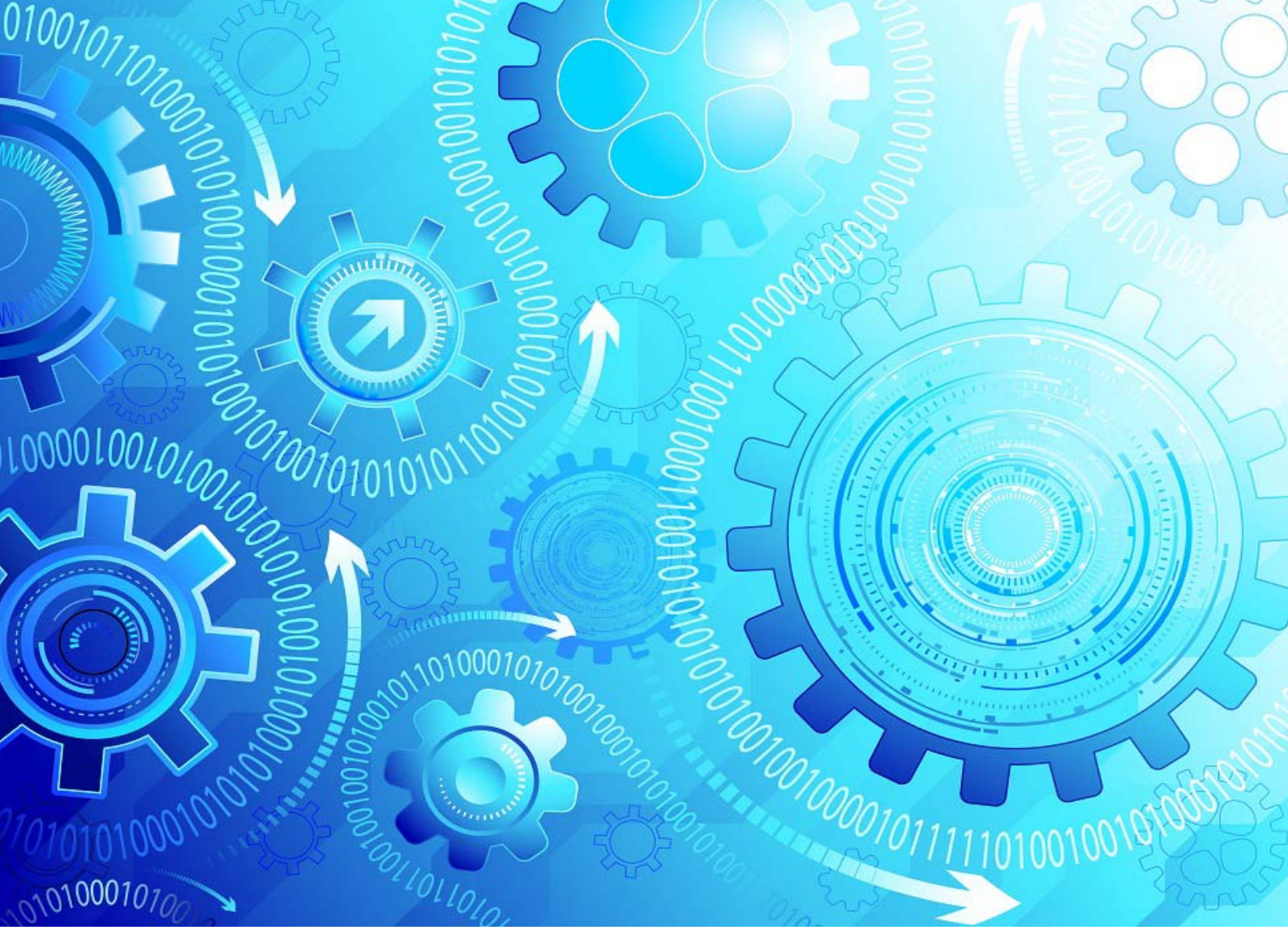


EANTC

MPLS & Ethernet World Congress 2012

Public Multi-Vendor Interoperability Event

White Paper

Editor's Note

Since our first interoperability event at the MPLS World Congress a decade ago, packet transport technology has constantly redefined itself. The industry has adopted new service provider requirements, standardized even its advanced solutions and multi-vendor interoperability has been achieved. This way, IP VPN services, Carrier Ethernet, mobile backhaul, and network management aspects have been incorporated into MPLS. We are at a point where the technology has reached an unequalled level of maturity.

It may be surprising to the uninitiated that the conference is more crowded than ever this year; the agenda is bursting with new topics such as cloud services; and the sister conference, IPv6 World Congress, has grown substantially as well. In fact, the reason is that service providers have redefined requirements and ramped up their expectations yet again.

MPLS and Carrier Ethernet are now used extensively for mobile backhaul and mobile core networks, supporting the huge migration to data-heavy mobile services and LTE/4G networks. Service providers compete in deploying cloud services as fast and reliable as possible. Some are even getting ready for the future of IPv6 networking.

Our annual interoperability events in Paris reflect the market developments:

- We have again seen substantial interest in mobile backhaul testing as before, including the transport protection aspects.
- IPv6 migration has been accepted as a test topic for the first time, hoping for a fast start given the age of the technology.
- And there is again some innovation in the MPLS-TP space.

Thirteen vendors participated this time, bringing aggregation, access, CPE, and test equipment to our two-week hotstaging test in Berlin, Germany in January. Some test areas went really well, for example the Ethernet Ring Protection (ERPS) tests. Others — as the reader will witness reading our report — could have gone better.

Why is that? Quality assurance bears some basic truths. If a standard has been well-defined, the vendors have tested with each other from the beginning, and there are no unreasonable performance expectations, there is a good chance to achieve a stable plateau of interoperability quickly. This is the case for ERPS, for example.

Interestingly, we saw hiccups in IPv6 which has been around for a long time. It is obvious that only a few vendors have deployed IPv6 in production environments at this point. Also, there is a variety of migration scenarios; it is a challenge to support them all.

Synchronization in mobile backhaul networks is a different story. Meanwhile most vendors worldwide have submitted their implementations to our interop tests since we started with this test topic in Paris in February 2008. The experience of the past four years is that Synchronous Ethernet works great in multi-vendor environments, but IEEE 1588:2008 packet-based synchronization remains a challenge. It is simply a very demanding, non-trivial problem that is solved by IEEE 1588:2008 — and there is no alternative to using this protocol.

Progress in sync testing has been made again this time: We were able to show the complete chain — master clock, boundary clock, transparent clock, and slave clock — as a multi-vendor solution in our test bed. That said, we will continue our interoperability test program in this area.

We hope you like this report and look forward to meeting some of our readership in Paris. As usual, we are open for feedback and suggestions!

Table of Contents

Introduction

EANTC interoperability test events are guided by a test plan, which we began developing in September 2011 for this event. Discussions with service providers, vendor requests, and the experience of EANTC's testing team contribute to the test areas. Through a series of conference calls with sometimes more than 50 attendees from all interested equipment manufacturers, the test areas are mutually agreed upon. In the process, EANTC defines test cases for each of the test areas.

Later on, once vendors commit their participation to a test, the EANTC team creates a configuration guide based on protocol support of each device under test, the vendors' interest in specific test combinations, and physical layer aspects.

More than 30 support engineers from all participating vendors then met at EANTC's lab in Berlin, Germany, for an intense two-week hotstaging (January 9–20). They tested their solutions with each other in an open and constructive environment, covered by a multi-party Non-Disclosure Agreement.

We strictly require validation of all test results by EANTC staff as a precondition for being documented in this white paper. Vendors generally welcome this requirement since it ensures a leveled playing field for all. EANTC has set industry standards by the extent of our planning, execution and documentation efforts.

Interoperability Test Results

The following sections of the white paper describe the test areas and results of the interoperability event. The document generally follows the structure of the test plan — a document edited by EANTC and reviewed together with vendors in preparation for the event.

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 terminated by equipment from a single vendor on both ends.

Test Equipment. In order to perform our tests we had to generate, measure, impair, and analyze Ethernet and MPLS traffic and perform synchronization analysis. We thank Calnex Solutions, Ixia, Spirent Communications, Symmetricom and VeEX for their test equipment and support throughout the hot staging.

Participants and Devices

Vendor	Devices
Calnex	Paragon-X
Cisco	ASR 1002 ASR 9006 Linksys E4200
Ericsson	MINI-LINK SP 110 MINI-LINK SP 210 MINI-LINK SP 310 SPO 1410 SPO 1460 SmartEdge 100 SmartEdge 1200
Extreme Networks	E4G-200 E4G-400
Hitachi	AMN1710
Huawei	ME60 NE40E Home Gateway
Ixia	ImpairNet IxNetwork XM12
Metaswitch	Metaswitch DC-PCE
Spirent Communications	Spirent Anue 3500 Spirent TestCenter Spirent XGEM
Symmetricom	Cesium Reference CsIII SSU 2000e TimeProvider 500 TimeProvider 1500 TimeProvider 5000 TimeProvider 5000 Expansion TimeProvider E10 TimeProvider E30
Telco Systems	EdgeGenie T5C-XG T-Marc 340 T-Marc 3208SH T-Metro 7124S T-Metro 7224
VeEX	TX130M
ZTE	ZXR10 M6000-3S ZXR10 M6000-5S ZXR10 M6000-8S

IPv6 Migration

The introduction of IPv6 as a technology happened many years ago. RFC 2460, titled "Internet Protocol, Version 6" was published in December 1998. It was not until recently that service providers, content providers and vendors started to show a real interest in IPv6, interest that was surely fueled by the imminent IPv4 address exhaustion, as the Internet Assigned Numbers Authority (IANA) assigned the last 5 remaining address blocks to the five Regional Internet Registry (RIR) authorities on February 3rd 2012.

The "World IPv6 Day" event, which took place on June 8th 2011, demonstrated the commitment of top websites and Internet service providers around the world to the promotion of IPv6. At that day, IPv6 was temporary enabled for 24 hours - in this year on June 6th during the "World IPv6 Day" event IPv6 will be permanently enabled on more than 400 websites, and by at least 18 Internet service providers (see www.worldipv6launch.org for a current list of participants and additional information).

In light of the increasing importance of IPv6 implementations and the real need for migration strategies and insight, we focused our IPv6 interoperability testing on two aspects - IPv6 migration technologies and IPv6 control protocols implementations.

The question, "how do we migrate to IPv6?" cannot be answered by a one-size-fits-all solution. A number of technologies were introduced in order to address the different needs of service providers and to

provide a strategy for common scenarios. Service providers who plan to keep maintaining IPv4-only in they network will be interested in IPv6 Rapid Deployment on IPv4 Infrastructures (6rd). Providers who plan to maintain both IPv4 and IPv6 protocols in their entire network follow the Dual Stack (DS) strategy. And finally the service providers that migrate completely to IPv6-only network whilst still providing both IPv4 and IPv6 access, may follow the Dual Stack Lite (DS-Lite) strategy.

We see a number of vendors that are just entering the IPv6 market and also a number of new products from vendors that have already entered the IPv6 market a while ago. Therefore, we see a need for interoperability testing of IPv6 control protocols even if one could consider them mature.

In this event we verified the three major IPv6 migration scenarios — 6rd, DS, and DS-Lite. We also verified implementations of DHCPv6, ICMPv6, Neighbor Discovery, and Port Control Protocol (PCP) - all essential for hosts using IPv6.

We believe that a complete, end-to-end IPv6 migration scenario require customer premise equipment - CPEs. Therefore we procured an AVM Fritz!Box and D-Link CPEs for the test, and used them in a number of IPv6 tests. These CPEs were not supported in the test by their respective manufacturers.

6rd: Rapid IPv6 Deployment using IPv4 Infrastructure

The IETF has addressed the important question of how to transition from IPv4 to IPv6 in a number of RFCs. 6rd, originally described in RFC 5569 and standardized in RFC 5969, is a facility to transport IPv6 traffic over an IPv4 network, enabling IPv6 devices to communicate.

6rd consists of two main components: 6rd Customer Edge (6rd CE) and 6rd Border Router (6rd BR).

- The 6rd CE provides Dual Stack connectivity towards the LAN, while taking care of the tunnel encapsulation and the mapping scheme between IPv4 and IPv6 addresses. Although the standard allows for any means of configuring the 6rd tunnel on the CE, including manual configuration, it suggests the use of a single DHCPv4 option to auto-configure the device.
- On the provider side, the 6rd BR serves as the gateway to the global IPv6 network and terminates the tunnels to the individual 6rd CEs.

This test was performed in order to verify the algorithmic mapping between IPv6 and IPv4 addresses that are assigned for use within the Service Provider Network. This mapping allows for automatic determination of IPv4 tunnel endpoints from IPv6 prefixes, so IPv6 packets with destination addresses within the same 6rd domain will traverse 6rd CEs, while packets destined to the IPv6 Internet will be sent via the Border Router.

We connected two 6rd CEs to a 6rd BR. The 6rd CEs obtained a 6rd prefix and a WAN IPv4 address from the DHCP server residing on the 6rd BR.

A tester was used to emulate one of the 6rd CEs. On the emulated 6rd CE the 6rd delegated prefix was statically created by combining the 6rd prefix and all or part of its CE WAN's IPv4 address, as opposed to being dynamically created.

A commodity CPE was used as the other 6rd CE creating the 6rd delegated prefix automatically. All or parts of the CE IPv4 address was used depending on the value of the IPv4MaskLen.

The 6rd BR was configured with one IPv4 interface, a 6rd tunnel Interface for multi-point tunneling and a dual-stack IPv4/IPv6 interface connected to a traffic generator.

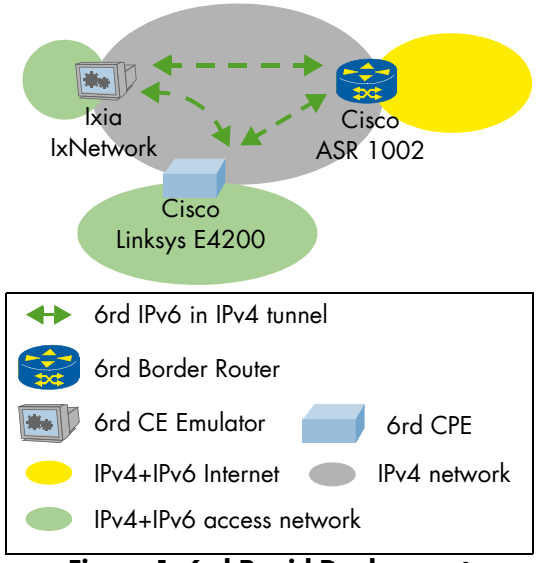


Figure 1: 6rd Rapid Deployment

The test was performed for two different IPv4MaskLen:

- 0: Using all 32 bits of the WAN IPv4 address,
- 8: Stripping the leading 8 bits of the WAN IPv4 address.

We also performed the test for two different IPv4 address ranges: a global IPv4 range and a private IPv4 range.

In all tests we first verified that the 6rd delegated prefix for use by the client behind the commodity 6rd CE was correctly created. We successfully tested that traffic within the same 6rd domain was directly forwarded between the 6rd CEs, by sending bidirectional traffic between both 6rd CEs. Packets were captured, and we verified that IPv6 packets were encapsulated in an IPv4 tunnel using the WAN IPv4 addresses of the corresponding 6rd CEs.

In addition we validated that traffic destined to the IPv6 Internet was encapsulated in an IPv4 tunnel using the 6rdBRIPv4Address and sent to the 6rd BR, which decapsulated and forwarded it to the destination IPv6 address. This was also verified by capturing the packets.

We verified that IPv4 traffic originating from a client behind 6rd CEs was routed natively over the IPv4 network.

We tested successfully: Ixia IxNetwork emulating one 6rd CE, Cisco Linksys E4200 acting as the

commodity 6rd CE and Cisco ASR 1002 as 6rd BR.

In a test run with other devices we encountered an issue with a 6rd CE. After receiving the 6rd prefix from the 6rd BR, it was unable to correctly perform the concatenation of the delegated prefix with the WAN IPv4 address obtained from the DHCP server. The resulting concatenated prefix was shifted by one bit.

Dual-Stack Lite Broadband Deployments

The Dual-Stack Lite (DS-Lite) technology has been standardized by the IETF in RFC 6333 and is aimed at providing an IPv4 service over an IPv6 network, using an Address Family Transition Router (AFTR) and a Basic Bridging BroadBand (B4) to provide the IPv4 service. The B4 is a function implemented within a DS-Lite capable CPE and provides the tunnel link to the AFTR.

DS-Lite combines two well-known technologies: IPv4-in-IPv6 (4in6) tunneling and NAT44, to enable the deployment of IPv6 in the core network independent of IPv6 deployment in the global Internet. Since the 4in6 tunnels can terminate anywhere in the provider network, it allows a horizontal scaling of the DS-Lite technology.

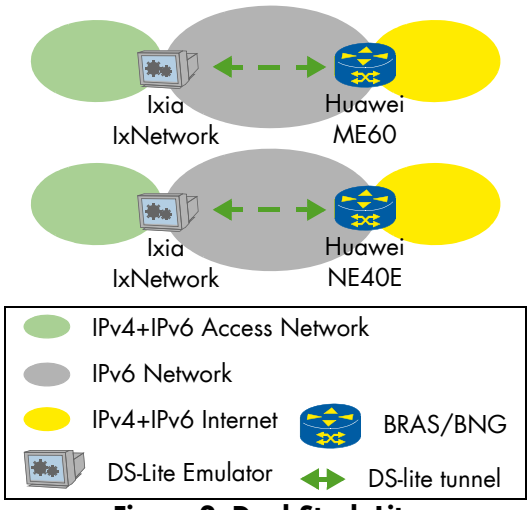


Figure 2: Dual-Stack Lite

To test this function, a device under test was configured to provide the DS-Lite AFTR function, as well as IPv6 native connectivity.

The emulation of a client in the access network and the DS-Lite CPE, in particular the tunnel encapsulation part of the B4 function, was performed using a traffic generator. Traffic was also captured on both ends of the DUT and analyzed.

We observed that the IPv6 traffic was routed natively, and the IPv4 traffic was encapsulated in a 4in6 tunnel and had NAT44 applied to it. A NAT policy was configured to allow up to 50 sessions, which was the minimum configurable value on the DUT. Using the traffic generator we first generated 50 sessions with no traffic loss, then 51 sessions with the expected traffic loss for one session.

Huawei NE40E and Huawei ME60 successfully tested for the DS-Lite AFTR function. Ixia IxNetwork successfully emulated the tunnel encapsulation functionality of the DS-Lite CPE.

Dual-Stack PPPoE

The Point-to-point Protocol (PPP) provides multi-protocol access and can be used to provide IPv4 access through the IP Control Protocol (IPCP), as well as IPv6 access through the IPv6 Control Protocol (IPV6CP) as defined in RFC 5072. We tested these protocols and options for Dual-Stack PPPoE:

- IPCP
- IPV6CP
- DHCPv6 Prefix Delegation

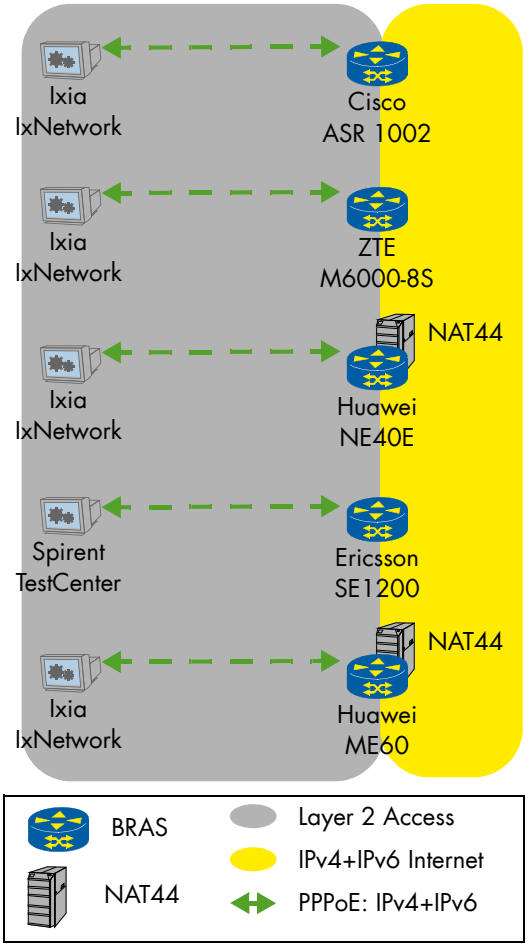


Figure 3: Dual Stack PPPoE

One noticeable difference between IPCP and IPV6CP is the address negotiation. While the addresses are directly negotiated in IPCP, IPV6CP is limited to negotiating the 64-bit interface identifier, which is used to construct the Link-Local addresses. After the link is ready, IPv6 addresses and IPv6 delegated prefixes can be negotiated using DHCPv6.

The device under test was a PPPoE server, DHCPv6 server and gateway to the IPv4/IPv6 public network. The client was emulated using a tester. Results for this test were verified using packet captures.

We observed the establishment of the PPPoE service, followed by the PPP session negotiation and establishment, in particular the IPCP for IPv4 and IPV6CP for IPv6.

After the establishment of the IPV6CP, we observed the DHCPv6 delegated prefix address negotiation.

The following devices were tested successfully: Cisco ASR 1002, Ericsson SE1200, Huawei ME60, Huawei NE40E and ZTE M6000-8S. Huawei ME60 and Huawei NE40E both successfully tested NAT44 in this setup. Ixia IxNetwork and Spirent TestCenter have successfully emulated the PPPoE IPv4+IPv6 CPE sessions

Dual-Stack PPPoE and Dual-Stack L2TP

The Layer Two Tunneling Protocol (L2TP) as defined in RFC 2661 facilitates the tunneling of PPP packets across a network. This could be used to establish IPv6 sessions across an IPv4 core network.

To aid in the deployment of IPv6 without changes to the existing IPv4 core network, it is possible to use IPv6 enabled L2TP Access Concentrators (LACs) and L2TP Network Servers (LNSs).

Within this test we verified the functionality of PPPoE and L2TP in the context of Dual-Stack:

- Dual Stack PPPoE and Dual Stack L2TP using IPv4 transport
- IPCP
- IPV6CP
- DHCPv6 Prefix Delegation

A device under test was configured to provide the function of the PPPoE server and LAC, and a second DUT was configured as the LNS. The L2TP tunnel from the LAC to the LNS was configured using IPv4 addresses.

The client was emulated using a tester apart from one test instance in which we used AVM Fritz!Box WLAN 3370 CPE and a client PC connected to the AVM CPE.

To verify the results we used a packet capture. We observed the establishment of the PPPoE service, followed by the establishment of an L2TP session within the L2TP tunnel and the PPP session negotiation and establishment of IPCP for IPv4 and IPV6CP for IPv6. After the establishment of the IPV6CP, we observed the DHCPv6 delegated prefix address negotiation.

The devices to successfully participate in the test were: Cisco ASR 1002 and Ericsson SE1200 as both LAC and LNS; Huawei NE40E and ZTE M8000-8S as LAC; Huawei ME60 as LNS.

In two of the test runs which included Ericsson SE1200 and Huawei ME60 providing the LNS functionality we also verified NAT44 operation.

Ixia IxNetwork have successfully emulated the PPPoE IPv4+IPv6 CPE sessions.

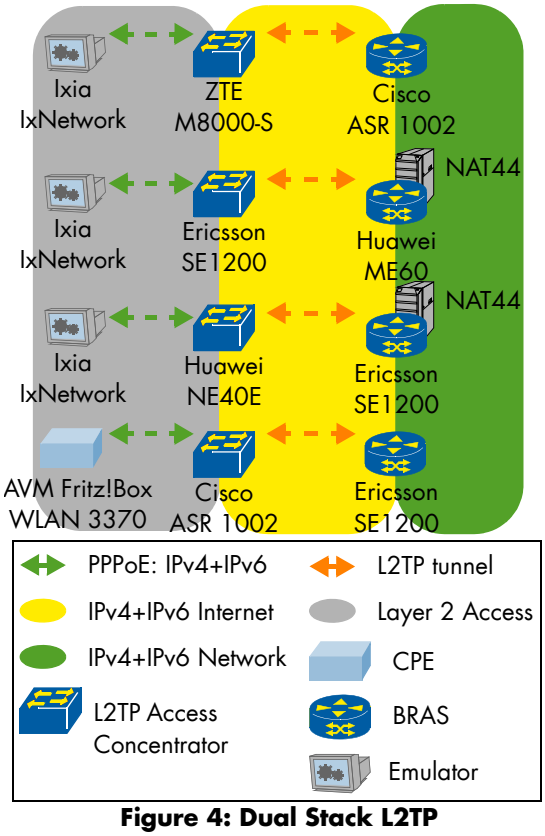


Figure 4: Dual Stack L2TP

6VPE L3VPN Traffic Isolation

The concept of 6VPE was introduced in order to address the need to transport IPv6 Layer 3 Virtual Private Networks (L3VPNs) over an IPv4 or IPv6 core network. It was published by the IETF in RFC 4659. IPv6 L3VPN is a direct descendant of IPv4 L3VPN. It aims to provide the same functionality for the IPv6 address family as is available for IPv4 L3VPNs. In this test we aimed to verify the correct isolation of IPv6 traffic in a multi-vendor IPv4 MPLS core.

We connected four PE devices in a physical mesh network. In addition, a tester emulating a PE router was connected as a leaf node. These 5 devices formed an MPLS IPv4 core network.

The tester also captured the iBGP packets, which we later analyzed, as well as emulating two CE devices per 4 of the 5 PEs.

We defined two Dual Stack VPNs, "green" and "yellow" with both IPv4 and IPv6 overlapping and non-overlapping addresses.

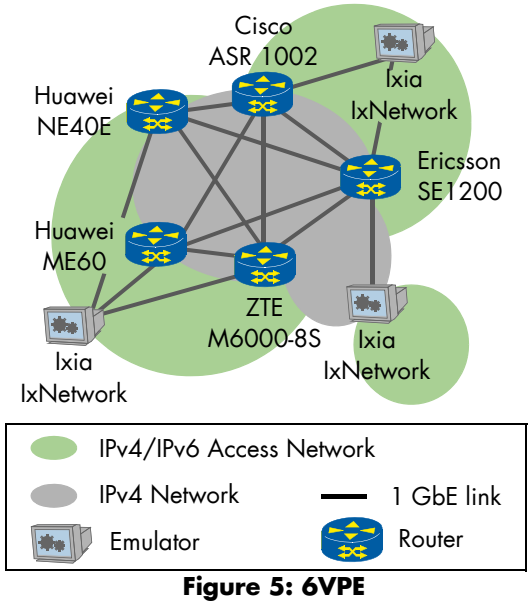


Figure 5: 6VPE

We then established an MP-BGP session between the emulated CEs and the PEs. After the devices learned the routes in the topology, we proceeded to generate traffic in a full mesh between all possible end-points, resulting in over 300 traffic pairs flowing simultaneously.

Traffic destined to addresses within the two different VPNs reached its destination with no packet loss. Traffic originating from "green" destined to "yellow", and vice versa, was dropped as expected.

We also observed that the IPv6 routes were advertised with the IPv4 next hop encoded as an IPv6 address, in accordance with RFC 4291.

There were 2 successful test runs: one with Cisco ASR 1002, Ericsson SE1200, Huawei ME60, Ixia IxNetwork and ZTE M6000-8S, and the other with Cisco ASR 1002, Ericsson SE1200, Huawei NE40E, Ixia IxNetwork and ZTE M6000-8S.

Ixia IxNetwork successfully emulated the CE devices.

DHCPv6

The Dynamic Host Configuration Protocol for IPv6 (DHCPv6) including the Relay Agent function has been formalized by the IETF in RFC 3315. The protocol is a direct descendant of DHCPv4 and provides auto-configuration in the context of IPv6 networks.

The Prefix Delegation option has been published by the IETF in RFC 3633. It provides a mechanism for automated delegation of IPv6 prefixes using DHCPv6. Both of these protocols were the subject of our next test area.

Stateful DHCPv6

Within this test we verified the following DHCPv6 functionality:

- IPv6 Address Auto-Configuration (IA Address)
- IPv6 DNS Auto-Configuration

The role of the device under test was to serve as the DHCPv6 server while the tester emulated the client. During the test execution we initiated the DHCPv6 client and server negotiation in order to obtain the IP addresses and configuration. We verified the successful negotiation by observing the DUT CLIs and DHCPv6 client GUI.

We also analyzed the captured packets and verified the exchange of DHCPv6 messages starting with the Solicit message and finishing with the Reply as is specified in RFC 3315 section 17 and 18.

We observed the assigned address under the Identity Association Address (IAADDR) field (DHCPv6 option 3), the IPv6 address for the DNS server (option 23) and the domain search list (option 24).

Cisco ASR 1002, Huawei NE40E, Huawei ME60 and ZTE M6000-5S were successful in accomplishing the test goals. Ixia IxNetwork and Spirent TestCenter successfully emulated the DHCPv6 CPE.

We encountered one issue during this test. Since we defined the address pool in the test to start from an arbitrary number in order to reserve the first 10 addresses - one vendor's device allocated addresses based solely on the network, starting with the first available address on the network.

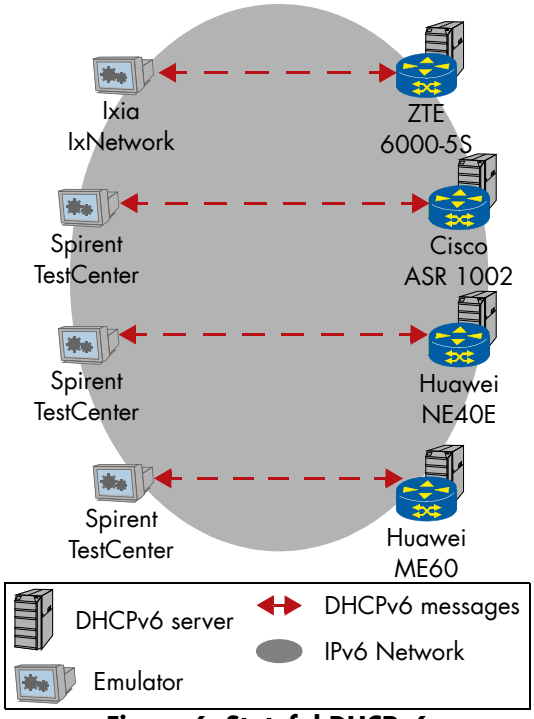


Figure 6: Stateful DHCPv6

Stateless DHCPv6

The Stateless DHCPv6 has been standardized by the IETF through RFC 3736 and provides a method to auto-configure devices without maintaining any state information.

To signal a stateless auto-configuration, an IPv6 router can send a Router Advertisement with the "Managed Address" flag bit cleared (0) and the "Other Configuration" flag bit set (1).

The receiving client will process this message accordingly and will start a Stateless Address Auto-configuration (SLAAC) process while obtaining other information such as DNS server address and Prefix Delegation using DHCPv6.

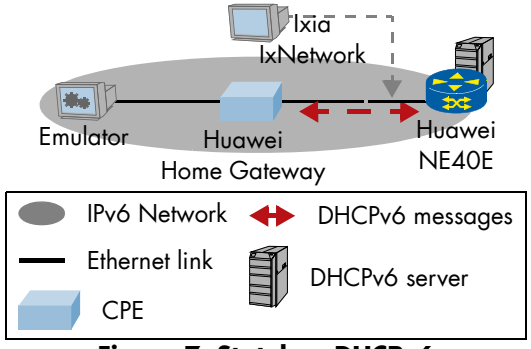


Figure 7: Stateless DHCPv6

Huawei demonstrated successfully stateless DHCPv6 using Huawei NE40E and Huawei's Home Gateway CPE. During the verification we used a client PC and analyzed the protocol exchange by captures.

Stateful DHCPv6 Relay Agent

In this test we verified the following DHCPv6 functionality in a Relay Agent setup:

- IPv6 Prefix Delegation (IA_PD)
- DHCPv6 Relay Agent functionality

The Relay Agent functionality differs from DHCPv4, as the entire DHCPv6 message to be relayed is included intact within the relay message itself.

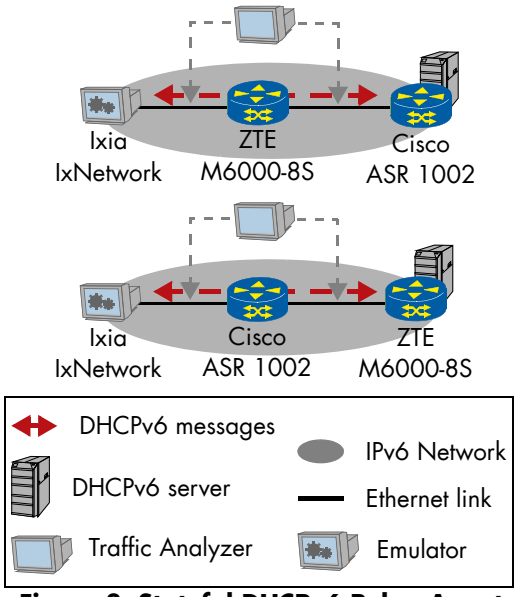


Figure 8: Stateful DHCPv6 Relay Agent

A device under test was configured as DHCPv6 server and a second device under test was connected to it providing Relay Agent functionality.

The client was emulated using a tester connected to the DHCPv6 Relay Agent, and the data plane packets were then captured and analyzed. Bidirectional traffic was generated to confirm proper routing on the Relay Agent.

During the test execution we initiated the DHCPv6 client and server negotiation in order to obtain the IP addresses and the configuration. We verified the successful negotiation by observing the DUT CLIs and DHCPv6 client GUI as well as by capturing the packets and analyzing them.

We observed the prefix delegation under the Identity Association Prefix Delegation field (IA_PD) containing the prefix with the prefix length (/48). A matching static route was installed on the router serving as the relay agent and was verified using the CLI as well as bidirectional traffic we sent to and from these routes.

Cisco ASR 1002 and ZTE M6000-8S were both successfully tested as DHCPv6 server and relay agent. Ixia IxNetwork successfully emulated a DHCPv6 CPE.

IPv6 Neighbor Discovery

The Neighbor Discovery Protocol (ND), standardized in RFC 4861, is the counterpart of the well-known Address Resolution Protocol (ARP) used by IPv4 speakers. In contrast to ARP, it is integrated into the IPv6 control layer provided by ICMPv6 and uses multicast groups determined by the low-order 24 bit of the address (as described in RFC 4291) instead of broadcasting the requests in the Ethernet layer.

The Stateless Address Auto-configuration (SLAAC-RFC 4862), describes the use of the Duplicate Address Detection (DAD) mechanism, which uses the ND Protocol.

In this test we verified the following IPv6 Neighbor Discovery mechanisms:

- Neighbor Discovery (ND)
- Duplicate address detection (DAD) for both Link local and global IPv6 addresses.

ND serves to discover IPv6 neighbors in order to learn their link-addresses and to determine if the address, which was manually or automatically configured, is already in use.

By default, IPv6 addresses are auto-configured using the link-layer address. For Ethernet interfaces, IPv6 addresses use MAC address in their construction as described in RFC 2464.

The address is considered tentative until the DAD timer expires and the number of retries, by default one, have been performed. If no response is received, the address is considered unique.

In this test case, two DUTs of different vendors were connected via an Ethernet switch that also provided a third port used for monitoring. We connected a traffic analyzer to the monitor port in order to evaluate the exchange of messages between the two devices under test. Initially the device under test ports were in “down” state.

As we enabled the ports, the traffic was captured and analyzed offline. We observed the correct auto-generation of addresses and the neighbor solicitation for the address owned by each device interface. We initiated the “ping” command and observed that neighbor solicitation messages to discover the link-layer address of the target were sent to the correct IPv6 multicast address and link-layer address, as specified in RFC 2464 section 7, and were answered with the correct reply.

When we verified DAD functionality for link-local addresses, we configured the same IPv6 link-local addresses on both DUT ports. We first observed that the status of the IPv6 address on the port was “tentative”. When we brought up the port of one of the devices we observed the status to be equivalent to “in use”. When we then brought up the port of the second DUT we observed the status equivalent to “duplicate” on one of the ports. DAD was also tested using global addresses in a similar setup.

The following devices successfully participated in the test: Cisco ASR 1002, Ericsson SE1200, Extreme E4G-200, Huawei ME60, Huawei NE40E and ZTE M6000-5S.

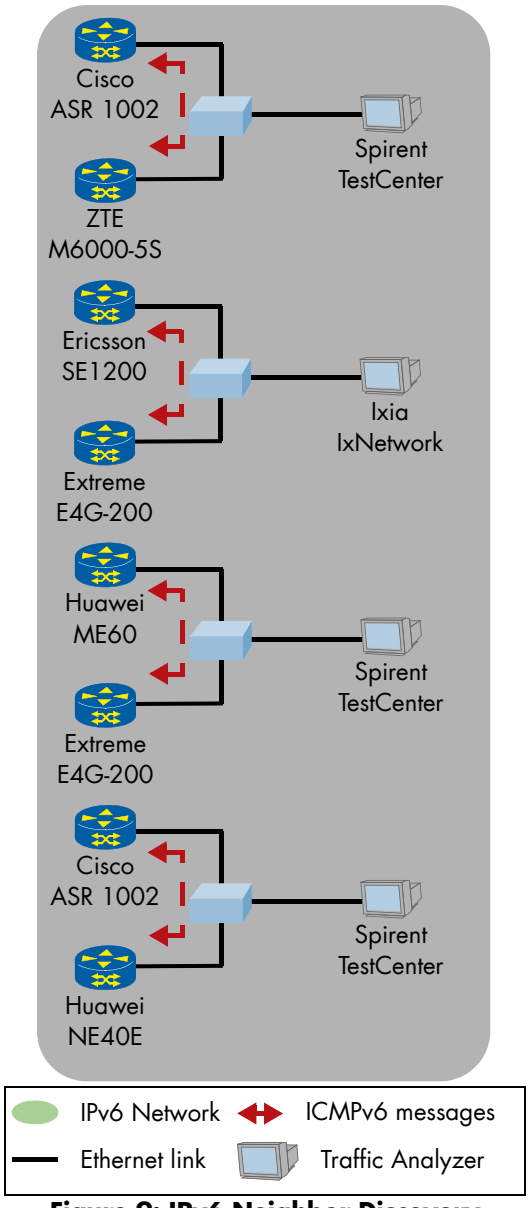


Figure 9: IPv6 Neighbor Discovery

ICMPv6: IPv6 Ping and Traceroute

The Internet Control Message Protocol (ICMPv6 - RFC 4443) is used by IPv6 nodes to report errors encountered in processing packets, and to perform other internet-layer functions, such as diagnostics (ICMPv6 "ping").

In this test we verified the following ICMPv6 functionality:

- ICMPv6 Echo Request (Type 128) and Echo Reply (129)
- Traceroute functionality utilizing increments of Hop Limit and the receipt of corresponding ICMPv6 "Hop limit exceeded in Transit" messages (Type 3, Code 0).

The ICMPv6 Echo Request and Echo Reply, as well as the traceroute mechanism are analogous to IPv4 ping functions. The Hop Limit (HL) field in IPv6 is similar to the IPv4 Time to live (TTL) field.

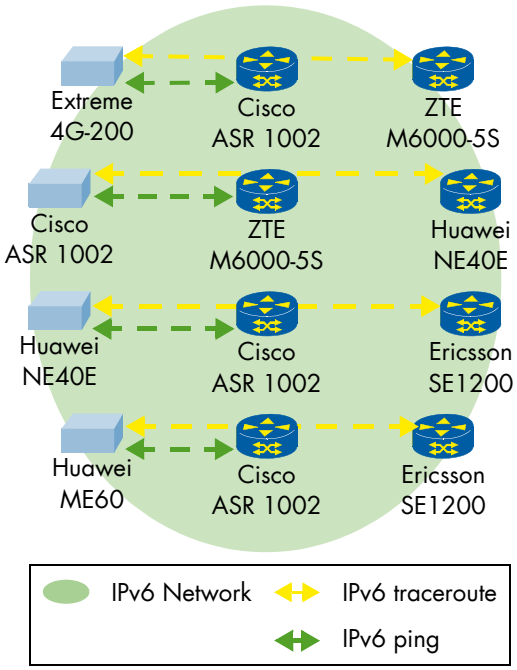


Figure 10: IPv6 Ping and Traceroute

In each test we connected three devices in order to test the traceroute mechanism in IPv6 while the ICMPv6 "Ping" was tested with the adjacent router.

During the test we observed the receipt of Echo Reply messages corresponding to the Echo Requests that were sent along with the round trip time and hop limit. We also observed the IPv6 addresses of the intermediate node as well as the traceroute endpoint node. We verified the results using the CLI output.

The following devices successfully participated in the test: Cisco ASR 1002, Ericsson SE1200, Extreme E4G-200, Huawei ME60, Huawei NE40E and ZTE M6000-5S.

In one test execution we observed that a mid point device did not send an ICMPv6 type 3, code 0 ("HL exceeded in transit").

Port Control Protocol in Context of Carrier-Grade NAT for Dual-Stack Lite

Port Control Protocol (PCP) is a simple protocol that is being defined by at the IETF. It provides a mechanism for applications to create a pinhole on a carrier grade NAT (CGN) device from an external IP address and port to an internal IP address and port. These pinholes are required for successful inbound communication destined to a host behind the CGN device.

PCP is a client-server protocol, where the PCP client issues PCP request to a PCP server for dynamic port allocation. Upon receiving a request, the server will allocate the requested ports for a specific lifetime. The pinhole created handles incoming packets destined to the port associated with that pinhole.

Our goal was to verify that PCP dynamically creates pinholes for a specific lifetime on a CGN device and ensure that a PCP-constructed pinhole is used when handling incoming packets destined to the associated port. We also wanted to verify that the pinhole was cleared when the lifetime expired.

In all test scenarios we first configured a DS-lite tunnel connecting the CPE including the Basic Bridging Broadband (B4) element and the Address Family Translator Router (AFTR). This was done by manually configuring the AFTR IPv6 address on the CPE. A tester emulated the CPE, including B4 element. We started the PCP server to control the DS-Lite Carrier Grade NAT, running on the AFTR. The CGN device was configured with a global NAT table. A proprietary software-based PCP client was connected to an AFTR and configured with the port information to be requested.

Initially all UDP ports were closed on the CGN, and we verified that UDP traffic generated from the Internet side of the DUT was dropped and no pinhole for the UDP traffic existed on the CGN.

The PCP client explicitly requested a port number to allow forwarding of UDP traffic through the CGN from the internet to the internal network. After the CGN received the PCP request from the client, it created a server-map with an aging timer of 600 seconds and the IP and port mapping. This was verified via Command Line Interface (CLI).

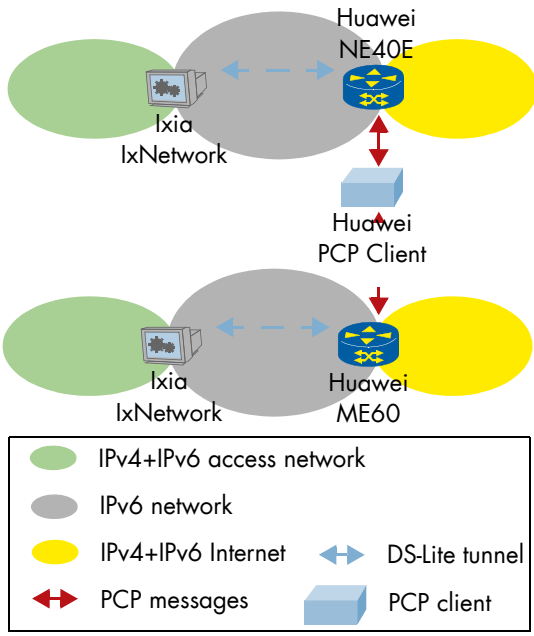


Figure 11: Port Control Protocol

We then started the UDP traffic over the DS-Lite tunnel and observed that a NAT session table was created with the same timer as the server-map table. The UDP traffic was correctly forwarded through the CGN.

We stopped the UDP traffic. After the pinhole lifetime was expired, we restarted the UDP traffic and observed 100% traffic loss.

The following devices demonstrated this functionality: Huawei NE40E and Huawei ME60 acting as AFTR, CGN and PCP server and Ixia IxNetwork emulating a CPE including the B4 element.

We noticed that the CGN managed two timers, one for the server-map and another for the NAT table map. Both timers could be set independently. The NAT table map expires before the server-map timer start to decrement. The pinhole is cleared when the server-map timer expires as expected.

We observed that the pinhole lifetime was not accurate. Immediately after the expiration of the pinhole lifetime, incoming packets destined to the port associated with that pinhole was still forwarded.

Mobile Backhaul Transport

Not all tests were focused on IPv6 migration scenarios - we still have a lot of requests for mobile backhaul transport tests which this area of this report is focusing on. Our Mobile Backhaul Transport tests were concentrated on aggregation network technologies.

In particular we tested new and updated implementations with new features of MPLS-TP and ERPS. We also tested interworking scenarios between MPLS-TP and ERPS as well as IP/MPLS and ERPS.

ITU-T Y.1731 performance monitoring is a long runner in our events. The protocol is very important for monitoring Layer 2 services in packet based aggregation networks. It has a rich set of features and is still being extended by ITU-T. In this event we tested for the first time interoperability for two way delay and delay variation measurement both with per CoS ID.

Another first timer in our event was Path Computation Element (PCE), which is being specified by PCE IETF working group. PCE can be used in IP/MPLS aggregation networks to centralize path computation for MPLS LSP Traffic Engineering LSPs. We see PCE as a piece of a network management system so its standardization could improve interoperability between management systems of various vendors.

Multi-Service Ethernet Ring Protection Switching

Ethernet Ring Protection Switching (ERPS) introduced by ITU-T G.8032, was developed to achieve protection switching within 50 milliseconds in an Ethernet Ring.

In an Ethernet ring network, proper Ethernet traffic forwarding requires a loop free topology. This is achieved by placing a logical block on a specific link called ring protection link (RPL) under normal operation. Each end of the RPL is connected to two ring nodes, called RPL owner and RPL neighbor, which block and unblock their RPL port depending on the ring state. Ring automatic protection switching (R-APS) is a protocol defined by ITU-T to coordinate protection activities in the ring. R-APS

protocol messages are transmitted over a control channel called R-APS channel, which is separated from the service channel by the use of different VLAN IDs.

When a link or node failure occurs in the ring, other than RPL, ring nodes detecting the failure report this failure to the other nodes in the ring using Ring Automatic Protection Switching Signal Fail (R-APS(SF)) messages which triggers protection switching on the ring by releasing the block of RPL and blocking the ports associated with the failed link. The RPL is then used to forward the traffic. R-APS is a protocol defined by ITU-T G.8032 to coordinate ring state and events across the Ethernet ring.

Upon recovery of the failed link, the node adjacent to the recovered link transmits R-APS No Request (NR) messages, which cause the RPL owner to start the Wait-to-Restore timer (WTR). Once the WTR timer expires, the RPL owner blocks RPL and transmits R-APS No Request, RPL Block (NR, RB) messages, indicating that the ring is in the idle state. Ring nodes receiving these messages, flush their Filtering Database (FDB), and the nodes adjacent to the previously failed link unblock their ports. The ring returns to the normal operation.

The current ITU-T recommendation G.8032 describes Flush Optimization, which significantly reduces the amount of flushing of the FDB in the ring. In some failure scenarios, like the failure of the RPL, the protection switching will not change the active topology of the ring, thus the flush operation is unnecessary. When a node detects an RPL failure, it sends R-APS messages with the do-not-flush (DNF), indicating state transition without FDB flush.

The ITU-T G.8032 standard also defines the protection switching process, where two or more Ethernet rings are interconnected to extend their coverage. An Ethernet Ring connected to other rings is called a sub-ring, and the ring to which a sub-ring is connected is called the major ring. Sub-rings and major rings are connected through the use of interconnection nodes. A physical ring in an interconnected ring can maintain several ERP instances separated by VLANs, so each ring instance uses different RPL ports.

Our test focused on verifying protection switching in a multiple instance ERPS as well as on verifying support of Flush optimization.

In this test scenario two physical Ethernet rings, one major and one sub-ring were interconnected. On both physical rings, two groups of services - Red and Black - were associated with four configured ERP instances, protecting both services. Each ring instance was provisioned with its own RPL owner and RPL neighbor to block the RPL link. The R-APS communication between all ERP instances was disambiguated by using different R-APS VLAN IDs. The sub-ring was configured as an open ring, meaning that R-APS communication in the sub-ring was terminated at interconnection nodes.

A failure of the link between the interconnection nodes was introduced. The expected failover and restoration time for each ERP instance in the major ring was less than 50 milliseconds.

We observed protection switching and restoration times ranging from 6 to 9 milliseconds, but the failover time ranged from 40 to 54 milliseconds.

The following devices participated in the test:

- In the major ring: Telco Systems T5C-XG as RPL owner for both instances, Telco Systems TMARC-3208SH and Cisco ASR 9006 acting as interconnection nodes.
- In the sub-ring: Extreme Networks E4G-400.

To test flush optimization, we configured an additional port on Cisco ASR 9006 to provide a monitoring function allowing us to capture packets and evaluate the R-APS messages received. We then disconnected the RPL for the Red service, captured packets and analyzed them. We observed the proper receipt of the R-APS (SF, DNF).

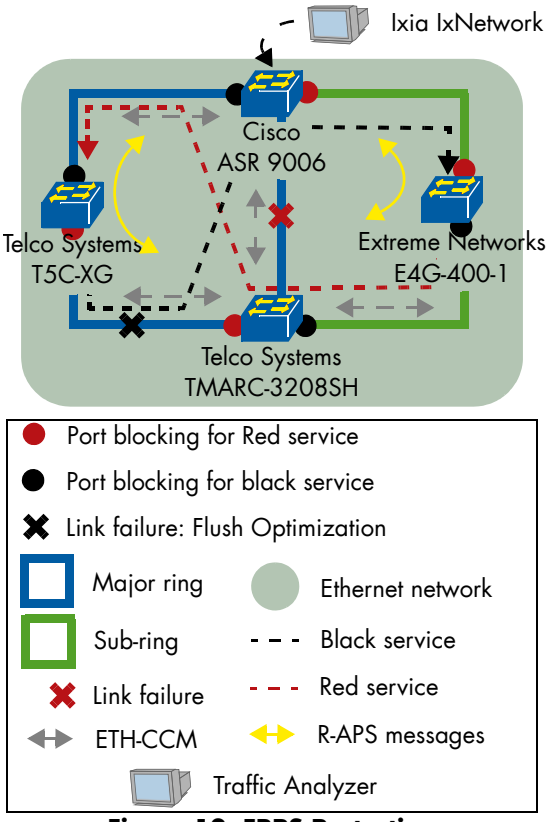
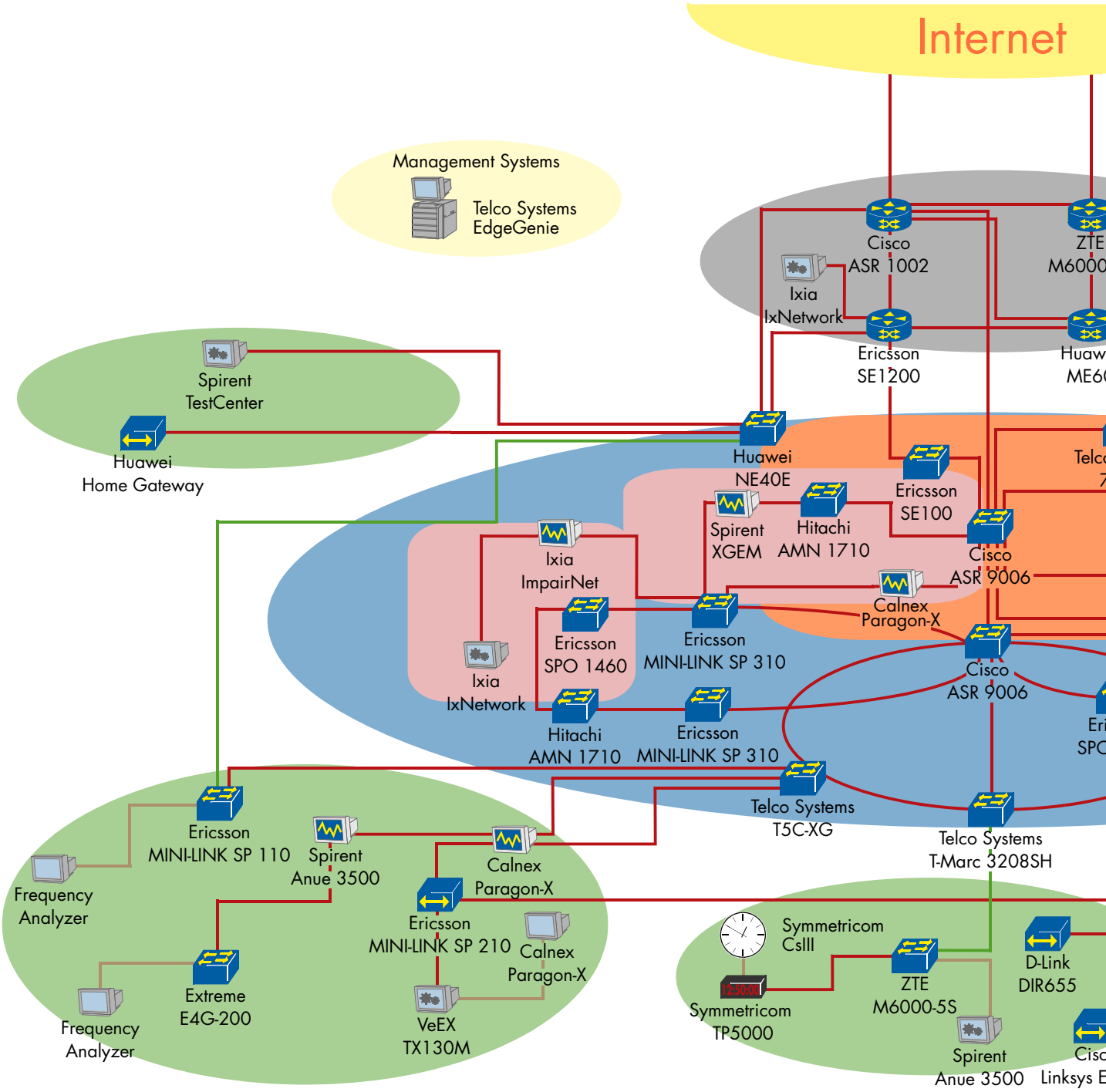


Figure 12: ERPS Protection

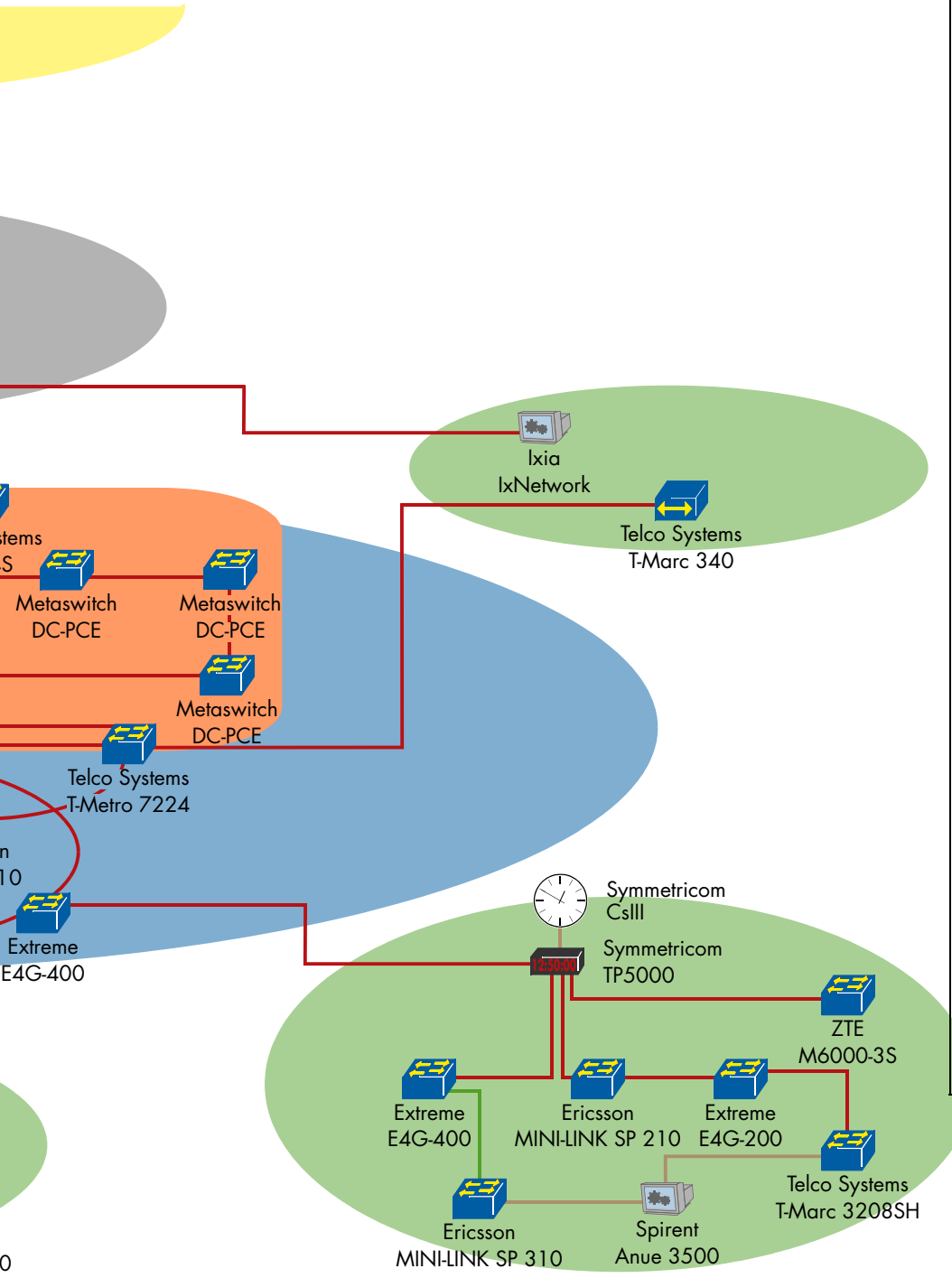
Since the link failure emulation was performed via link disconnection, there was a need for CFM configuration. We observed some differences between implementations of the ETHDi/ETH_A function which extracts and generates R-APS messages. Some vendors required MEP to be configured and running in order to configure their ERP and some did not. For this reason CFM was configured between some nodes at 100 milliseconds interval.

During the test we observed that some vendors were unable to correctly process the CFM messages format sent by others vendors.

Physical Netw
Multi-Vendor MPLS & Etherne



rk Topology
Interoperability Event 2012



Device Types

- Provider/ Provider Edge device
- Aggregation device
- IPv4/IPv6/Ethernet CE/CPE
- Analyzer/Traffic Generator
- Emulator
- IEEE 1588v2 Grandmaster
- Impairment tool
- Reference Clock

Connection Types

- Ethernet link
- Aggregated Ethernet link
- Clock link

Network Areas

- IP/MPLS core network
- IP/MPLS aggregation network
- MPLS-TP aggregation network
- Ethernet aggregation network
- IPv4+IPv6+Ethernet Access network
- ERPS Rings

ERPS and VPLS Interworking

At this event we wanted to focus on how ERPS can be interconnected with an IP/MPLS or MPLS-TP aggregation networks. This could be a rather standard configuration where a redundant access network running ERPS could be dual-homed to an IP/MPLS or MPLS-TP aggregation network.

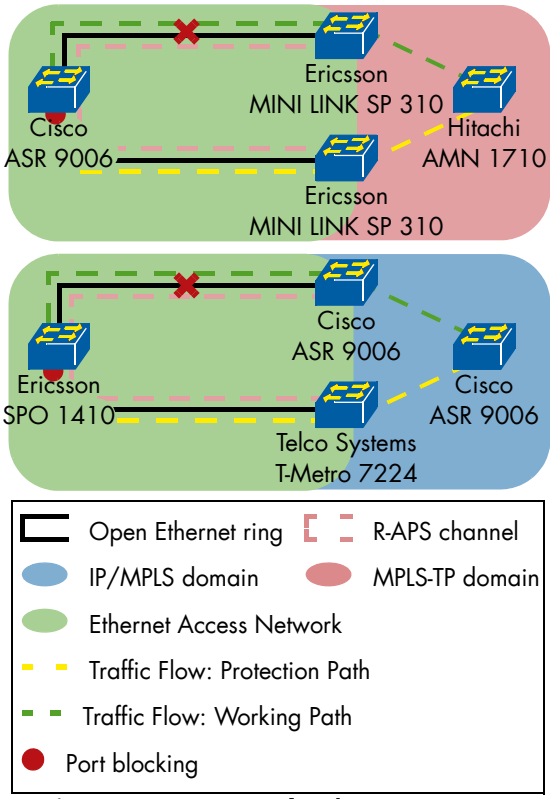


Figure 13: ERPS and IP/MPLS VPLS or ERPS and MPLS-TP VPLS Interworking

In this test, one open ring was constructed and one node was provisioned as RPL owner. The open ring was interconnected to either an IP/MPLS or MPLS-TP domain via two interconnection nodes, one on the working path and the other on the protection path. There was no Ring Automatic protection switching Virtual Circuit between the interconnection nodes. The service was built using 802.1Q standard in the access network and was dual-homed to an MPLS-TP or IP/MPLS aggregation network, using a Virtual Private LAN Service (VPLS) with static or dynamic Pseudo Wires (PW). Two VLANs were implemented in the open ring, one for R-APS messages and the other for ERPS traffic. Continuity Check Message (CCM) were not configured in the ring for any of the scenarios. Loss of Signal (LoS) was used to trigger a fault in the ring.

We emulated a fault condition in the ring and validated the Ethernet Ring Protection Switching (ERPS) protocol exchanges as well as measured the failure duration in both directions. We then removed the fault conditions and measured the restoration time. We expected to record an out of service time value of less than 50 ms during failover and reversion for both directions.

In one scenario we recorded less than 50 ms failure duration for ring to VPLS direction, and less than 500 ms out of service time for VPLS to ring direction, since not all devices supported MAC Address Withdrawal implementation.

We recorded less than 52 ms failure duration for both directions in the other scenario.

For both test runs we observed that the restoration duration was less than 50 ms for both directions.

Also, we validated the G.8032 protocol exchanges for the two scenarios. We verified that the ring was in "Idle" state at the beginning. As soon as we disconnected the cable, protection switching was triggered and the RPL node in the open ring unblocked the RPL port. The ring was in "Protection" state at this step. We verified that the port adjacent to the fault blocked the faulty link and sent Signal Failure (SF) message to the neighbor node. We observed that for the IP/MPLS scenario, the interconnection node on the forwarding path sent the LDP MAC Address Withdraw message to the VPLS peer and performed MAC flushing. As soon as the VPLS edge router has received a MAC Address Withdrawal message, it performed MAC flushing and flooded traffic to both the forwarding and the blocking paths. For the other scenario, the interconnection node on the working path performed MAC flushing as soon it was in the Signal Failure State. The VPLS peer re-learnt the new location as soon as service frame traffic began to be received over the protection path.

For both scenarios, we verified that the adjacent node detected that the failure was resolved and sent R-APS No Request (R-APS(NR)) message triggering the RPL owner to start the Wait to Restore (WTR) timer. The RPL owner blocked its RPL port upon expiration of the WTR timer. Nodes adjacent to the formerly failed link unblocked their ports, and the ring then resumed its original operation. At this point all nodes flushed their MAC tables, and the traffic switched back to the working path. As in the failure procedure, the interconnection node on the protection path can support MAC Address Withdrawal messages to notify the VPLS peer that the failure was resolved. The switchover based on the MAC Address Withdrawal functionality was observed in one scenario.

The following devices were tested successfully: Cisco ASR 9006, Ericsson MINI-LINK SP 310, Ericsson SPO 1410, Hitachi AMN 1710 and Telco Systems T-Metro 7224.

MPLS-TP Fault Management

MPLS-TP fault management automatically indicates a disruption on a link or node along the path to an MPLS-TP LSP endpoint. This indication is important to suppress alarms and to activate protection as defined in RFC 6427.

The MPLS-TP Alarm Indication Signal (AIS) message and corresponding Link Down Indication (LDI) flag are generated in response to detecting defects on a path. The AIS/LDI is sent when a link along the path is disrupted. Lock Report (LKR) is generated when a

link along the path is administratively shutdown. All these are actions of a mid point device along the path. It sends a signal telling the endpoint which defect is detected, so that the endpoint is immediately notified about the failure.

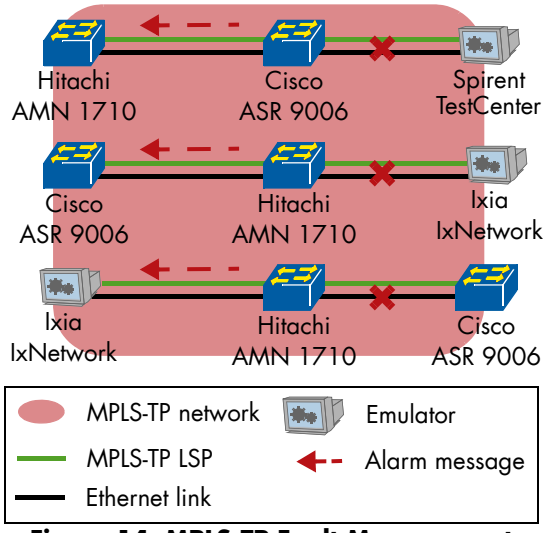


Figure 14: MPLS-TP Fault Management

The following devices participated in this test case: Cisco ASR 9006, Hitachi AMN 1710 and Ixia IxNetwork.

For all scenarios, the service was configured using Pseudo Wires (PWs). We tested the AIS/LDI messages by disconnecting one of the links along the path. The intermediate node propagated the AIS/LDI On message as soon as the link was down, and the AIS/LDI Clear message as soon as the link was up. We tested the AIS/LDI message for all scenarios successfully. In order to test the LKR message, an administrative command was required. This has been tested only in one scenario where the intermediate node used "Shutdown Port" command as an administrative command. As soon as we performed the port shutdown command, the intermediate node generated an LKR message that was then observed on the edge node. Then as soon an the lock condition was cleared, we recorded an "AIS/LDI On" message, followed by an "AIS/LDI Clear" message, and finally an "LKR Clear" message. The vendor explained that the reason for sending the AIS/LDI is due to the fact that after a port is enabled, the interface goes from the "admin down" state, through the "down" state, and then to the "up" state.

MPLS-TP 1:1 OAM-Based Protection

Resiliency in MPLS-TP network is an important topic for network operators. Due to the progress in standardization, we decided to add an MPLS-TP 1:1 OAM-based Protection test in our interoperability event. We focused on "draft-ietf-mpls-tp-linear-protection-draft 07" in our previous events, but by focusing on RFC 6378, vendors with new implementations could participate in this test case.

The test scenario was straightforward from a testing perspective: we built two MPLS-TP LSPs between two DUTs using 1:1 protection according to RFC 6378 on two different links between the DUTs. We first failed the active LSP and measured the out of service time during the failover. We then restored the LSP and measured the out of service time while the devices reverted to the initial LSP. We expected the out of service time to be less than 50 ms during failover and restoration.

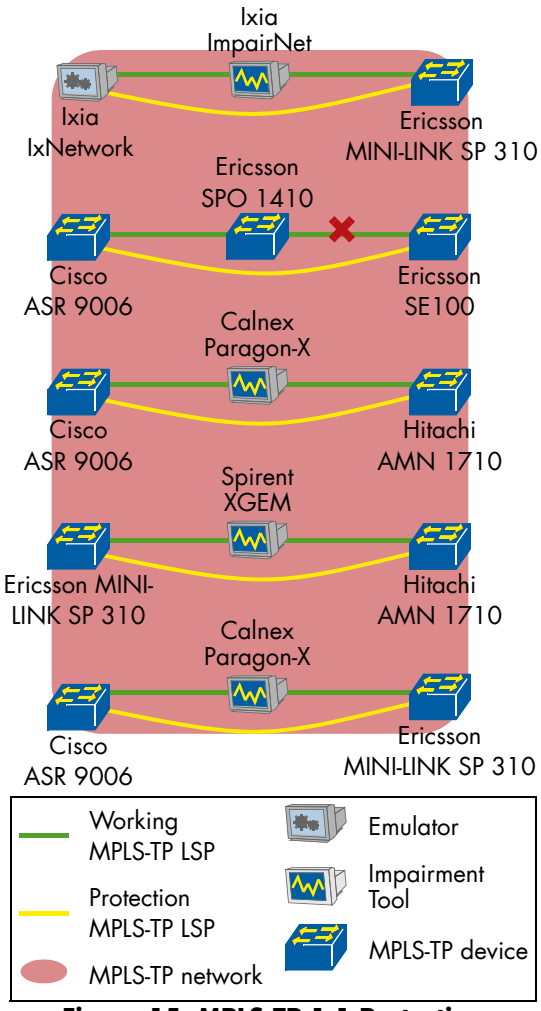


Figure 15: MPLS-TP 1:1 Protection

Besides protection switching we also tested MPLS-TP administrative commands based on RFC 6378. These commands were "Lockout of protection", "Forced Switch" and "Manual Switch".

For all test scenarios, the test was performed by statically configuring primary and protection Label Switched Paths (LSPs) in each direction. In addition we configured a Pseudo Wire (PW) which was transported over the LSPs. We generated bidirectional test traffic in order to measure out of service times.

BFD Continuity Check (CC) sessions were running on both primary and secondary LSPs to monitor the liveliness of the LSPs. BFD-CC was transmitted over the Generic Associated Channel (G-ACh) using Generic Associated Label (GAL). During BFD session establishment sessions follow the BFD slow start procedures. Once the BFD sessions transitioned to the "UP" state the transmit and receive intervals were negotiated to either 3.33ms or to 100ms.

For all scenarios we emulated a unidirectional link failure on the primary path by dropping all traffic if we used an impairment tool. When we used an

intermediate node, we disconnected a link between the intermediate node and one of the DUTs.

In all test scenarios, we used one of the three impairment generators — Calnex Paragon-X, Ixia ImpairNet and Spirent XGEM to provide impairment functions. When the impairment generator was not available (due to other testing requirements for example) an intermediate node was used.

The following devices successfully demonstrated interoperability in this test case: Cisco ASR 9006, Ericsson MINI-LINK SP 310, Ericsson SE100, Hitachi AMN 1710 and Ixia IxNetwork.

For some combinations, we observed less than 50ms out of service times when the BFD interval was negotiated to 3.33 ms. For other combinations, we observed an out of service times less than 570 ms when the BFD interval was negotiated to 100 ms.

We also observed that some BFD implementations supported a discrete set of transmission intervals (e.g. 3.3, 10, 100, 1,000 ms) while others supported a continuous range (e.g. any value from 15 to 1,000 ms). As a result, the lowest common transmission interval between an implementation with a discrete set of transmission intervals and an implementation with a continuous range was negotiated by BFD. In the example above this interval was negotiated to be 100 ms. Therefore even if all implementations were able to achieve an out of service times of less than 50 ms, not all implementation combinations were able to demonstrate this.

The administrative commands were successfully tested based on "draft-ietf-mpls-tp-linear-protection-draft 03" for one pair combination and based on "draft-ietf-mpls-tp-linear-protection-draft 06" for another pair combination. No result was recorded for administrative commands based on RFC 6378. The processing priority of Forced Switch command against Signal Failure (SF) on protection path is a difference between the draft versions and their later RFC6378.

During the test we observed that one implementation did not revert to the working LSP after the fault condition was resolved and the WTR timer has expired. Another implementation required receiving Protection State Coordination (PSC) protocol message from the paired node to revert. As PSC was not supported by one vendor the both-end revertive mode was not observed for that scenario.

Performance Monitoring

When operating Ethernet Virtual Connections (EVC), network operators require tools that can monitor performance in order to verify Service Level Agreements (SLAs). ITU-T Y.1731 provides such tools.

In this event we focused on Y.1731 performance monitoring functions that we did not test in previous events: performance monitoring per Class of Services (CoS). We also tested single ended Y.1731 Loss Measurement.

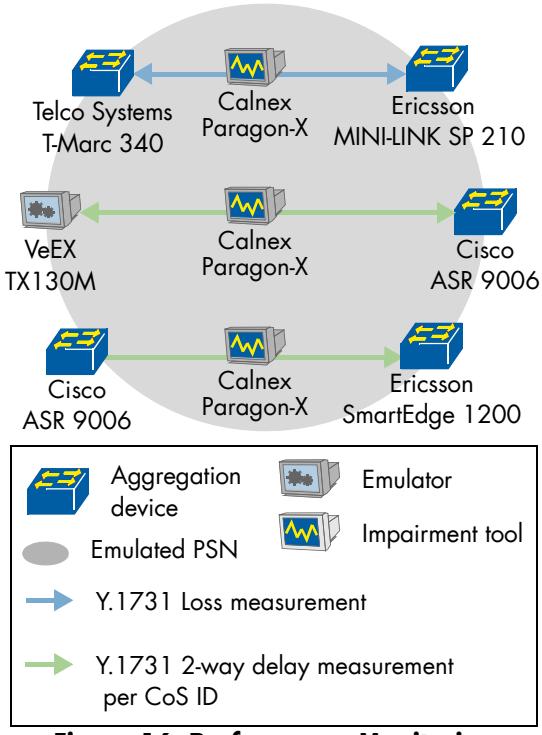


Figure 16: Performance Monitoring

We distinguished between three measurement types for the test: frame loss per EVC, two-way frame delay and two-way frame delay variation per VLAN CoS ID.

As a baseline reference test, we first validated performance monitoring implementations per CoS ID without introducing any impairment. Then we introduced impairment per CoS ID with artificial constant delay, delay variations, or packet loss. For each specific type of impairment we verified the measurement results provided by the implementations under test against the emulator configuration. We expected the delta between the impaired configuration and the reference test to be equivalent to the impairment tool settings.

For delay measurement we added a unidirectional constant delay of zero, 10 and 20 milliseconds (ms) per CoS ID 5, CoS ID 3 and CoS ID 0 respectively. We expected the devices to observe that the average delay stayed unchanged for CoS ID 5, increased by 10 ms for CoS ID 3 and increased by 20 ms for CoS ID 0.

In order to test per CoS ID delay variation, we configured the impairment tool to introduce no packet delay for CoS ID 5. For CoS ID 3 and 0 we introduced an unidirectional packet delay of 5 ms (CoS ID 3) or 15 ms (CoS ID 0) to every second Delay Measurement Message (DMM) and 15 ms (CoS ID 3) or 25 ms (CoS ID 0) to other DMMs.

We expected the average delay for CoS ID 3 and 0 at 10 ms and 20 ms respectively, and delay variation at 10 ms.

As we performed loss measurement we sent bidirectional Ethernet traffic over the network service and introduced 10% frame loss in one direction with the impairment tool. We verified whether the far-end

and the near-end frame loss displayed on the device under test showed the same loss values.

The following combinations were successfully tested for two-way delay measurement per CoS ID: Cisco ASR 9006, Ericsson SmartEdge 1200; Cisco ASR 9006, VeEX TX130M;

In all test scenarios, we used the Calnex Paragon-X to inject the required frame loss and delay onto the target OAM packet streams.

Ericsson MINI-LINK SP 210 and Telco Systems T-Marc 340 were successfully tested for single-ended frame loss measurement.

During the tests we observed that one implementation supported concurrent delay measurement probes per CoS ID, while the other implementation performed the measurement on demand at a time per single CoS ID.

Interconnecting MPLS-TP and IP/MPLS

In this test we verified interconnection of MPLS-TP and IP/MPLS networks as well as fault propagation over IP/MPLS and MPLS-TP domains in a multi-vendor environment.

The following devices participated in this test case: Cisco ASR 9006, Ericsson SE100, Ixia IxNetwork and Telco Systems T-Metro 7124S. We tested two device configurations. In both, Cisco ASR 9006 was used as a stitching point and Ixia IxNetwork configured a pseudowire in the MPLS-TP part of the network.

Either Ericsson SE100 or Telco Systems T-Metro 7124S was used to configure Access Circuit (AC) in IP/MPLS domain. In order to verify that the stitching point worked properly, we generated traffic in both directions. As expected we did not observe any packet loss.

The following device combinations successfully performed the IP/MPLS and MPLS-TP interconnection part of the test: Ixia IxNetwork, Cisco ASR 9006, Ericsson SE100, and Ixia IxNetwork-, Cisco ASR 9006, and Telco Systems T-Metro 7124S

Due to mismatch of pseudowire status notification code points in different implementations, we could not successfully perform the second part of the test: fault propagation between MPLS-TP and IP/MPLS.

Path Computation Element Interoperability

Path computation in large, multi-domain networks can be very complex. Traditionally, path computation takes place on the network nodes. This may produce only an approximation to the best path, for two reasons. First, the network nodes do not usually have a full view of the entire network topology, but only see a summarized version of the topology outside their own area. Second, in some types of network, computation of an exact path may require more computational power than is usually available in a network node.

A Path Computation Element (PCE) is defined by the Internet Engineering Task Force (IETF) as an entity like a component, application or network node that is capable of computing a network path or route based on a network graph and applying computational constraints.

We used a procedure that allowed us to validate that the protocol exchange between two implementations worked as expected, focusing on RFC 5440. At the beginning of the test a Path Computation Client (PCC) requested a Path Computation Element (PCE) Communication Protocol (PCEP) session from a PCE server. As soon as the PCEP session was up, the PCC attempted to set up a TE LSP with 50% of the bandwidth available for reservation. All requested LSPs were originated from area 0 and destined for area 1. As soon as the first TE LSP was established, the PCC attempted to set up the second TE LSP with additional 40% of the reservable bandwidth. As the second TE LSP was established, the PCC requested to set up a third TE LSP with additional 20% of the reservable bandwidth. The PCE server replied then indicating that no path satisfying the required constraints was found.

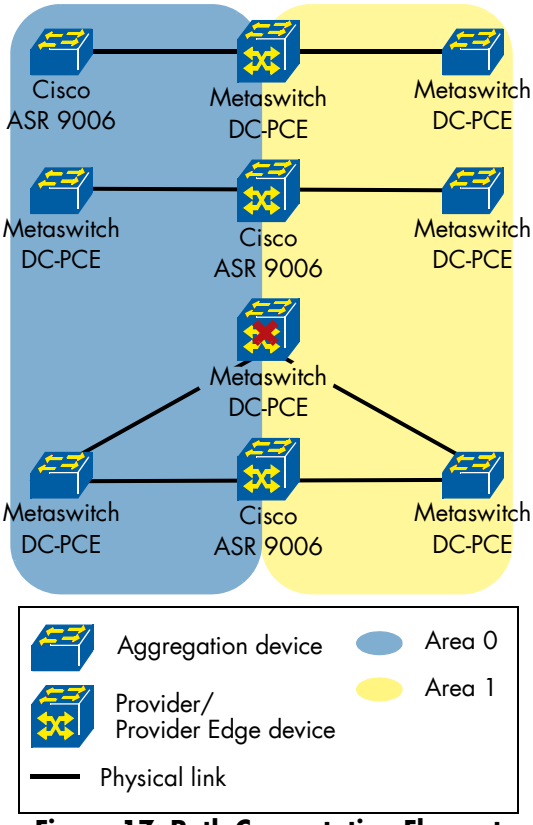


Figure 17: Path Computation Element Interoperability

The following combinations were successfully tested for the single PCE server interoperability: Cisco ASR 9006 as PCE server and Metaswitch DC-PCE as PCC; Cisco ASR 9006 as PCC and Metaswitch DC-PCE as PCE server. For both scenarios, another Metaswitch DC-PCE was used as the LSP end point.

Each PCE server reserved 200 Mbit/s of a 1 Gbit/s interface in area 0 and 100 Mbit/s of a 1 Gbit/s interface in area 1 for all LSPs requested by the PCC

from area 0 to area 1. We verified that for both scenarios, the first and second TE LSPs were established with 50 Mbit/s and 40 Mbit/s respectively. The PCE server sent a negative reply for the third TE LSP request for 20 Mbit/s due to the bandwidth constraints in area 1.

As a part of this test we also verified PCE server failure. For this test two PCE servers were set up. One of the PCE servers was prioritized over the other. The identity and priority of each server was dynamically advertised to the PCCs via OSPF, using the PCE Discovery TLV as specified in RFC 5088.

The redundancy of PCE server was successfully tested in the following configuration:

One Metaswitch DC-PCE as a PCC, connected to Metaswitch DC-PCE primary server, and to Cisco ASR 9006 as secondary server.

In the test we observed that each PCE server set maximum bandwidth available for reservation on each interface to 200 Mbit/s in area 0 and to 100 Mbit/s in area 1. The links on the route through the secondary PCE server were configured to a lower TE link metrics than the links on the route through the primary. We then verified that the link metrics were distributed via OSPF and that the primary PCE server indicated establishment of the first and the second TE LSPs using the links with the lower TE metrics.

As LSPs were established, information about the remaining available bandwidth was distributed between PCE servers via OSPF. We verified that a secondary PCEP session was established from the PCC to the secondary PCE server and that no LSPs were lost on the node running secondary PCE server upon the failure of primary PCE server. As the primary PCE server came back up, the PCC attempted to establish a TE LSP with 20 Mbit/s. We observed that the primary PCE server replied indicating a path with higher TE metric as no path satisfying the required constraints was available along the links with lower TE metric. The LSP was then established over the path with higher TE metric.

Mobile Backhaul Synchronization

In the area of Mobile Backhaul Synchronization we continued our successful test program for testing ordinary, boundary, transparent, and grandmaster clocks. New vendors and new implementation showed interoperability with the implementations that were already tested in previous events.

We extended our test program this time with two new scenario tests, which are Synchronous Ethernet over Ethernet Link Aggregation Group (LAG) and Synchronous Ethernet Islands connected through PTP (IEEE 1588-2008).

SyncE over LAG is of interest to service providers that offer Carrier Ethernet services to mobile operators over protected Ethernet User Network Interface (UNI). In addition to the protected data service, mobile operators might wish to receive also synchronization over the same UNI by using SyncE. In failure scenarios it is desirable that synchronization is not disturbed when the LAG of the protected UNI switch over from one link to another.

Transport of SyncE clock through PTP is another use case in this test area. In order to provide SyncE clock across a network, all devices has to support SyncE. This is, however, often not the case. In order to bridge the SyncE clock through the non-SyncE nodes, IEEE 1588-2008 could be used.

Precision Time Protocol: Frequency and Phase Synchronization

Two major roles that PTP devices play are grandmaster and ordinary clocks. Connecting ordinary clocks to a grandmaster clock over a packet based network delivers frequency and phase synchronization at certain level on the location of the ordinary clocks.

The synchronization quality depends on the packet delay variation as well as asymmetry of delays in down- and upstream directions between the grandmaster and the ordinary clocks.

In our test scenario, the Symmetricom CsIII provided a reference source to the PTP Grandmaster, the role of which was filled by the Symmetricom TimeProvider 5000. The same reference source was also provided to a Calnex Paragon-X (combined frequency and phase analyzer) or a Spirent Anue 3500 (frequency analyzer) and a stand-alone phase analyzer. In order to emulate certain network delays, we configured impairment which was introduced between the Grandmaster and Ordinary clocks by either a Calnex Paragon-X or Spirent Anue 3500. The impairment profile was configured according to the test case 12 of the ITU-T G.8261.

The Ordinary Clock (OC) DUT was allowed to synchronize from free-running status under impairment. Once the DUT's clock output was stable, frequency and phase measurements ran for 4 hours.

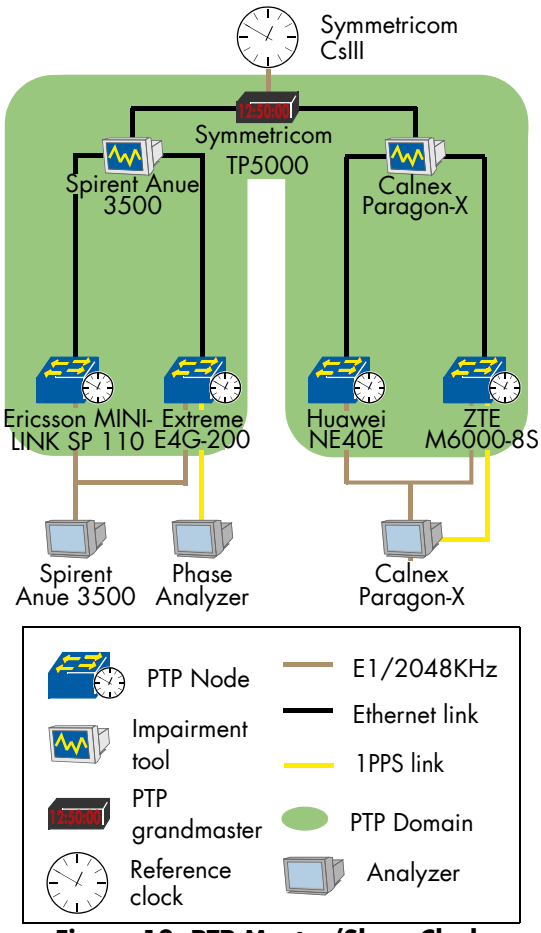


Figure 18: PTP Master/Slave Clock

Tested devices included: Ericsson MINI-LINK SP 110, Extreme E4G-200, and ZTE M6000-8S. All passed the MTIE G.823 SEC mask for frequency.

The Extreme E4G-200, the Huawei NE40E and the ZTE M6000-8S Ordinary Clocks passed the 16 ppb frequency accuracy requirement.

The Extreme E4G-200 passed the +/- 1.5µs phase deviation requirement.

The ZTE M6000-8S passed the +/- 5µs maximum phase deviation requirement.

Precision Time Protocol: Transparent Clocks

One option to improve synchronization quality at the location of the ordinary clocks is to deploy the transparent clock functionality on the network devices that are transporting the PTP messages, between the grandmaster and ordinary clocks.

The PTP transparent clock measures time that the PTP messages need to traverse the PTP transparent clock device, and adds this time to the value located in the correction field of the PTP messages.

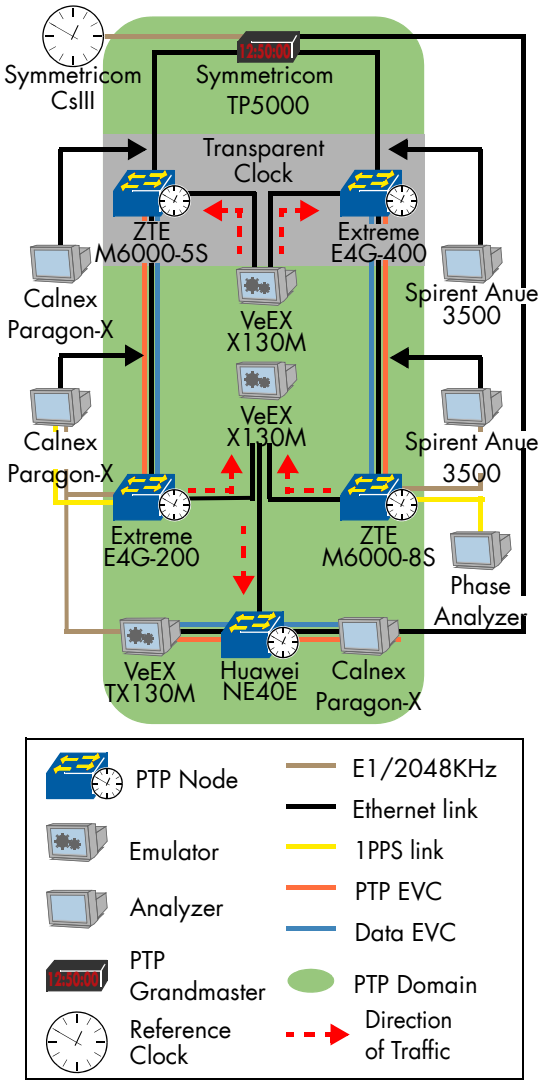


Figure 19: PTP Transparent Clock

As in the test with ordinary clocks, the Transparent Clock (TC) test featured a reference clock sourced from the Symmetricom CsIII. The PTP Grandmaster was once again the Symmetricom TP5000.

A Calnex Paragon-X or Spirent Anue 3500 was introduced on either side of the Transparent Clock, on one side in between the TC and the PTP Grandmaster, and on the other in between the TC and the PTP Slave. This was to ensure that the Correction Field on the PTP Sync packets was 0 before it reached the TC, and was adjusted to some acceptable value when leaving the TC for the PTP Slave device.

The PTP Slave device was allowed to synchronize to the PTP Grandmaster, using the offset measured in the Correction Field. In one test, VeEX demonstrated the ability of their device to analyze the correction field in PTP Sync packets and function as a PTP Slave on the same device.

Verification that this synchronization happened correctly was provided in the form of frequency and phase measurement. The devices used for this were a Calnex Paragon-X (combined frequency and phase) and a Spirent Anue 3500 (frequency measurement).

There were two parts to this test: one without any traffic, and one with test traffic flowing from the Transparent Clock to the Slave Clock at 50% line rate on a separate EVC from the PTP packets. This test traffic was introduced into the topology by a VeEX TX130M. Once the clock output was stable in each part frequency and phase measurements were allowed to run for half an hour.

Tested transparent clocks included: Extreme E4G-200, Huawei NE40E, and ZTE 6000-5S. Slave devices were Extreme E4G-200, VeEX TX130M, and ZTE 6000-8S. Vendor pairing is depicted in Figure 19. All Maximum Time Interval Error (MTIE) graphs for these tests passed the ITU-T G.823 SEC mask. In addition, none of the test traffic was dropped.

Precision Time Protocol: Boundary Clocks

Another option to improve synchronization quality at the location of the ordinary clocks is to deploy boundary clocks. This option can be used together with the transparent clock deployment.

A boundary clock (BC) has at least two PTP ports each in different PTP domains and maintains the timescale used in each domain. In one PTP domain BC synchronizes to another clock. It then acts as a slave or ordinary clock, while in another domain it serves as a source of time and also acts as a master clock.

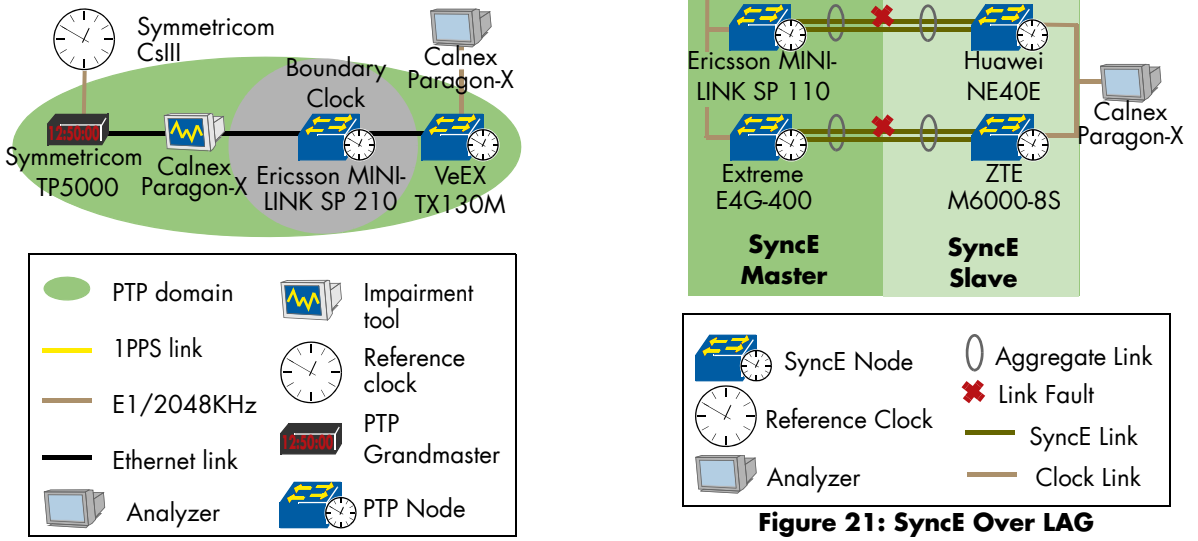


Figure 20: PTP Boundary Clock

Similar to the ordinary and the transparent clock tests, Symmetricom provided the CsIII as a reference clock and the TimeProvider 5000 as the PTP Grandmaster. Impairment was introduced between the Grandmaster and the Boundary Clock (BC) by a Calnex Paragon-X. The impairment profile was set to test case 13 of the ITU-T G.8261.

The BC and OC were allowed to synchronize from a free-running state under impairment, and the frequency offset of the OC was measured by the Calnex Paragon-X. After we observed the signal to be stable, the test measurement was allowed to continue for four hours.

There was one boundary clock test run: with the Ericsson MINI-LINK SP 210 boundary clock and VeEX TX130M ordinary clock. At the test's conclusion, the MTIE graph passed the ITU-T G.823 SEC mask.

There were a few interesting interoperability issues uncovered during testing. The most pronounced was that two vendors supported entirely different ranges of PTP Domain IDs, which meant that our options for BC and OC pairings were somewhat restricted.

Synchronous Ethernet over LAG

In a synchronized network it is very important that any changes in the physical topology do not affect the stability of the overall signal. When one link in a redundant LAG goes down, there would be undesired consequences on the final signal if it caused the frequency to be offset by much. This test was designed to ensure when a LAG member fails, the Synchronous Ethernet clock is not impaired.

Figure 21: SyncE Over LAG

The test was performed by establishing a LAG connection between a SyncE master node and slave node, and allowing the system clocks to synchronize and become stable. We began a measurement of the SyncE slave's frequency output to ensure that the synchronization state was indeed stable.

At this point, we began to alter the state of the links in the LAG. First the primary link was disconnected, then re-connected. Then the secondary link was disconnected and re-connected. At least one minute passed between each disconnect and re-connect, to ensure that any WTR (Wait To Restore) timers expired and the correct links were being used for the SyncE traffic. This was also verified in the DUTs' command line interfaces.

Once the standby link was reconnected, we continued to measure the SyncE slave's frequency for an additional 60 minutes with either a Calnex Paragon-X or Spirent Anue 3500 to ensure that no aberrations occurred later.

Vendor pairings are show in Figure 21. The DUTs that had passing results were: Ericsson MINI-LINK SP 110, Ericsson MINI-LINK SP 310, Extreme E4G-400, Huawei NE40E, Telco Systems T-Marc-3208, ZTE M6000-5S, and ZTE M6000-8S. All MTIEs for the frequency measurements passed the ITU-T G.8262 EEC Option 1 mask.

SyncE Islands Synchronization via PTP

This test was designed to verify clock transfer from one Synchronous Ethernet network to another across a larger PTP-enabled network.

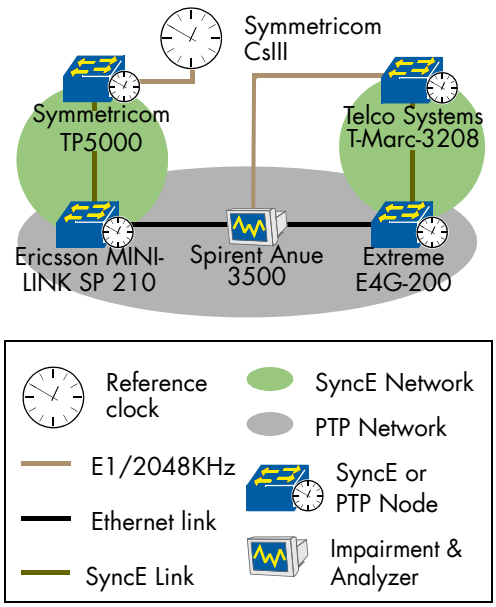


Figure 22: SyncE Island Synchronization via PTP

Symmetricom's CsIII was provided as a reference clock source. The role of the first SyncE Master was filled by the Symmetricom TimeProvider 5000. Its SyncE Slave and the PTP Master was the Ericsson MINI-LINK SP 210. The PTP Slave and second SyncE Master was the Extreme E4G-200. The second SyncE Slave was the Telco Systems T-Marc-3208. The impairment of PTP Sync messages and the frequency measurement was done by a Spirent Anue 3500.

Impairment was enabled on the Spirent Anue 3500, and all SyncE and PTP Slave nodes were brought from free-run to locked status in a little over 90 minutes. From there, the frequency measurement running on the same Spirent Anue 3500 was allowed to run for over four hours.

The MTIE graph compiled from the measurement passed the G.823 SEC mask.

Demonstration Network

During the two weeks of hotstaging we achieved many successful test results as reported in this paper.

Based on some of the results we created an end-to-end live demonstrations, which we present at the congress and reflect in our physical topology.

In the "PTP across multi-instance ERPS" demonstration we connected IEEE 1588-2008 PTP grandmaster and ordinary/boundary clocks to a multi-instance ERPS ring, doing frequency and phase synchronization. Ericsson MINI-LINK SP 110, Extreme E4G-200, Telco Systems T5C-XG, Cisco ASR 9006, Telco Systems T-Marc 3208SH, Extreme E4G-400, Symmetricom TP5000, Symmetricom CsIII, Spirent Anue 3500, VeEX TX130M and Calnex Paragon-X are part of this setup.

For the "Per CoS ID Performance Monitoring over IP/MPLS VPLS and ERPS" demonstration we show ERPS-VPLS interworking, where the VPLS is running over an IP/MPLS network. The demonstration includes performance monitoring per CoS ID. Ericsson MINI-LINK SP 210, Ericsson SPO 1410, Cisco ASR 9006, Telco Systems T-Metro 7224 and Telco Systems T-Marc 340 are participating.

The "IPv6 6rd over IP/MPLS and ERPS" demonstration shows 6rd CE and BR connected over an Layer 2 aggregation network. The aggregation network provides a resilient Ethernet Virtual Connection (EVC) between the CE and BR via ERPS and IP/MPLS VPLS. This demonstration includes Cisco ASR 1002, Cisco ASR 9006, Telco Systems T-Metro 7224, Ericsson SPO 1410, Cisco Linksys E4200 and D-Link DIR655.

The path computation demonstrated using Path Computation Element (PCE) between Metaswitch DC-PCE and Cisco ASR 9006 on the IP/MPLS aggregation network.

Furthermore, we demonstrate IP/MPLS L3VPN for IPv4 and IPv6 (6VPE). We connected CPEs with a dual stack PPPoE access to the 6VPE VRFs. The vendors participating in this demonstration are: ZTE M6000-8S, Huawei ME60, Ericsson SE1200, Cisco ASR 1002, Ericsson SE100, Cisco ASR 9006 and Ixia IxNetwork.

Finally we created a demonstration that shows interworking between IP/MPLS VPLS and MPLS-TP VPLS. Hitachi AMN1710, Ericsson MINI-LINK SP 310, Ericsson SPO 1410 and Cisco ASR 9006 are participating.

Acknowledgements

We would like to thank Kari Asheim from nLogic AS and Stephen Murphy from UNH-IOL for their support during the testing.

Editors. This report has been edited by Kari Asheim, Sergej Kälberer, Stephen Murphy, Ronsard Pene, Carsten Rossenhövel, Shima Sajjad and Eldad Zack.

Acronyms

Term	Definition
6rd	IPv6 Rapid Deployment on IPv4 Infra-structures
AC	Access Circuit
AFTR	Address Family Transition Router
AIS	Alarm Indication Signal
ARP	Address Resolution Protocol
B4	Basic Bridging BroadBand
BC	Boundary Clock
BFD	Bidirectional Forwarding Detection
BFD-CC	BFD Continuity Check
BGP	Border Gateway Protocol
BR	Border Router
CCM	Continuity Check Message
CE	Customer Edge
CFM	Connectivity Fault Management
CGN	Carrier-Grade NAT
CLI	Command Line Interface
CoS	Class of Service
CPE	Customer Premises Equipment
DAD	Duplicate Address Detection
DHCP	Dynamic Host Configuration Protocol
DHCPv4	Dynamic Host Configuration Protocol for IPv4
DHCPv6	Dynamic Host Configuration Protocol for IPv6
DMM	Delay Measurement Message
DNF	Do Not Flush
DNS	Domain Name System
DS-Lite	Dual-Stack Lite
DUT	Device Under Test
ERPS	Ethernet Ring Protection Switching
ETHDi/ ETH_A	Ethernet Diagnostic/Ethernet Adoption function
EVC	Ethernet Virtual Connection
FDB	Filtering Database
GAL	Generic Associated Label
GUI	Graphical User Interface
HL	Hop Limit
IA	Identity Association
iBGP	Internal Border Gateway Protocol
ICMPv6	Internet Control Message Protocol for IPv6
IETF	Internet Engineering Task Force
IPCP	IP Control Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
IPV6CP	IPv6 Control Protocol
L2TP	Layer Two Tunneling Protocol
L3VPN	Layer 3 Virtual Private Network
LAC	L2TP Access Concentrator
LAG	Link Aggregation
LAN	Local Area Network
LDI	Link Down Indication
LKR	Lock Report
LMM	Loss Measurement Messages
LMR	Loss Measurement Replies
LNS	L2TP Network Server
LOS	Loss of Signal
LSP	Label Switched Path
MAC	Media Access Control
MEP	Maintenance entity group End Point
MP-BGP	Multiprotocol Border Gateway Protocol
MPLS	Multiprotocol Label Switching
MPLS-TP	MPLS Transport Profile
MTIE	Maximum Time Interval Error
NAT44	Network Address Translation - IPv4 to IPv4
NDP	Neighbor Discovery Protocol
NR	No Request
OC	Ordinary Clock
OSPF	Open Shortest Path First
PCC	Path Computation Client
PCE	Path Computation Element
PCEP	PCE Communication Protocol
PCP	Port Control Protocol
PD	Prefix Delegation
PE	Provider Edge
PPP	Point-to-Point Protocol
PPPoE	PPP over Ethernet
PSC	Protection State Coordination
PTP	Precision Time Protocol
PW	Pseudo Wire
R-APS	Ring Automatic Protection Switching
RB	RPL Block
RFC	Request for Comment
RPL	Ring Protection Link
SEC	SDH Equipment Clock
SF	Signal Failure
SLA	Service Level Agreement
SLAAC	Stateless Address Auto-configuration
SSM	Synchronization Status Messages
TC	Transparent Clock
TCP	Transmission Control Protocol
TE	Traffic Engineering
TTL	Time to Live
UDP	User Datagram Protocol
VLAN	Virtual Local Area Network
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
WAN	Wide Area Network
WTR	Wait-to-Restore Timer

References

"Transmission of IPv6 Packets over Ethernet Networks", RFC 2464
"Layer Two Tunneling Protocol (L2TP)", RFC 2661
"Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315
"IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633
"Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736
"IP Version 6 Addressing Architecture", RFC 4291
"Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443
"BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659
"Neighbor Discovery for IP version 6 (IPv6)", RFC 4861
"IPv6 Stateless Address Auto-configuration", RFC 4862
"IP Version 6 over PPP", RFC 5072
"Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440
"IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)", RFC 5569
"IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) – Protocol Specification", RFC 5969
"Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333

"Ethernet ring protection switching", ITU-T G.8032 version 2
"MAC Address Withdrawal over Static Pseudowire", draft-boutros-pwe3-mpls-tp-mac-wd, work in progress.
"Bidirectional Forwarding Detection (BFD)", RFC 5880
"MPLS Transport Profile Data Plane Architecture", RFC 5960
"A Framework for MPLS in Transport Networks", RFC 5921
"MPLS-TP Linear Protection", RFC 6378
"MPLS Fault Management Operations, Administration, and Maintenance (OAM)", RFC 6427
"Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428
"Precision Time Protocol (PTP)", IEEE 1588-2008
"Port Control Protocol (PCP)", draft-ietf-pcp-base
"The Control of Jitter and Wander Within Digital Networks Which Are Based On 2048 kbit/s Hierarchy", ITU-T G.823
"The Control of Jitter and Wander Within Digital Networks Which Are Based On 1544 kbit/s Hierarchy", ITU-T G.824
"Timing characteristics of synchronous Ethernet equipment slave clock (EEC)", ITU-T G.8262
"OAM Functions and Mechanisms for Ethernet Based Networks", ITU-T Y. 1731

EANTC AG
European Advanced Networking Test Center

Salzufer 14
10587 Berlin, Germany
Tel: +49 30 3180595-0
Fax: +49 30 3180595-10
info@eantc.de
http://www.eantc.com

upperside conferences

Upperside Conferences

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

20120131 v1.0