12
Spirent TestCenter
Path Computation Element (PCE)
Ixia IxNetwork
Huawei Agile Controller
Juniper Northstar
NETCONF
Cisco NSO Enabled by Tail-f
Network Applications & Orchestration
Data Center
Ixia RackSim
Cisco Nexus 5672
Cisco Nexus 5672
Cisco Nexus 9396
Metaswitch vRouter
Microsemi TimeProvider 5000
ADVA OSA 5421
Ericsson SSR 8004
Huawei NE40E-M2E
Juniper MX 80
Juniper MX 80
Huawei NE40E-X2-M8
Ixia IxNetwork
Spirent TestCenter N4U
Metaswitch vRouter
Juniper MX 240
Ericsson Virtual Router
Segment Routing
Nokia 7750 SR
Cisco Nexus 7702
Nokia 7750 SR
Microsemi Edge Assure 1000
ECI NPT 1800
Ixia IxNetwork
Huawei NE40E-X8A
Juniper MX 240
ADVA FSP150-GE114Pro
Huawei NE40E-X3
ECI NPT 1800
Metaswitch vRouter
Ericsson Router 8801
Spirent Attero
Huawei ATN 980B
Cisco IOS XRv
Cisco ISR 4000
IP/MPLS
Ericsson Router 6672
Data Center
Huawei ATN 950B
Huawei VNE
Ixia IxNetwork
Huawei NE40E-M2F
Cisco Nexus 9396
Meinberg LANTIME M4000
ADVA OSA 5401
Huawei ATN 910B
Calnex Paragon-X
Meinberg LANTIME M1000S
Ericsson Router 6672
Omnitron XM5
Microsemi TimeProvider 2300
Calnex Sentinel
Frequency
Microsemi TimeProvider 2700
Calnex Paragon-X
Meinberg LANTIME M1000S
Assisted Partial Timing Support (APTS)
Ericsson MINI-LINK 6691
Calnex Paragon-X
Omnitron XM5
Microsemi IGM 1100
Ericsson RBS 6501
Ericsson MINI-LINK TN
Microsemi TimeProvider 2700
Full Timing
Data Center
Access Network
Core Router
Emulator
Synchronous Node
Controller
Devices Managed by OpenFlow
IP/MPLS Core Network
Segment Routing Core Network
Access Device
Phase/Frequency Analyzer
Clock Node
Data Center Server
Devices Managed by PCE
Physical Link
Orchestration & Controllers
Impairment Tool
Virtual Network Function
Devices Managed by NETCONF

After that we sent a programmable network attack by using a traffic generator to generate a malformed UPDATE message to the device under test. We used the ORIGIN attribute in the IPv6 UPDATE message and introduced an undefined value (3) into it. The device under test was expected to remove only the IPv6 prefix from its table, without resetting the whole BGP session.

We tested this scenario using the Spirent TestCenter to send the Prefixes and the attack. We tested the Ericsson Virtual Router and the Ericsson Router 8801 and both devices behaved as expected: the IPv6 prefix was removed from the table and the IPv4 was left without resetting the session. The IPv4 traffic was not affected by the attack.

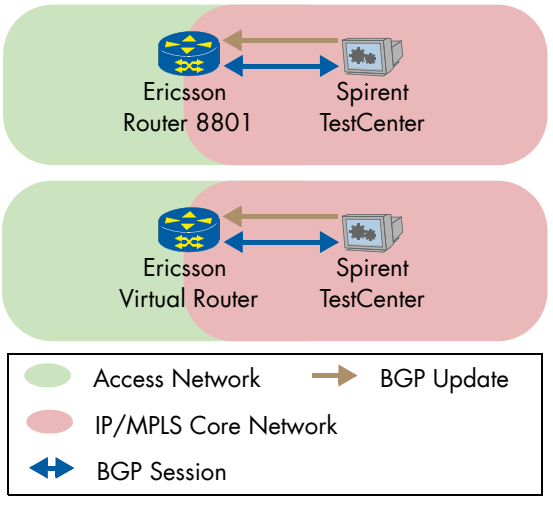


Figure 13: BGP Error Handling

Software Defined Networking

SDN is increasingly cementing its position at service providers as part of the virtualized network environment but there is still incomplete standardization or standardization that needs further clarification with regard to SDN and NFV. Many of the remaining issues facing SDN and virtualization in general center around network management with operators keen to avoid lock-ins to a single vendor by deploying open, interoperable network management systems.

Test scenarios in this area include rate-limiting, on-demand bandwidth management and bandwidth guarantees with OpenFlow, Segment Routing with Path Computation Element (PCE), BGP-LS and Path Computation Element (PCE) Integration.

This year, we built on the introduction of testing for NETCONF/YANG data models at last year's hot testing. Service provider interest in these technologies has increased and NETCONF/YANG models are regarded as less complex and therefore easy to understand and implement for production environments. Nevertheless, standardization and interoperability challenges are numerous.

OpenFlow

A range of testing was performed within OpenFlow environments and to test interworking between OpenFlow devices. We ran tests to verify that rate limiting can be applied to a specific flow using OpenFlow and that flow based bandwidth guarantee can be managed through the network using OpenFlow.

OpenFlow: Rate Limiting. Maintaining quality of service is a crucial task for service providers and therefore meters were introduced by the ONF in OpenFlow. Meters provide a way to classify the traffic so different policies can be implemented for different classes of traffic. Two meter types are present in OpenFlow: DROP and DSCP REMARK. For the DROP type the OF Forwarder simply drops the traffic when the band rate exceeds the chosen value. We didn't test the DSCP REMARK band type (which is optional in OpenFlow 1.3) because of lack of support from the vendors.

We successfully tested rate limiting using two meters. We verified that the OF Controller installed the correct flow entries and the corresponding meters for each class of traffic. We then generated bidirectional traffic at constant rate (100 Mbit/s for high priority and 250 Mbits/s for low priority) and checked that no traffic loss was observed.

In order to check the rate limiting functionality we sequentially increased the rate of both traffic classes and monitored that the additional traffic was dropped as expected.

We used two classes of traffic (high priority and low priority) and those classes were distinguished through the DSCP value as described in the following table.

Table 3: OpenFlow, Classes of Traffic

Traffic Class	DSCP Value	Per Direction Band Rate	Band Type
High	48	100[Mbit/s]	Drop
Low	0	250[Mbit/s]	Drop

Two pairs participated in the test, in each test we used ADVA FSP150-GE114Pro acting as OpenFlow Forwarder; Ixia IxNetwork and Spirent TestCenter separately participated in the tests as OpenFlow Controller.

OpenFlow: Bandwidth on Demand/ Bandwidth Guarantee. An application may have changing requirements and therefore OpenFlow provides a way to interface with customer applications. Service attribute modifications may be made on-demand (i.e., fulfillment requested immediately) or scheduled (i.e., fulfillment requested at a designated time in the future).

We successfully tested bandwidth guarantee without any dynamic meter change from the application (after a quota is achieved). We used the same classes of traffic as in the previous test. In order to check the bandwidth on demand functionality we doubled the rate of the high priority traffic and monitored that the additional traffic was not dropped during the provisioned time (5 minutes) as

expected. After five minutes we checked that half of the traffic was dropped as expected.

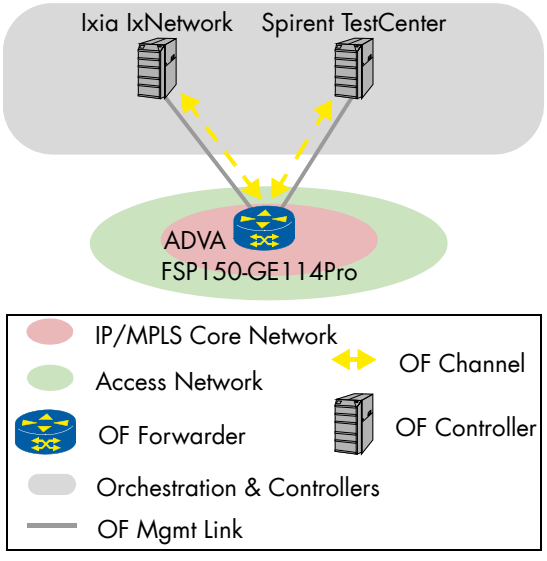


Figure 14: OpenFlow: Rate Limiting and Bandwidth Guarantee

We tested devices in two pairs: ADVA FSP150-GE114Pro acted as the OpenFlow Forwarder in both cases; Spirent TestCenter and then Ixia IxNetwork acted as the OpenFlow Controller. We observed that both Spirent and Ixia OF Controllers managed to install meters for a provisioned time period as expected. Both Spirent and Ixia OF Controllers provided a way to check the amount of data used for a particular flow but did not provide a way to dynamically change the meter band once the quota is consumed.

Path Computation Element Protocol

The Path Computation Element Protocol (PCEP) has been defined in RFC 5440 and specifies the communications between a Path Computation Client (PCC) and a Path Computation Element (PCE). Other drafts are developed to add stateful PCEP, PCE initiated LSP as well as extension to support Segment Routing LSP. A PCE is an entity (component, application or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints. A PCC is a client application that will request a path computation to be performed by a Path Computation Element. The PCE controller receives knowledge of the network topology via the Traffic Engineering Database (TED) and of the previously established paths via the LSP database (LSPDB). The Path Computation Element Protocol facilitates the deployment of Software Defined Networks. In the first two tests the PCEP is used to control RSVP-TE tunnels, in the third test it is combined with Segment Routing and in the last one it is combined with BGP-LS (in multiple domains).

PCE-initiated Paths in a Stateful PCE Model. In order to meet the changing demands of the applications running in an MPLS network, the PCE needs to be able to dynamically program the LSPs. The application can communicate its needs to the PCE and the PCE will program the traffic engineered paths accordingly, by instructing the PCCs to instantiate and signal the path. The PCE will tear down the path when it is not needed anymore.

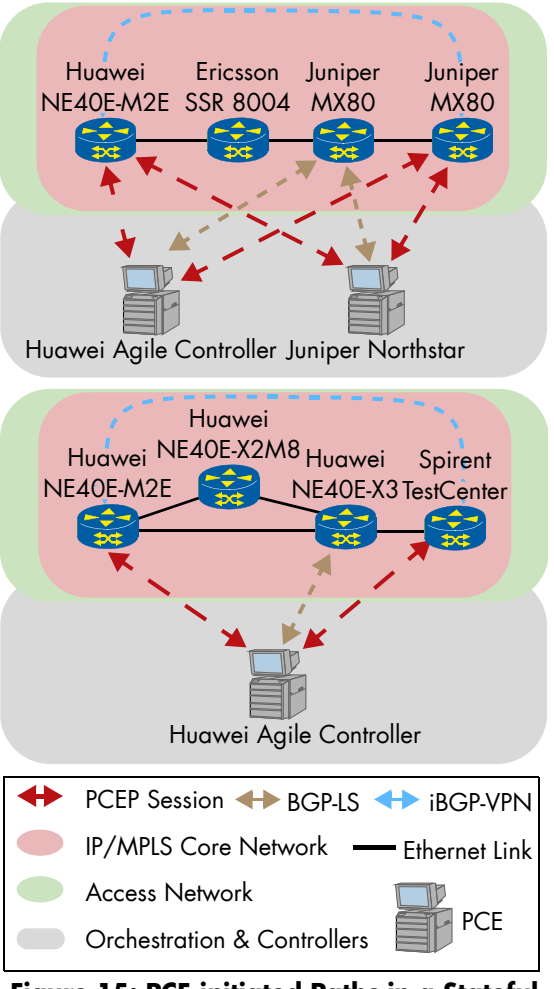


Figure 15: PCE-initiated Paths in a Stateful PCE Model

We tested that the PCEP session can be established between the PCE and the PCCs and that the PCE fully synchronised the LSP states as well as the TED (Traffic Engineering Database) information. We then verified that the PCE properly computed the paths between the PCCs and that the LSPs were installed on all nodes used by these paths. We later checked that the PCCs did not tear down the paths even if the PCE is disconnected from them; and that PCE synchronized with the current state of the network when the connection was brought back. Additionally, we assessed that the PCE can request a change of path and that the nodes on this path will receive the updates corresponding to the new path. Finally, we checked that the PCE can instruct the PCCs to tear down the paths and that LSPs are deleted as expected. We verified the paths by generating bidirectional traffic.

Three vendors participated in three combinations of the test: one with Juniper Northstar acting as PCE, Huawei NE40E-M2E and Juniper MX80 acting as PCCs, Ericsson SSR 8004 and another Juniper MX80 as Transit Nodes.

The second test was run with Huawei Agile Controller acting as PCE, Huawei NE40E-M2E and Juniper MX80 acting as PCCs, Ericsson SSR 8004 and another Juniper MX80 acting as Transit Nodes.

In both cases the L3VPN service was configured by Huawei and Juniper (Egress) and the controllers pushed the paths required for the service to work.

The third combination was tested with Huawei Agile Controller acting as PCE, Huawei NE40E-M2E, Huawei NE40E-X2M8, Huawei NE40E-X3 and Spirent TestCenter acting as PCCs. The Spirent TestCenter and Huawei NE40E-M2E created the L3VPN service.

In one combination, after a simulated PCC-PCE connection flap, one PCE did not manage to take back control of an orphan LSP. As expected, the remote PCC sent a PcRpt message with the D flag=0 after the relegation timer expired, but control of the LSP was never returned to the PCE.

PCC-initiated Paths in a Stateful PCE Model. In this test we tested creation, delegation, revocation and deletion of PCC-initiated LSP. We used a similar scenario to the previous test but this time we verified that the PCC could initiate a LSP delegation to the PCE. Later we verified that the PCC could trigger a change of path.

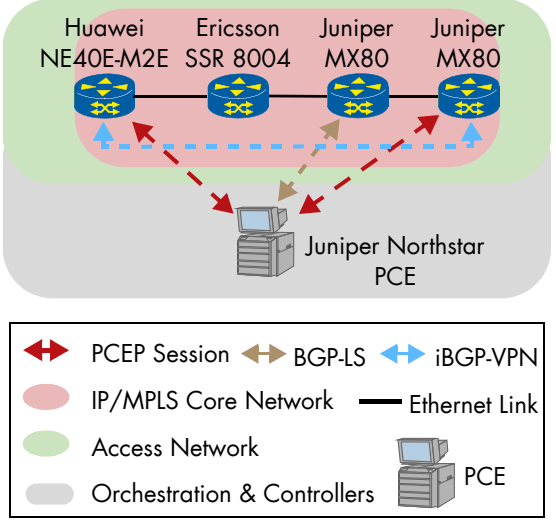


Figure 16: PCC-initiated Paths in a Stateful PCE Model

We tested Juniper Northstar as PCE, Huawei NE40E-M2E and Juniper MX80 as PCCs, Ericsson SSR 8004 and another Juniper MX80 as Transit Nodes. The L3VPN service was configured by Huawei and Juniper (Egress) and the controllers pushed the paths required for the service to work.

Segment Routing with Path Computation Element. The draft "PCEP extension for Segment Routing" from the IETF defines how the Segment Router Explicit Route Object (SR-ERO) can be used to carry a segment routing path. In this test we verified the setup of end-to-end service using a standard service control plane, while the transport is derived using segment routing without utilizing hop-by-hop signaling techniques (LDP or RSVP-TE).

The segment routing path is derived from a PCE controller. The PCE controller learns the network topology via the Traffic Engineering Database (TED) and previous established paths via LSP database.

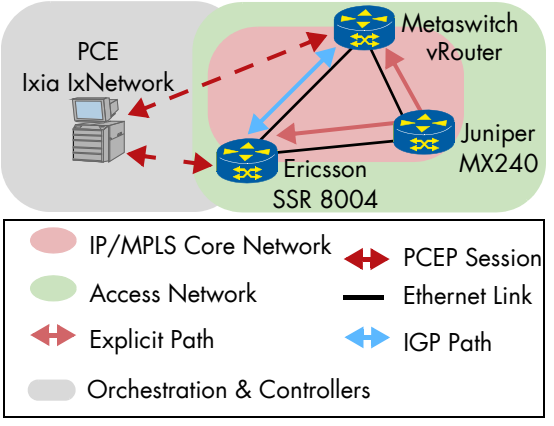


Figure 17: Segment Routing with Path Computation Element

We verified that the IGP information was correctly exchanged between the nodes and the PCEP session was established successfully between PCE and PCCs. We then checked that the PCE correctly computed the paths, using a shortest path scheme first and that the paths were correctly installed on the PCC nodes. Later we verified that the PCE could change the paths, using an explicit path different from the shortest path and finally that it could tear down the paths. Throughout the test we tested the paths by generating bidirectional traffic.

We tested the following combination: Ixia IxNetwork was acting as PCE, Ericsson SSR 8004 and Metaswitch vRouter were acting as PCCs and Juniper MX240 was a transit node.

We ran the test twice so that every device could be tested as part of the transit and the access network. During the first run Ixia generated traffic between Ericsson and Metaswitch and the traffic first went through the shortest path and then through Ericsson, Juniper and Metaswitch for the PCE-initiated change of path. During the second run Ixia generated traffic between Metaswitch and Juniper and the traffic first went through the shortest path and then through Metaswitch, Ericsson and Juniper for the PCE-initiated change of path. Some vendors supported the latest version 6 of the draft and some other vendors only supported version 5 and therefore some combinations could not be tested.

BGP-LS and Path Computation Element Integration. In this test we verified that Path Computation Element (PCE) combined with BGP Link State (BGP-LS) can be used to setup optimal end-to-end traffic engineering (TE) LSP across multiple network domains. An inter-domain TE LSP is a LSP that transits through at least two network domains. Topology, Traffic Engineering Database (TED) and Link State Database (LSPD) remain locally visible to a given network domain, and a head-end network node cannot compute an inter-domain end-to-end path. One key application of the PCE based architecture is the computation of inter-domain TE LSP.

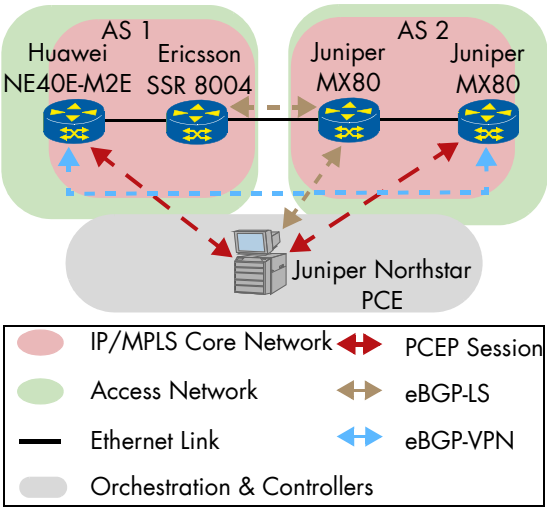


Figure 18: BGP-LS and Path Computation Element Integration

First we assessed that the IGP information was correctly exchanged between the nodes, that the PCEP session was established and that the PCE retrieved the TED information from each AS domain. We then verified that the BGP-LS session was established between both domains. Later we tested that the PCE correctly computed the inter-AS constrained shortest path and sent it to the network nodes. Finally, we verified that the PCE could tear down the paths. We tested the paths throughout by generating bidirectional traffic.

One vendors's combination successfully participated in this test. Juniper Northstar was acting as PCE, Huawei NE40E-M2E and Juniper MX80 were acting as PCCs, Ericsson SSR 8004 and another Juniper MX80 were transit nodes. There were eBGP-LS sessions between Ericsson SSR 8004 and the first Juniper MX80 and between this latter and Juniper Northstar. The L3VPN service was configured by Huawei and Juniper (Egress) and the controllers pushed the paths required for the service to work.

NETCONF

The NETCONF protocol defines a simple mechanism through which a network device can be managed, configuration data information can be retrieved and new configuration data can be uploaded and manipulated. It is defined in RFC 6241. The protocol allows the device to expose a full, formal application programming interface (API). Applications can use this straightforward API to send and receive full and partial configuration data sets. YANG is the data modeling language used by NETCONF and it is defined in RFC 6020. YANG provides ways to define configuration and state parameters, to represent complex data models such as list and unions and also supports nested data definitions.

Device Configuration Using NETCONF/YANG. Through this test we verified NETCONF/YANG functionality to manage configuration and operational state on a network device.

We tested that the NETCONF session was established, that the YANG schema (when the GET-SCHEMA operation is supported) and the running configuration could be retrieved. We also verified the usage of the subtree filtering to retrieve the running interface configuration. Finally, we assessed a configuration change and then a change deletion on the device.

Six pairs were tested. Each time Cisco NSO enabled by Tail-f was the NETCONF Controller and we tested it against ADVA FSP150-GE114Pro, Cisco IOS XRv, Cisco ISR 4000, Ericsson Router 6672, Juniper MX240 and Microsemi EdgeAssure 1000. One vendor only supported the read (and no write) operation

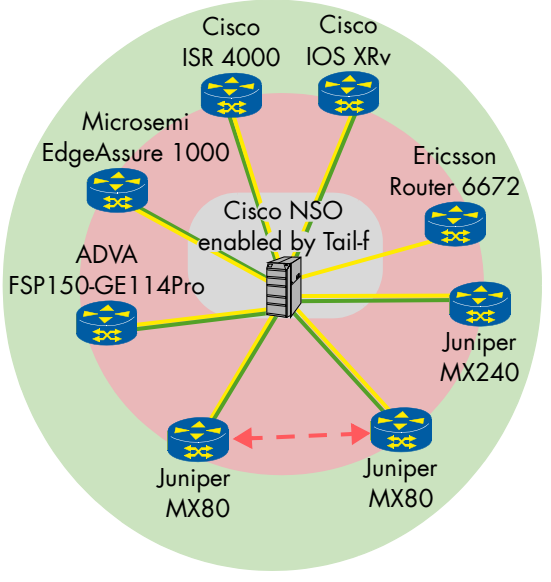


Figure 19: Device Configuration and L3VPN Service Creation Using NETCONF/YANG

L3VPN Service Creation Using NETCONF/YANG. The IETF's L3 Service Modeling Group (L3SM) introduces service models in YANG to provision BGP PE-based L3VPNs, described in RFC4110 and RFC4364, using NETCONF. The models aim to create a heterogeneous configuration layer which can simplify service creation in a multi-vendor environment. In this test a NETCONF compliant client is used as a centralized controller to configure a group of PE nodes and provision a predefined list of L3VPN services. The scenario is similar to the previous one but now we focused on the verification of the creation of the L3VPN service.

One pair successfully participated in this test: Cisco NSO enabled by Tail-f acted as the NETFCONF Controller and two Juniper MX80 acted as NETCONF Servers.

L2 Service Performance Monitoring Using NETCONF/YANG. Service Performance is a service attribute which allows monitoring of the performance received from an Ethernet Virtual Connection (EVC). The ITU-T specification Y.1731 introduces message types which are defined to measure frame loss, frame delay and frame delay variation for a point-to-point Ethernet connection. This test focused on these three performance attributes. The frame delay and frame delay variations can be measured via two different methods: one-way or two-way. The one-way measurement requires precisely synchronized clocks on the participating devices. This can be achieved via a common external clock (e.g. GPS), or via packet based synchronisation protocols like IEEE 1588 or NTP. NETCONF controller/client started the execution of performance monitoring measurement as well as read and display the measured values. NETCONF was used to read the Performance Monitoring Statistics from the devices.

Within this test we verified the following features:

- Loss measurement over a point-to-point EVC and per CoS ID
- One-way delay and delay variation over a point-to-point EVC

We successfully tested Frame Delay measurement and Frame Delay Variation measurement (Two-way ETH-DM) per point-to-point EVC and CoS ID with Cisco NSO enabled by Tail-f acting as NETCONF controller and two Microsemi EdgeAssure 1000 devices acting as NETCONF servers. Spirent Attero was used as impairment tool.

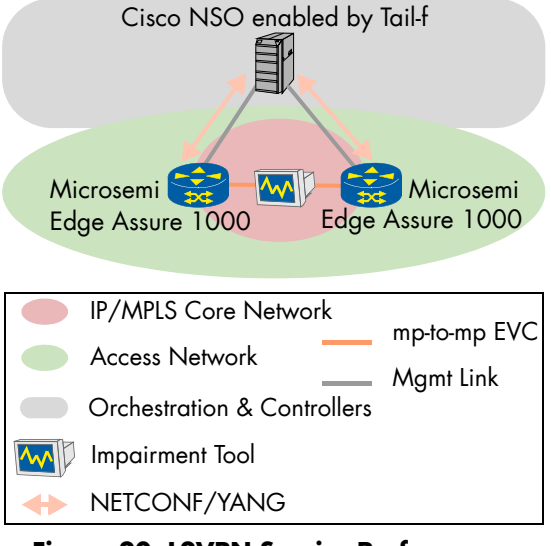


Figure 20: L2VPN Service Performance Monitoring using NETCONF/YANG

Clock Synchronization

This year, the ninth year of testing clock synchronization in our interoperability event, was marked by a rather unusual event. On Tuesday, January 26, we experienced first-hand the effect of a *GPS Ground System Anomaly,* as reported by an official press release from the US Air Force. A timing expert from one of the vendors used his GPS analysis software to determine the root cause: A couple of space vehicles were sending wrong UTC correction values, causing an offset amounting to 13 µs.

The effect of the anomaly manifested itself differently for each GPS receiver: some lost GPS lock as soon as the signals from these space vehicles were received, while others exhibited phase jumps (transients).

This issue, albeit rare, joins GPS jamming and GPS spoofing in highlighting two of our tests in this year — the assisted partial timing support providing for GPS-based synchronization using the Precision Time Protocol (PTP, IEEE standard 1588-2008) protocol.

Our clock synchronization testing program this year was focused on performance in optimal and suboptimal conditions: delay asymmetry, hold-over performance for phase and frequency, source failover between two different grandmasters for phase and frequency as well as microwave transport with adaptive modulation.

As always, we based our quality requirements on the recommendations of the ITU-T and the end applications. We considered applications for modern mobile networks, which include Time Division Duplex (TDD), enhanced Inter-cell Interference Coordination (eICIC), Coordinated Multipoint (CoMP) and LTE Broadcast. We borrowed the accuracy level of ±1.5 µs (ITU-T recommendation G.8275 accuracy level 4) as our end-application goal, and we defined 0.4 µs as the phase budget for the air interface. The requirement on the network limit, the last step before the end-application, had to be therefore ±1.1 µs.

For frequency synchronization, we continued using the G.823 SEC mask as a requirement, with the exception of the frequency hold-over tests, where we used the G.8261.1 option 3 mask. The primary time reference clock was GPS using an L1 antenna located on the roof of our lab.

Phase/Time Assisted Partial Timing Support: Delay Asymmetry

Using GPS, a synchronization accuracy of ±100 ns can be expected, an order of magnitude better than the ±1.1 µs requirement for LTE-Advanced applications. In case GPS is jammed or its signal weakened by severe weather conditions, PTP can be used in a mode called assisted partial timing support to provide backup to the local GPS from a separate reference source.

The goal of the test is to verify that a slave clock can maintain the required synchronization quality when its GPS is temporarily unavailable and the network delay asymmetry changes. To emulate the path lacking PTP support, we used a PDV profile according to G.8261 test case 12.

We started the test by allowing the boundary clock to acquire lock via GPS (higher priority), while also PTP signal (lower priority) was available. We then

disconnected the GPS antenna, forcing the boundary clock to switch to PTP. Afterwards we changed the delay asymmetry with the impairment tool, while GPS remains disconnected.

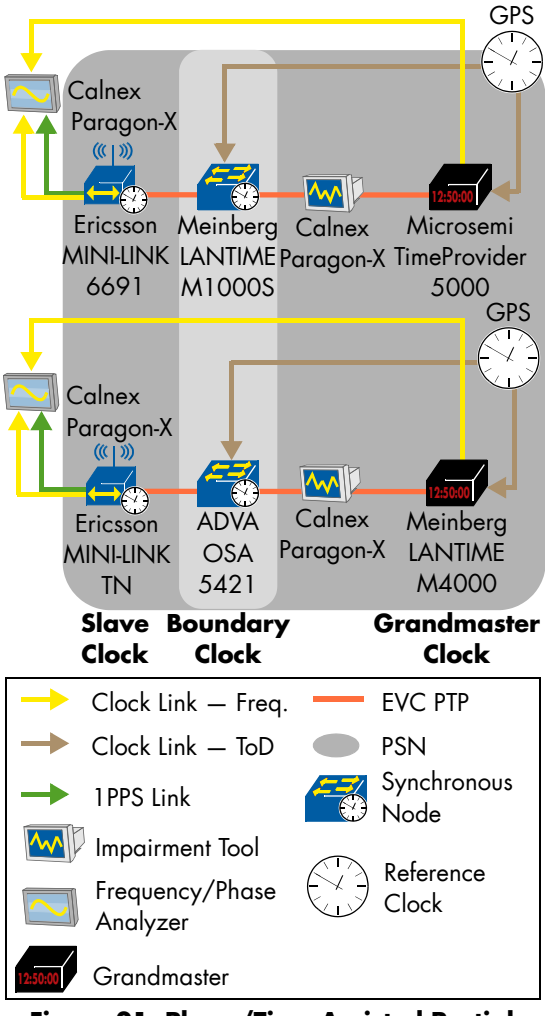


Figure 21: Phase/Time Assisted Partial Timing Support: Delay Asymmetry

All test runs depicted in the diagram complied with the ±1.1µs absolute phase error requirement as well as the G.823 SEC frequency mask.

Meinberg LANTIME M4000 and Microsemi TimeProvider 5000 passed as grandmaster clocks; ADVA OSA 5421 and Meinberg LANTIME M1000S passed as boundary clocks; Ericsson MINI-LINK 6691 and Ericsson MINI-LINK TN passed as slave clocks.

Hold Over Performance

In this section we will describe two Hold Over performance tests. The hold over time is considered to be the longest period that the slave clock maintains the required accuracy. The measurements were performed over night with the evaluated duration of 12 hours.

Phase/Time Assisted Partial Timing Support. In this setup we started from a free running situation, then let the slave synchronize by GPS and via PTP with the Grandmaster clock, without any physical frequency reference – such as TDM or SyncE. After disconnecting the GPS antenna, we emulated a PTP impairment and verified the holdover performance of a slave clock in relation to phase/time stability. In this test we used the Calnex equipment to emulate a packet delay variation (PDV) according to the profile defined in G.8261 test case 12.

After the overnight measurement, we removed the impairment on the PTP stream and verified the transient response of the slave clock matched the requirements after it re-locks to the grandmaster.

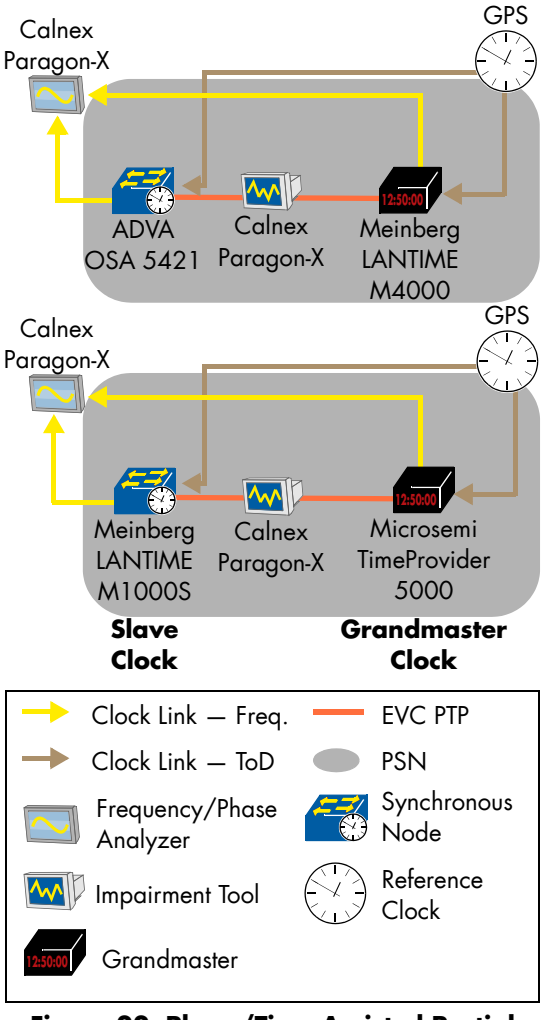


Figure 22: Phase/Time Assisted Partial Timing Support: Hold Over Performance

All test runs depicted in the diagram complied with the ±1.1µs absolute phase error requirement as well as the G.823 SEC frequency mask.

Meinberg LANTIME M4000 and Microsemi TimeProvider 5000 passed as grandmaster clocks; ADVA OSA 5421 and Meinberg LANTIME M1000S passed as slave clocks.

Frequency Synchronization: Hold Over Performance. We started from a free running situation, then let the slave synchronize via PTP with the Grandmaster clock, without any physical frequency reference – such as TDM or SyncE. We emulated a PTP impairment and verified the holdover performance of a slave clock in relation to frequency stability.

In this setup we introduced the impairment by disconnecting the cable to the Grandmaster and thus breaking the PTP signal.

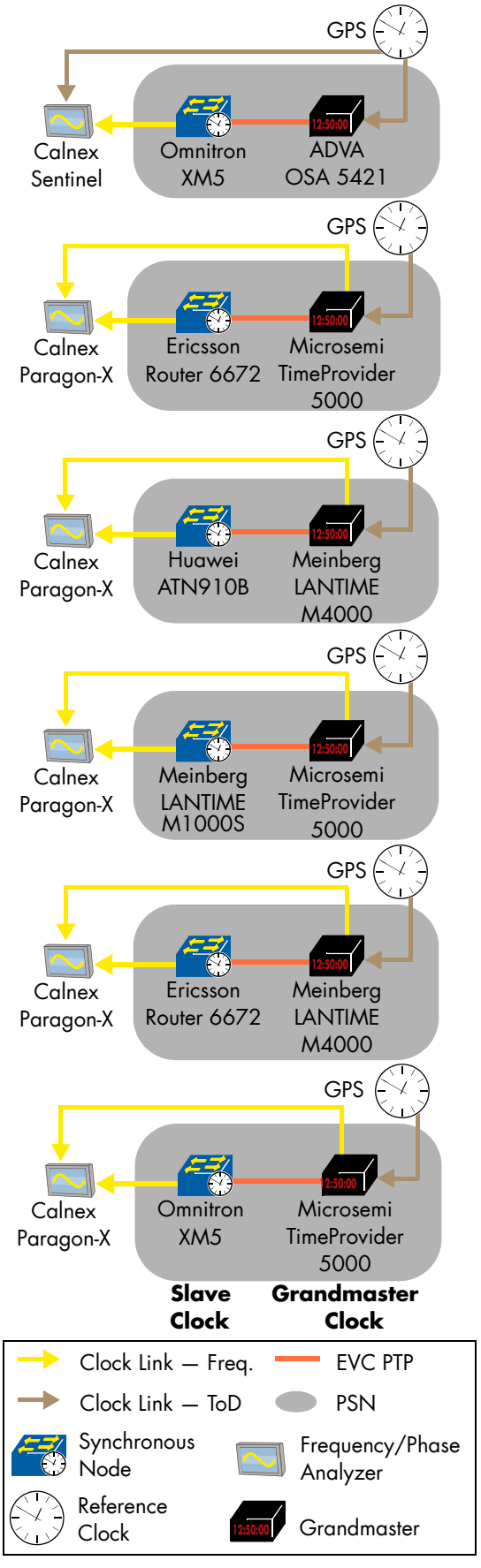


Figure 23: Frequency Synchronization: Hold Over Performance

After the overnight measurement, we removed the impairment on the PTP stream and verified the transient response of the slave clock matched the requirements after it re-locks to the grandmaster.

We observed one implementation that did not enter free-running mode and retained stable frequency compared to GPS even after cold-starting the device by removing power supply; although it is a positive feature of the implementation, we were unable to start the test with this device, since we require all slave clocks to start from free-running mode.

In the last step of the test, when the slave clock was switching from hold-over into locked mode, we observed a transient over 30 µs with the Omnitron XM5 as slave clock in both of its test runs.

All test runs depicted in the diagram complied with the G.8261.1 option 3 frequency mask.

ADVA OSA 5421, Meinberg LANTIME M4000 and Microsemi TimeProvider 5000 passed as grandmaster; Ericsson Router 6672, Huawei ATN910B, Meinberg LANTIME MM1000S and Omnitron XM5 passed as slave clock.

Source Failover

A slave clock may be connected to more than one grandmaster through a boundary clock. This is for resiliency purposes, protecting both against failure of the primary grandmaster, and also against failure of the GPS antenna connected to that grandmaster.

The goal of the following two tests was to verify that a slave clock maintains the required clock synchronization frequency (first test) and phase/time (second test) quality when it switches over from its primary to its secondary grandmaster following a signal degradation.

In these tests, both grandmasters were provided with a GPS signal. We allowed the slave clock to lock to the primary grandmaster and then degraded the primary grandmaster's quality by disconnecting its GPS input and measured the slave clock's transient response.

We also verified the correct clockClass values are being signalled by the grandmasters according to the telecom profiles, which allows the alternate best master clock algorithms running on the slave clock to correctly select the best grandmaster during each step of the tests.

Frequency Synchronization . This test was performed with the G.8265.1 profile.

We observed an interoperability issue of clockClass values with two different implementations. One implementation used the option 1 value (QL-PRS, clockClass 80) and the other used the option 2 value (QL-PRC, clockClass 84). QL-PRS is employed for hierarchies based on T1, common in America and Japan while QL-PRC is employed for hierarchies based on E1, common in Europe.

All test runs depicted in the diagram complied with the G.823 SEC frequency mask and the G.8265.1 BMCA.

ADVA OSA 5421, Meinberg LANTIME M4000, Microsemi TimeProvider 2700 and Microsemi TimeProvider 5000 passed as grandmaster clock; Ericsson MINI-LINK 6691, Ericsson Router 6672, Huawei ATN910B and Meinberg M1000S passed as slave clocks.

We observed one implementation that delayed its switchover, since it was only regarding announce messages (which are slower than sync messages) as input to its BMCA. The vendor fixed the issue during the hot staging.

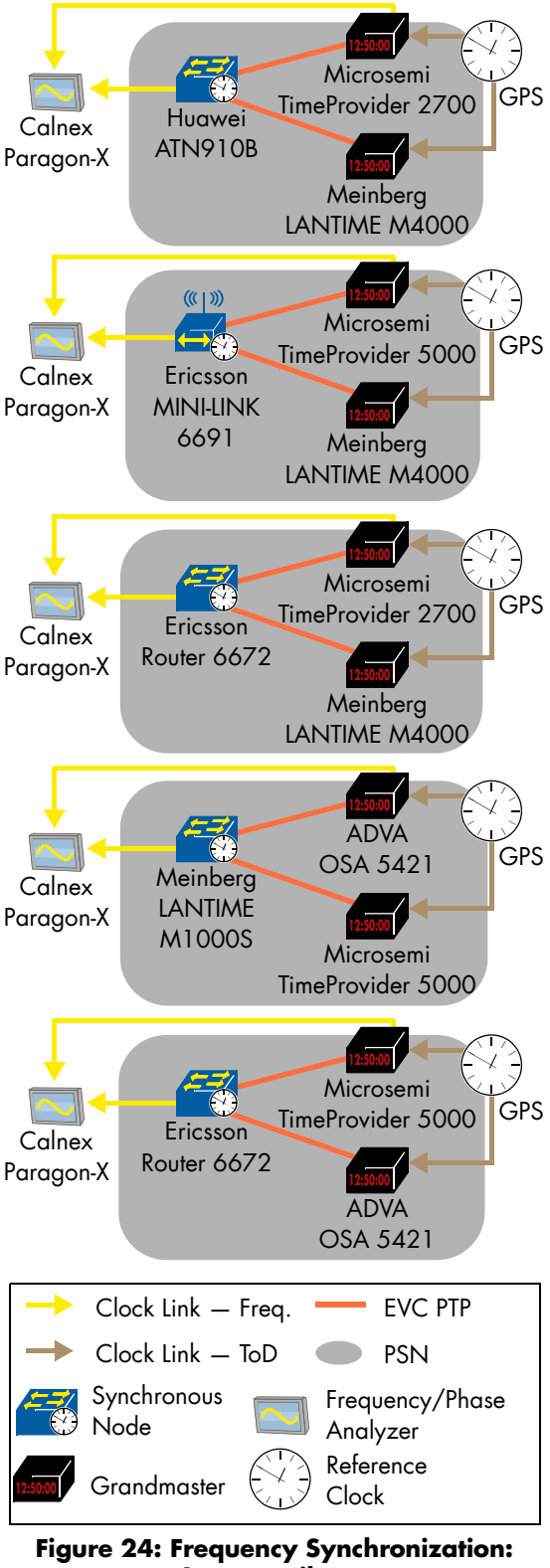


Figure 24: Frequency Synchronization: Source Failover

Phase/Time Synchronization. This test was performed with the G.8275.1 profile.

All the devices with a 1 pps output (Ericsson Router 6672, Ericsson MINI-LINK 6691, Huawei ATN 910B and Omnitron XM5) met the ±1.1µs absolute phase error requirement, all the devices complied with the G.823 SEC frequency mask and the G.8275.1 BMCA.

ADVA OSA 5421, Meinberg LANTIME M1000S, Meinberg LANTIME M4000, Microsemi TimeProvider 2700 and Microsemi TimeProvider 5000 passed as grandmaster clock; Ericsson MINI-LINK 6691 and Ericsson Router 6672 passed as boundary clock; Ericsson MINI-LINK 6691, Ericsson Router 6672, Ericsson RBS 6501, Huawei ATN 910B and Omnitron XM5 passed as slave clock.

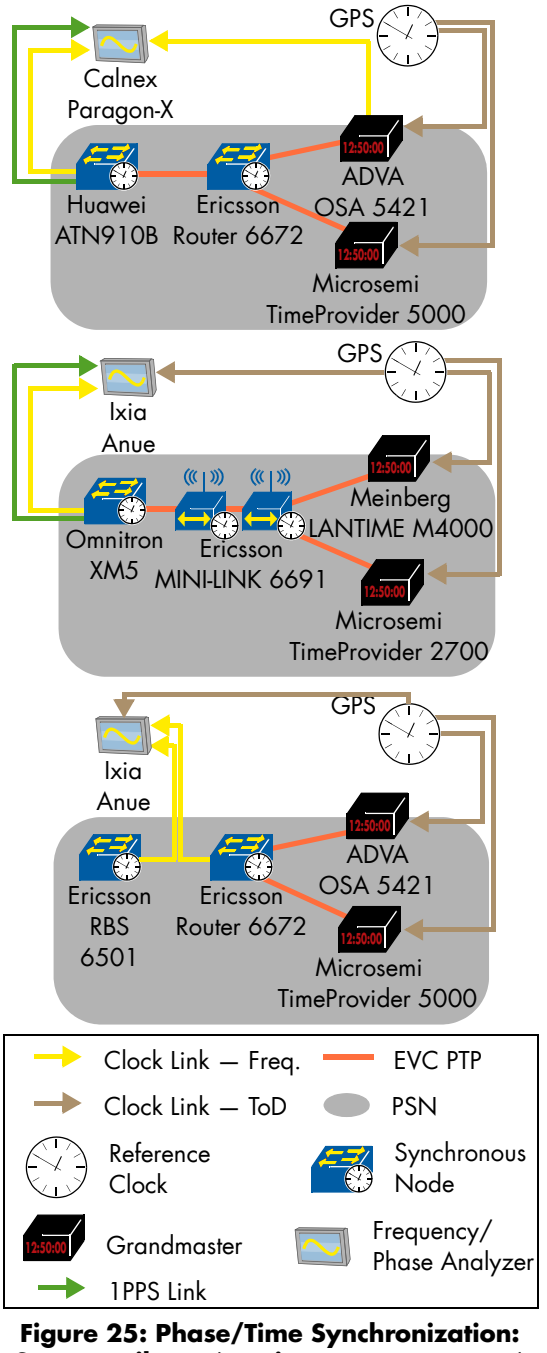


Figure 25: Phase/Time Synchronization: Source Failover (continues on next page)

We observed a premature switchover when the GPS source was disconnected from ADVA OSA 5421, as its hold-over period for phase is short; according to the ADVA engineer, this happened because its holdover performance parameter (this is user configurable) was set to a lower value than the default one.

After a short period of time, it switched from clock-Class 7 to 160. This behavior conforms to G.8275.1.

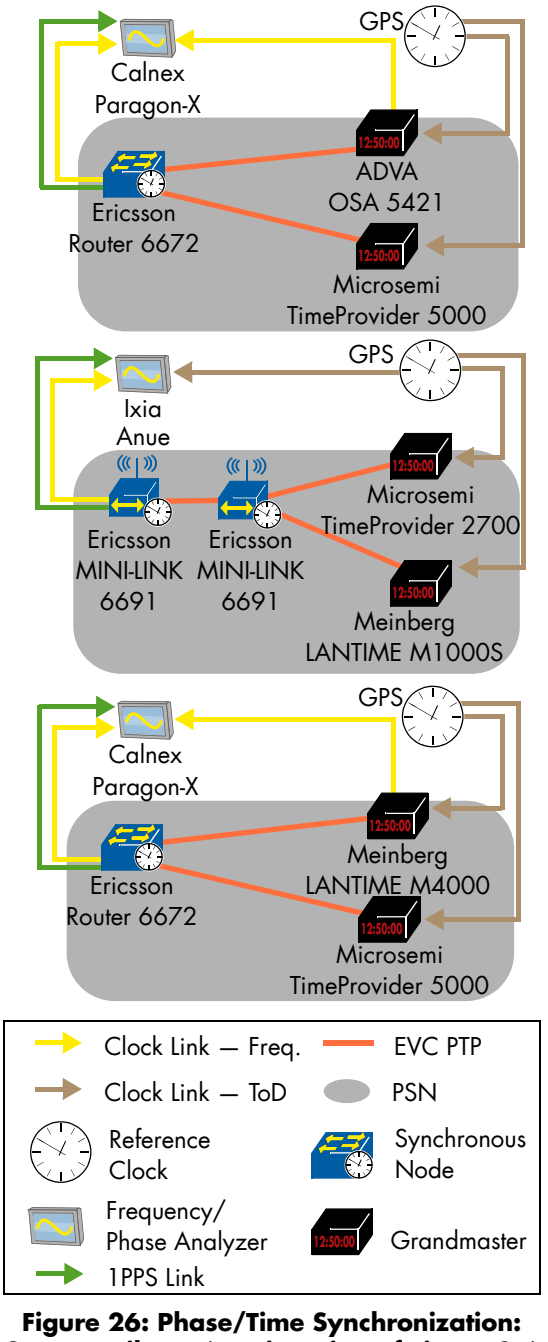


Figure 26: Phase/Time Synchronization: Source Failover (continuation of Figure 25)

Phase/Time Synchronization with Full Timing Support

Microwave Transport . A microwave system may undergo conditions that cannot be controlled by the network operation, such as severe weather conditions. When a microwave system is used as a transport for timing distribution, it is critical for the system to prioritize timing packets and to compensate for the delay variation when another modulation is chosen (and thus the throughput is reduced).

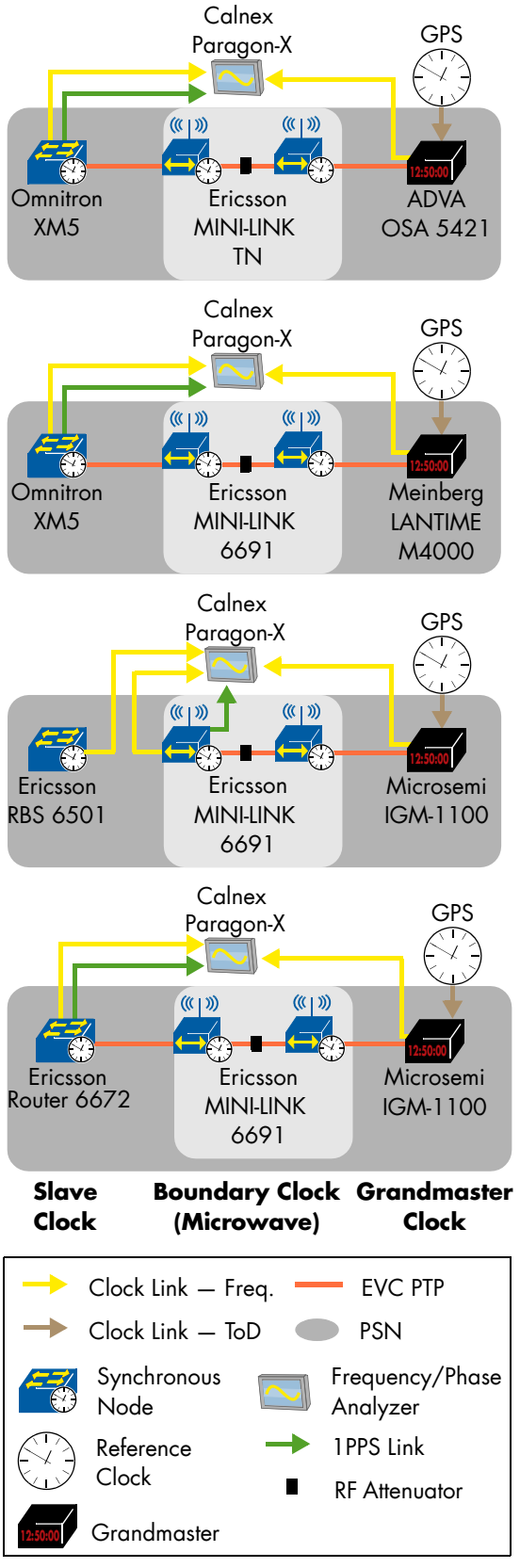


Figure 27: Phase/Time Synchronization with Full Timing Support: Microwave Transport

We started the test with the slave clock in free-running mode and generated traffic according to

G.8261 VI2.2 at the maximum line rate for the maximum modulation scheme and expected no traffic loss. After clock lock we took baseline measurements for phase and frequency lock from the slave clock. 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 (4096QAM to 4QAM for MINI-LINK 6691 and 1024QAM to 16QAM for MINI-LINK TN).

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 verified that data packets were dropped according to the available bandwidth.

All the test runs met the G.823 SEC frequency mask requirements. In addition, all the devices with a 1 pps output (Ericsson Router 6672, Ericsson MINI-LINK 6691, Huawei ATN 910B, and Omnitron XM5) complied with the ±1.1µs absolute phase error requirement.

OSA 5421, Meinberg LANTIME M4000 and Microsemi IGM-1100 passed as grandmaster clock; Ericsson Router 6672, Ericsson RBS 6501 and Omnitron XM5 passed as slave clock; Ericsson MINI-LINK 6691 and MINI-LINK TN passed as a microwave system.

Demonstration Scenarios

Point-to-Point EVPN. EVPN can be used to support virtual private wire service (VPWS) in MPLS/IP networks. EVPN enables the following characteristics for VPWS: single-active as well as all-active multi-homing with flow-based load-balancing, eliminates the need for single-segment and multi-segment PW signaling, and provides fast protection using data-plane prefix independent convergence upon node or link failure.

Nokia demonstrated this scenario with a pair of Nokia 7750SR in a single homing setup. Behind each device was an Ixia Traffic Generator. We tested at each side that the BGP EVPN Auto-Disc Routes were carrying the expected values (Route Distinguisher, IP of Next-Hop, Label). We then sent bidirectional unicast and multicast traffic and observed no packet loss. Upon a link failure, we tested the withdrawal the AD route.

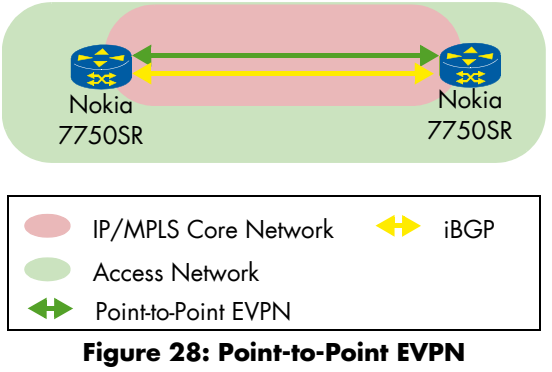


Figure 28: Point-to-Point EVPN

Label Switched Path (LSP) Ping/Trace for Segment Routing. The IETF draft (kumarkini-mpls-spring-lsp-ping) defines the LSP ping and traceroute method to Segment Routing (SR) on the MPLS data plane.

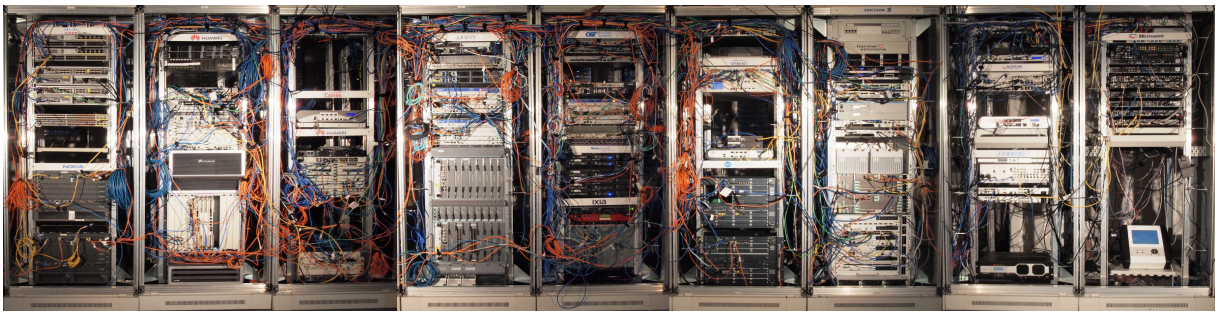
Ericsson demonstrated that LSP echo request and echo reply were performed on a segment routing network over MPLS data plane. The devices under test were Ericsson Virtual Router and Ericsson SSR 8004.

PCE-initiated Paths in a Stateful PCE Model. We verified that PCE triggered a change of path in the PCEP network.

Juniper demonstrated the control plane feature with Juniper Northstar acting as PCE. The Spirent TestCenter acted as PCC. We observed that PCEP session was established. Spirent PCC received the path change notification triggered by the PCE.

PCC-initiated Paths in a Stateful PCE Model. We verified that PCC triggered a change of path in the PCEP network.

Juniper demonstrated the feature with Juniper Northstar acting as PCE and Spirent TestCenter acting as PCC. We observed that a PCEP session was established. Juniper PCE received the path change notification triggered by Spirent PCC.



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

20160302 v5