



European Advanced Networking Test Center
Aktiengesellschaft



***Progressing Advanced
MPLS & Carrier Ethernet
Solutions***

***Public Multi-Vendor
Interoperability Test***

Paris, February 2011

EDITOR'S NOTE



Carsten Rossenhövel
Managing Director

Since our first MPLS multi-vendor interoperability event in Paris in 2003, we have tested virtually all known areas of MPLS and Carrier Ethernet technology for multi-vendor interoperability. We have witnessed major steps of evolution and have shared our experience in testing of many innovations with the great audience at MPLS & Ethernet World Congresses publicly.

It seems to me that MPLS and Carrier Ethernet, compared with the invention of the wheel, have reached the equivalent of the steel wheel with pneumatic tires now (in doubt, see front page for reference). At a time, mankind considered these wheels perfect. They certainly are very good and continue to be used. Nevertheless, innovation must continue for new applications and for better integration into the bigger picture.

To keep focused on those innovations, we rejected testing requests for mature protocols such as TDM / ATM over MPLS, basic Carrier Ethernet and VPLS services and the pre-draft standard MPLS-TP solutions with Y.1731-based protection which we had all tested for a few years.

The strategy proved useful: Four vendors brought new BFD-based pre-standard MPLS-TP implementations to the test which revealed interoperability issues based on ambiguous draft IETF standards, for example in BFD *slow start* procedures and the use of a control word in labels. More standardization work is required before the BFD-based implementations can be deployed in multi-vendor environments.

Fault management protocols for MPLS-TP such as LSP ping and traceroute require more testing in future to reach multi-vendor interoperability as well. It seems it is not trivial to retrofit existing IP/MPLS protocols to the new MPLS-TP suite.

On the Carrier Ethernet service side of things, performance monitoring based on Y.1731 showed great progress once more. Four of the Ethernet OAM implementations under test supported loss measurement in addition to delay measurements and worked mostly fine.

Ethernet ring protection based on G.8032:2010 was another area of great innovation. Since our last test event, the second edition of the standard has been completed. An overwhelming number of nine implementations all supported the new standard and tested with multiple rings. Reconvergence was always achieved below 300 milliseconds, which is great for non-MPLS networks. Interestingly, reconvergence following loss of carrier took longer due to MAC address flooding than the reconvergence in more advanced scenarios.

Service providers had urged us to verify an important design challenge: The integration of Optical Transport Networks (OTN) with MPLS and

Carrier Ethernet packet networks in multi-vendor networks. We invited vendors to join with their router and OTN solutions. Two vendors brought their solutions, one based on WDM systems interconnected with the router transparently, another system based on MPLS-TP integration of the transmission platform. A few of the vendors we had expected surprisingly declined for the second time — maybe it will take the industry some more time to develop consistent designs for standards-based integration of transmission and packet transport solutions.

For the first time, we evaluated 100 Gigabit Ethernet in a multi-vendor environment with two router and one test equipment vendor. The physical layer worked mostly fine in our interop tests. In one case, we noticed packet loss at low throughput that could not be resolved. Obviously we conducted only basic functional tests of MPLS services this time.

Finally, we continued our campaign of packet synchronization testing for mobile backhaul with more complex scenarios than before. This time, EANTC's test team raised the bar, requiring participants to pass the more stringent SEC mask of G.823 — in previous events we consented to the PDH traffic mask. The new requirement is specifically important for LTE deployments, and to increase overall confidence in the non-trivial IEEE 1588 solutions.

Major efforts went into testing transparent and boundary clocks in multi-vendor scenarios this time. It showed again that slave and boundary clock implementations are by far not trivial. Achieving multi-vendor interoperability and service performance between grandmaster clocks and slave clocks is difficult. Transparent clocks face another issue: There are cheap but very inaccurate implementations possible by setting the *correction field* statically. In a few cases, we were able to confirm that vendors had implemented dynamic delay correction.

Given the diverse and advanced areas of testing, it was not a surprise to see interoperability issues. Some were already resolved on site; most will be fixed by vendor teams in the coming weeks. It showed that there continues to be a need for interoperability testing at the forefront of MPLS and Carrier Ethernet technology, to ease and speed up multi-vendor deployments in service provider networks.

Our team and I hope that this report will be valuable and are happy to answer any additional questions.

TABLE OF CONTENTS

Participants and Devices on page 3
Demonstration Network on page 3
Interoperability Test Results on page 3
MPLS Services and Transport on page 4
Service Management on page 6
Clock Synchronization on page 9
Resiliency Tests on page 13
Acronyms on page 17
References on page 18

PARTICIPANTS AND DEVICES

Vendor	Devices
Brocade	MLXe Router NetIron CER
Calnex	Paragon Paragon-X
Cisco	7604 7606 ASR 9006 ASR 9010 CPT 50 CPT 600 Prime Network
Ericsson	OMS 1410 Mini-Link CN1010 SE 1200 SEA 10 SEA 20
Hitachi	AMN1710
Huawei	ATN910 ATN950 ATN980 ATN990 CX600-X1 CX600-X2 CX600-X3 NE40E-X8 NE5000E OSN550 OSN6800 OSN7500 II
Ixia	IxNetwork
Juniper	T1600
Metaswitch	DC-MPLS
Spirent Communication	Spirent TestCenter Spirent GEM Spirent XGEM
Symmetricon	Cesium Reference CsIII TimeProvider 500 TimeProvider 5000 SSU 2000e

Vendor	Devices
Telco Systems — a BATM Company	T5C-XG T-Marc-254H T-Marc-3208SH T-Marc-380 T-Metro-7124S T-Metro-7224
Vitesse	VSC7428 VSC7460

DEMONSTRATION NETWORK

Please see the physical topology on page Topology (centerfold) for reference.

With such a vast amount of devices and test cases, naturally they were accomplished in separate parallel test scenarios. Towards the end of our two weeks of testing, we stitched together several successful test scenarios to have a single physical topology which could demonstrate multiple test results, also in parallel. Therefore the diagram found in the centerfold of this document represents a small set of those results detailed in this whitepaper report, as planned to be presented at the MPLS & Ethernet World Congress in Paris 2011.

INTEROPERABILITY TEST RESULTS

In the following sections of the white paper we 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 run our tests we were required to generate, measure, impair, and analyze Ethernet traffic and perform synchronization analysis. We thank Calnex Solutions, Ixia, Spirent Communications, and Symmetricon for their test equipment and support throughout the hot staging.

MPLS SERVICES AND TRANSPORT

MPLS networks continue to grow and expand, and as new services are added, so are new standards including for more advanced services but also for higher interface speeds which are needed now more than ever. We have also focused on interworking of IP/MPLS and MPLS-TP, as well as testing of some BGP functionality for the IP/MPLS area.

MPLS Services over 100 Gigabit Ethernet

The need to transport more than 10 Gigabits/second of Ethernet traffic (10GE) over links in core networks has been present for some years, and so far bundling 10GE links has solved this, but to keep on increasing these bundles will have some practical limits. Also freeing up fiber by replacing these aggregated links is of great value to the providers owning or leasing them. With the ratification of the IEEE 802.3ba standard last summer, including a number of variants for the physical layer for 40G/100G Ethernet, there is a way to utilize higher bandwidth on single links.

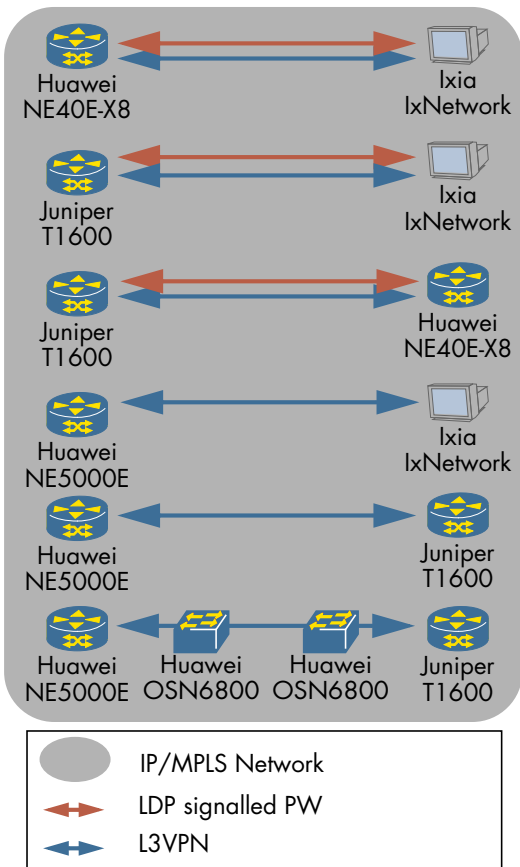


Figure 1: MPLS over 100 Gigabit Ethernet Tests

Our goal was to verify if one can expect to run the same MPLS services in deployment today using the 100GE interfaces currently available. Vendors brought 100GBASE-LR4 interfaces which utilize four

wavelengths of 25G on the fiber. In this test scenario the 100GE links were set up as core-facing links in an IP/MPLS network, and MPLS services were configured between the devices under test, transporting MPLS labeled traffic across the 100GE link.

There were successful test runs between the Huawei NE40E-X8, Huawei NE5000E, Ixia LxNetwork, and Juniper T1600. Juniper and Huawei set up Layer 2 Pseudowires (L2 PWs) and Layer 3 Virtual Private Network (L3VPN) services and forwarded traffic generated and verified by Ixia. Ixia LxNetwork was also used to emulate a Provider Edge router (PE) with L2 PW and L3VPN to interoperate with both Huawei and Juniper.

Huawei successfully demonstrated traffic transmission over 100 Gigabit Ethernet interfaces using both router and transmission equipment. Dense Wave Division Multiplexing (DWDM) was used over Huawei OSN6800 with 100GBASE-LR4 optics for the client side and Optical channel Transport Unit-4 (OTU4) interfaces on the network side running over two pairs of fiber using the 50 Ghz spacing system as defined in ITU-T recommendation G.694.1, and using Optical Channel data Unit-4 (ODU4) for conversion according to standard ITU-T G.709 internally.

IP/MPLS and MPLS-TP Interworking

As MPLS-TP continues to evolve, a key focus in standardization has been interoperability between the progressing MPLS-TP standards and current IP/MPLS deployments. This can be done either through a handoff between the domains, or a stitching implemented within a single device. In our testing vendors stuck with the latter configuration.

In this test, multi-segment pseudowires (MS-PW) were built between the IP/MPLS network and the MPLS-TP network. The pseudowire segments in the IP/MPLS network were signaled using LDP while the MPLS-TP network relied on static pseudowire segments. Bidirectional Forwarding Detection (BFD) for Operations, Administration and Management (OAM) was used in the MPLS-TP domain. One device was then connected to both networks and implemented the switching of the static and dynamic pseudowire segments. The two PE devices terminating the PWs were each connected to a tester, which was used to generate traffic.

Amongst the tests, Cisco ASR9010 and Metaswitch DC-MPLS each performed the stitching of PWs. Brocade NetIron CER, Ericsson SE1200 and Telco Systems T-Metro-7224 were used as PEs for LDP signaled PWs, while Cisco CPT600, Ericsson SE1200, and Ixia LxNetwork established PWs in the MPLS-TP network.

The control word - a mechanism used in PWs to ensure that packets arrive in the correct order - is negotiated within LDP when the PW is established up, along with the exchange of labels. The use of the control word in the MPLS-TP domain is defined by RFC 5586 and further explained in RFC 5960. According to these RFCs the control word is

mandatory during use of the Generic Associated Channel (G-ACh) mechanism when the OAM session is run within the PW, to differentiate control packets from data packets. When the OAM session is run within the PW, G-ACh uses the reserved G-ACh Label (GAL) - label 13, without the control word. More recently, draft-ietf-pwe3-mpls-tp-gal-in-pw removes the mandate from RFC 5586. In two test runs the control word was present on the MPLS-TP side and not used on the IP/MPLS side, while in a third test run both PWs were not using the control word.

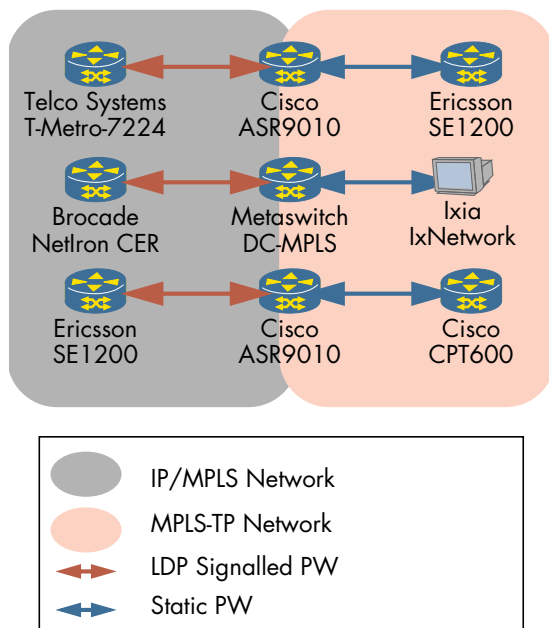


Figure 2: IP/MPLS and MPLS-TP Interworking Tests

In addition to testing the interworking between IP/MPLS and MPLS-TP, Metaswitch demonstrated MPLS-TP signalling between two Metaswitch DC-MPLS devices, using Generalized MPLS (GMPLS) signaling, Targeted LDP (T-LDP), and OSPF-TE. BFD was used for OAM on the LSP, where the BFD packets were encapsulated using the GAL label and G-ACh header. Traffic was generated and verified to be on the PW, using a tester. Captures from the Metaswitch equipment showed that data and signaling were using separate physical links.

BGP 4-Byte AS Numbers

When standardizing BGP in the early days, Autonomous System Numbers (ASN) were two Bytes long. This limits the total possible number of ASNs to 65536. In order to avoid running out of ASNs, Internet Assigned Numbers Authority (IANA) and the Regional Internet Registries have been assigning 4-byte ASNs by default since January 1, 2009. It is still possible to request a 2-byte ASN, but there will likely be a point where they are depleted. In order to maintain 4-byte ASN paths through BGP routers only capable of 2-byte ASNs, a special AS number 23456 (AS_TRANS) is reserved for use in the 2-byte AS_PATH and the transitive BGP-attribute

AS4_PATH is used to encode the 4-byte ASNs. In this test we verified the conversion between these two AS-PATH formats.

Three routers were set up in a row, where the middle router performed the conversion between 2-byte and 4-byte ASN, the 4-byte ASN being an IP/MPLS network. In some test runs the edge routers were emulated by testers. Testers were used to verify connectivity and to capture BGP updates for verification of the content within the BGP attributes.

Brocade, Ericsson, Huawei, Ixia and Spirent successfully participated in this test as depicted in Figure 3: BGP 4-Byte AS Number Tests. Router vendors showed that the conversion between a 2-byte and a 4-byte peering worked, while the test equipment vendors showed that they could emulate both a 4-byte and a 2-byte BGP router.

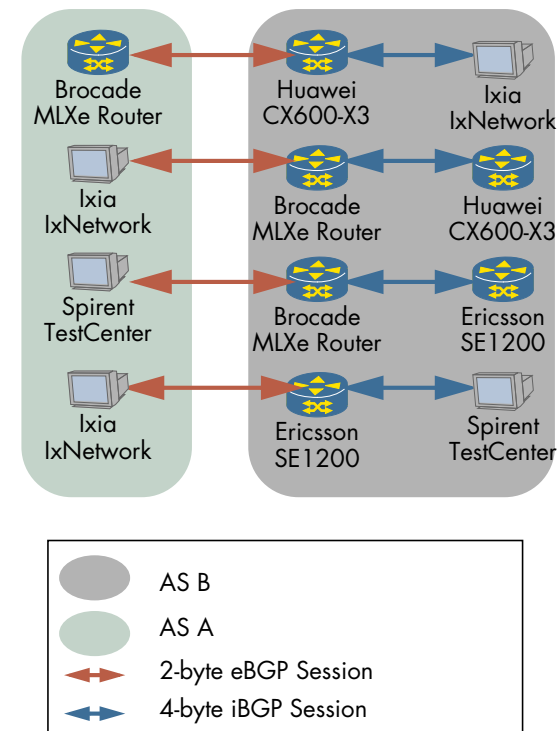


Figure 3: BGP 4-Byte AS Number Tests

During the testing we noted that once routers were upgraded to newer software supporting 4-byte ASNs, they would use 4-byte ASNs as default. One vendor implemented the ability to turn off negotiation of the 4-byte ASN capability. In one scenario we used this to emulate a 2-byte-only-speaker, though the router still displayed 4-byte ASNs in the AS-PATH.

BGP Multihoming

BGP multihoming is used between provider networks to create redundant connections between two Autonomous Systems (ASes), requiring that the ASes are contiguous. It is possible to remove this constraint, overriding the AS-PATH advertised from the split AS, which is also referred to as AS override. This should be used with caution in order to not introduce routing loops. If a single site should be multi-homed

and the AS number is overridden, a routing loop will occur. To avoid this the Site-of-Origin extended community attribute has been defined to distribute the information between the PEs, so the PEs can make the decision regarding which routes should and should not be distributed to a given Customer Edge router (CE).

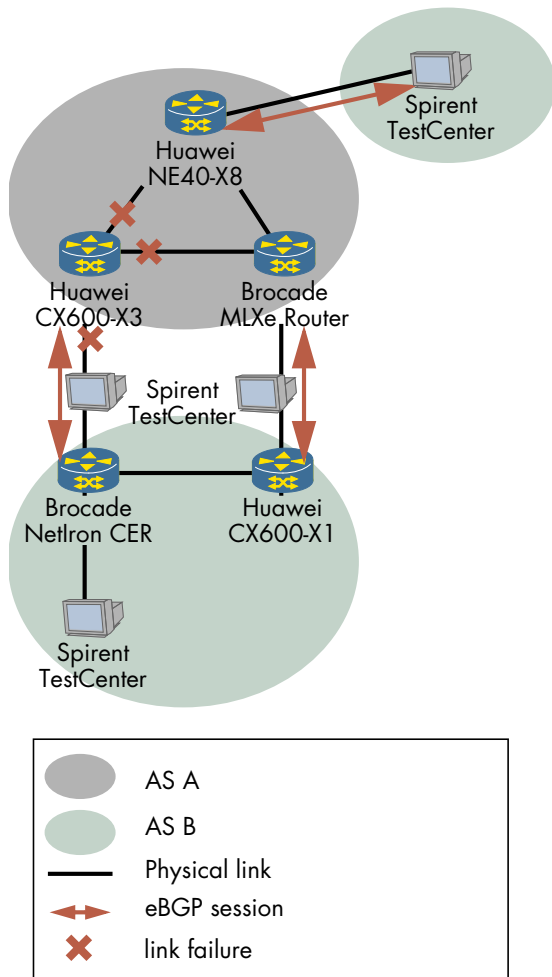


Figure 4: BGP Multihoming Test

In this test AS "A" was an IP/MPLS network connecting two disjoint IP networks belonging to the same AS (AS "B") — a potentially typical setup for a large enterprise customer using a L3VPN service from a provider for backbone connections. The AS number announced from the CE was overridden in the IP/MPLS network, and the PE routers used different Site-of-Origins for each site.

Brocade MLXe Router, Huawei CX600-X3 and Huawei NE40E-X8 were set up as PE routers in AS A, configured to override the AS with two different Site-of-Origin IDs to reflect the two customer sites. In AS B OSPF was used as the IGP, and eBGP was used on the PE-CE links. Routes were then redistributed between BGP and OSPF on the CEs. On one site Brocade Netron CER and Huawei CX600-X1 were used as CEs and multi-homed to the IP/MPLS L3VPN service. A Spirent TestCenter was used to emulate the CE on the other site by running eBGP with the Huawei NE40E-X8. The Spirent TestCenter also emulated a router within the AS B site which

was multihomed by advertising IP routes via OSPF. Finally the Spirent TestCenter was used in-line to monitor the PE-CE links to do packets capture and observe that the flow of routing information was correct.

The link between the Huawei CX600-X3 and the Brocade Netron CER was configured with a lower MULTI_EXIT_DISC (MED) than the other PE-CE link, so that it was the preferred path. We verified that the customer prefixes were correctly advertised and not re-advertised back to AS B. The three links on the Huawei CX600-X3 were shut down to mimic a failure of the entire router (since other tests were going on at the same time, it was not possible to power off the router as specified in our test plan). The traffic failed over to the backup PE-CE link, and a small amount of packet loss was observed as expected. This test focused on BGP operation and distribution of routes, and therefore techniques to do faster failover such as the tweaking of counters to lower rerouting time were not used. When restoring the three links on the Huawei CX600-X3, we saw no packet loss as the traffic reverted to the preferred path.

SERVICE MANAGEMENT

Alongside the development of modern Ethernet services and MPLS-TP based transport there is a continuing development of management protocols to provide the necessary tools for monitoring, statistics collection and troubleshooting. This time around, the tests focused on tools for end-to-end performance, MPLS-TP Operations, Administration and Maintenance (OAM), and interworking between traditional dynamic MPLS and MPLS-TP OAM protocols.

Performance Monitoring

In a service provider environment, performance monitoring is an important tool for operations staff to give an early warning of network issues. Packet loss and changes in delay indicate that unwanted events are taking place in the network, and being able to monitor and detect such events enables the operations staff to take action before it becomes a real problem for customers. The tools defined in ITU-T Y.1731 are not only useful for this, but also proving Service Level Agreement (SLA) attributes to their customers.

The test setup connected two devices through an impairment tool, and an E-Line service provisioned between them. An OAM session was then configured over that E-Line. A traffic generator added traffic to the service while the impairment tool introduced frame loss, delay or delay variation depending on the type of measurement being tested.

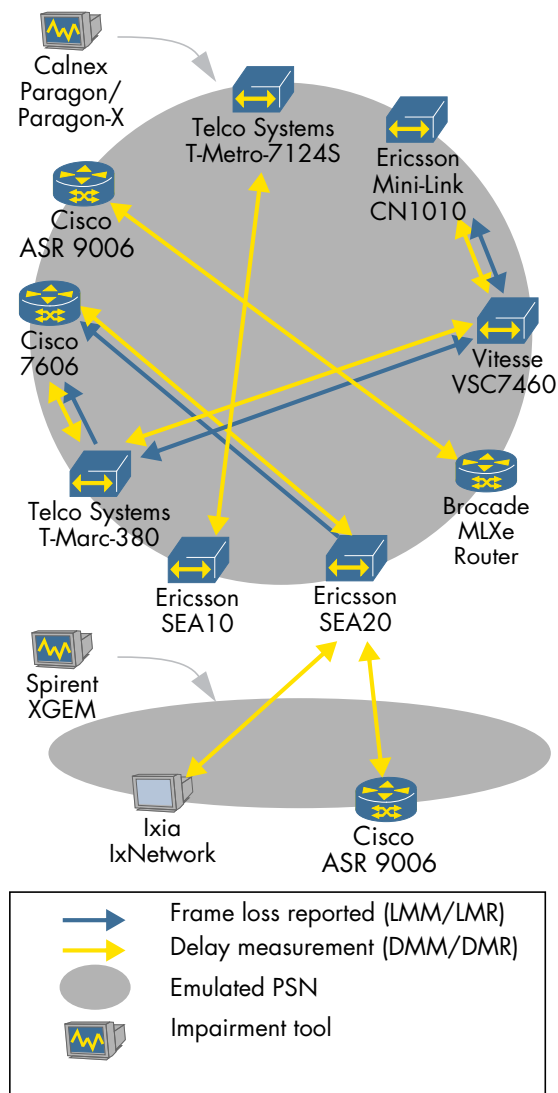


Figure 5:
Y.1731 Performance Monitoring Tests

Ixia IxNetwork and Spirent TestCenter were used to verify loss measurements, while the Calnex Paragon, Calnex Paragon-X and Spirent XGEM were used for impairment. Ixia IxNetwork was tested for interoperability, as shown in Figure 5: Y.1731 Performance Monitoring Tests.

Two-way frame loss measurements using Loss Measurement Message (LMM)/Loss Measurement Reply (LMR) OAM frames were successfully tested on the Cisco 7606, Ericsson SEA20, Ericsson Mini-Link CN1010, Telco Systems T-Marc-380 and Vitesse VSC7460. In one case, the receive count was not inserted properly into the LMM/LMR frame, thus causing both devices in that test to always report 100% frame loss in the direction from that device.

Delay Measurement Message (DMM)/Delay Measurement Reply (DMR) OAM frames were used to measure delay and delay variation measurements by Brocade MLXe Router, Cisco 7606, Cisco ASR 9006, Ericsson Mini-Link CN1010, Ericsson SEA10, Ericsson SEA20, Ixia IxNetwork, Telco Systems T-Marc-380, Telco Systems T-Metro-7124S and Vitesse VSC7460. During the delay and delay

variation tests, we noticed that the measurements performed by most vendors are now so accurate that their results are within a few tens of microseconds from each other.

Most tests were performed over VLAN interfaces using Tag Protocol ID (TPID) 0x8100 or 0x88a8, but additionally, the test between Cisco ASR 9006 and Brocade MLXe Router was performed over an MPLS LDP-signalled pseudowire (PW) as well as over a VLAN.

We discovered that although all the devices performed the measurements in the same way, results were presented to the user in different ways. Some devices performed measurements on demand, and therefore did not calculate running averages, some did not record min/max values, and some presented the round trip delay while others presented the delay in each direction. This implies that comparing measurements between two different vendors is possible, but requires a good understanding of each respective vendors' reporting.

MPLS-TP OAM

Monitoring and troubleshooting layer two services and network links can be a challenge when the appropriate tools are not available. Tools like OAM ping and traceroute may be used to isolate failures, thus simplifying troubleshooting and reducing the time to fix an outage. Continuity checks additionally enable rapid detection and notification of failures on a virtual or physical connection. Our test was designed to verify the tools for these respective purposes, allowing vendors to be flexible on the specific protocol used, based on support. In the end the focus was on those OAM tools as defined by the IETF MPLS working group for MPLS-TP. While EANTC has tested multiple author drafts of MPLS-TP OAM in the past (all whitepaper reports remain online), we focused on newer implementations based on draft-ietf-cc-cv-rdi and draft-ietf-mpls-tp-on-demand-cv.

Vendors configured scenarios within an MPLS-TP network consisting of three devices, with an MPLS PW over an LSP. Proactive Continuity Check (CC) or Connectivity Verification (CV) was used for continuity checks and MPLS on-demand CV for ping-like tests. Trace tests were done using MPLS on-demand Route Tracing. An impairment tool was then used to create 100% packet loss in order to simulate failure without causing the devices to experience a Loss of Signal (LOS) on the physical Ethernet line. We then verified that the failure was properly detected and reported by the continuity checks. We tested OAM ping and traceroute in a similar fashion with impairment, and recorded also the ability to locate failures.

Ixia IxNetwork and Spirent TestCenter were used to generate traffic, while Calnex Paragon, Calnex Paragon-X and Spirent XGEM were used for impairment. Ixia IxNetwork was additionally tested for interoperability, as shown in Figure 6: MPLS-TP OAM Tests.

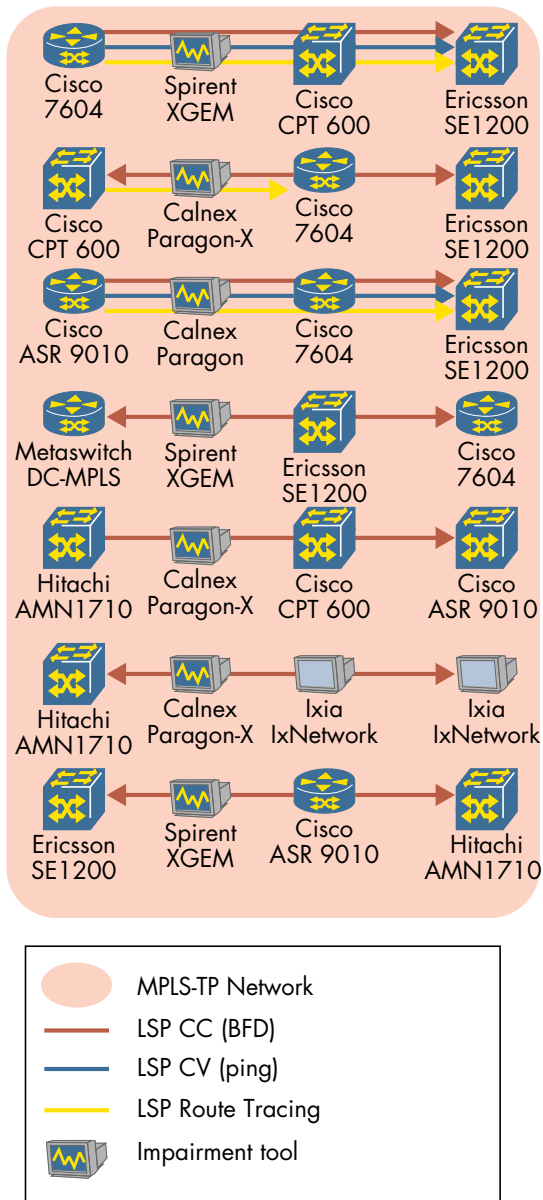


Figure 6: MPLS-TP OAM Tests

Successful tests of CC with Bidirectional Forwarding Detection (BFD) on the LSP according to draft-ietf-cc-cv-rdi-00 were conducted on the following devices: Cisco 7604, Cisco ASR 9010, Cisco CPT 600, Ericsson SE1200, Hitachi AMN1710, Ixia IxNetwork and Metaswitch DC-MPLS. Arrows in the diagram are drawn from the device that showed the expected state or result. Since unidirectional failures were introduced, we expected one device to show a BFD "down" state, and the other as "init". This was not always the case. In some cases, even though BFD successfully detected a failure, the device did not bring down the service. Instead, traffic continued to flow and was received by the tester in the direction without impairment.

Cisco ASR 9010 and Hitachi AMN1710 also tested interoperability using version 02 of draft-ietf-cc-cv-rdi, where the BFD session initiation procedures have been changed. The draft no longer specifies a poll/final sequence at the end of the slow start initialization phase.

MPLS on-demand CV pings from draft-ietf-mpls-tp-on-demand-cv were performed on the Cisco 7604, Cisco CPT 600, Cisco ASR 9010 and Ericsson SE1200. The draft specifies different ways of sending the pings, and since vendors implement different encapsulations from the draft, not all combinations were able to produce the expected results. In addition, some tests failed because both ends did not understand the same set of Type-Length-Value (TLV) options. Finally, we successfully tested the following devices for route tracing according to draft-ietf-mpls-tp-on-demand-cv: Cisco 7604, Cisco CPT 600, Cisco ASR 9010 and Ericsson SE1200.

OAM Failure Propagation

End-to-end services are often built either over multiple network domains, or over a single large network domain consisting of multiple technology domains. For example, different transport technologies could be used in the aggregation and core networks. This requires a unified approach to OAM; with the interconnecting node mapping failure status between one protocol and another - in our case, between IP/MPLS mechanisms, and MPLS-TP OAM in order to enable a seamless end to end service.

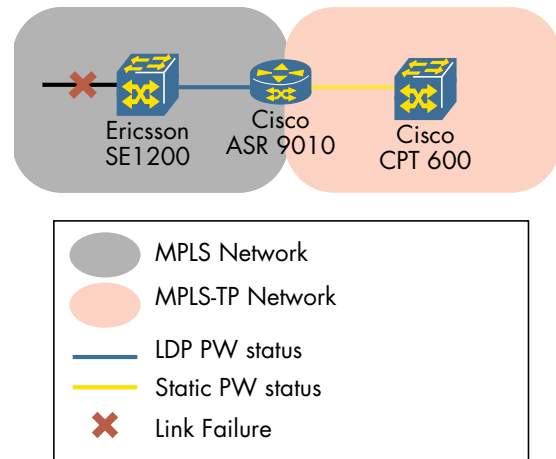


Figure 7: OAM Failure Propagation

In order to test this scenario vendors connected three devices in series, and a network service across them. At one end, a device was configured to propagate a physical link failure into an OAM protocol towards its peer, and this was expected to then propagate the failure status using another OAM protocol to the device at the far end of the physical link failure, which was expected to display an alarm.

In the test, failure propagation was performed between a dynamic LDP-signaled PW stitched to a statically configured PW in the MPLS-TP domain. The Cisco ASR 9010, Cisco CPT 600 and Ericsson SE1200 participated in this test. We were able to see that generating LOS on the attachment circuit would bring the LDP PW down, and that the fault was propagated properly to the static PW using draft-ietf-pwe3-static-pw-status, where the failure was reported by the far end device.

CLOCK SYNCHRONIZATION

Over three years have passed since EANTC started testing clock synchronization at our interop events, and there's no signs of slowing down. If the mobile network industry and the need to synchronize base stations' air interfaces was not already a big enough need for a robust solution to synchronize packet networks, additional industries have similarly identified such requirements. With such a growing market, the IEEE 1588 standard has also included a breadth of complex options and functional roles to keep a test lab busy for years. In this event we added tests for such roles - Boundary Clock and Transparent Clock - in addition to our continuation of tests for both frequency and phase on the Ordinary Clock (slave).

IEEE 1588 Master/Slave - Phase and Frequency Synchronization

Before moving devices onto Boundary Clock and Transparent clock tests, we started with a base phase and frequency synchronization test based on IEEE 1588 Precision Time Protocol (PTP). Using a Symmetricom TimeProvider 5000 as the PTP Grandmaster, we tested PTP slaves (Ordinary Clocks, or OCs) one by one, verifying that they were able to lock to the Grandmaster's reference by comparing the slave's output to that of the Symmetricom CsIII (cesium) primary reference clock. Measurements were performed by connecting each slave device to a wander analyzer for frequency, via either E1 or 2048 KHz interfaces, and a frequency counter via 1 PPS (pulse per second) for phase (time of day, or TOD) deviation measurement. All test scenarios were required to pass the ITU-T G.823 SEC mask for MTIE and TDEV for frequency synchronization and a maximum time error of 3 μ s for the time of day deviation. A Telco Systems T5C-XG was used to distribute the clock (PTP packets) to the slaves.

Impairment was introduced into the network according to ITU-T G.8261 (Test Case 12) by using the Spirent XGEM, the Calnex Paragon or the Calnex Paragon-X, generating delay between 59–206 μ s in the direction from the Grandmaster to the slave device, and 59–165 μ s in the reverse direction. After a device was first verified to be free running, we first enabled impairment and then allowed the clock to lock. Once the slave indicated that its clock was locked to the grandmaster, we ran measurements for frequency and/or phase for at least four hours minimum.

Multiple "profiles" were configured on the grandmaster, each on a different VLAN, to accommodate the various protocol implementation differences between vendors such as PTP over multicast/unicast, different rates of PTP messages, and 1 Step vs 2 Step. Using multicast PTP the Ericsson Mini-Link CN1010, Ericsson SEA10 and Telco Systems T-Marc-254H passed the frequency synchronization test. In addition, the Ericsson Mini-Link CN1010 and the Ericsson SEA10 passed the test for phase. Both

Cisco 7606 and Huawei ATN910 passed the frequency and the phase tests using unicast PTP.

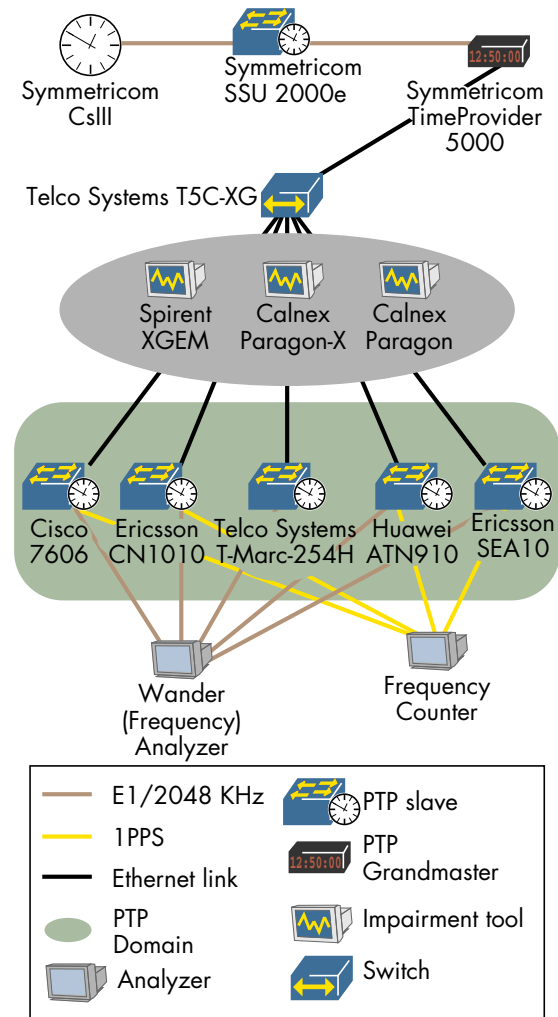


Figure 8: Master/Slave Clock Tests

IEEE 1588 Transparent Clocks

To help increase the ability for the slave-only ordinary clocks to support the required clock quality, without being impaired by the series of nodes that exist in the network between the slave and the master, the standard defines ways of bringing PTP awareness throughout exactly those network nodes. One of such roles is that of a Transparent Clock (TC) - a network node which acknowledges the delay it adds to PTP frames by entering that delay into a field within the packet called the Correction Field (CF). This field should then be used by the slave in its calculation for network delay and adjust its clock accordingly.

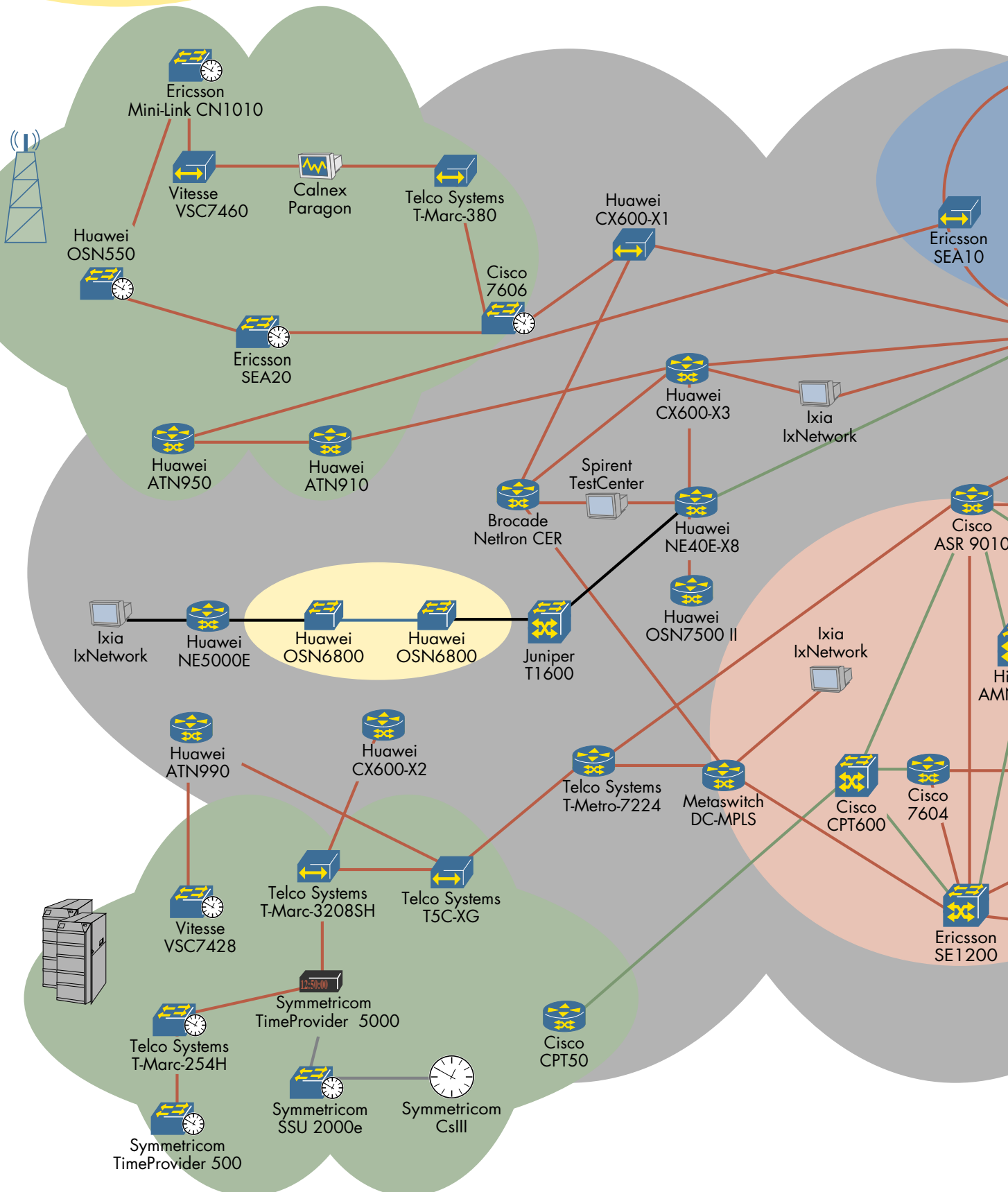
For our tests, we placed all Transparent Clock devices between the Symmetricom TimeProvider 5000 PTP Grandmaster and the slave device. To verify TC functionality, the Correction Field in the Grandmaster's PTP Sync packets was checked before and after the TC device with a Calnex Paragon and Paragon-X. This was to ensure that the Correction Field was 0 before it reached the TC device, and an acceptable value when forwarded towards the slave device.

Physical Network Multi-Vendor MPLS & Ethernet

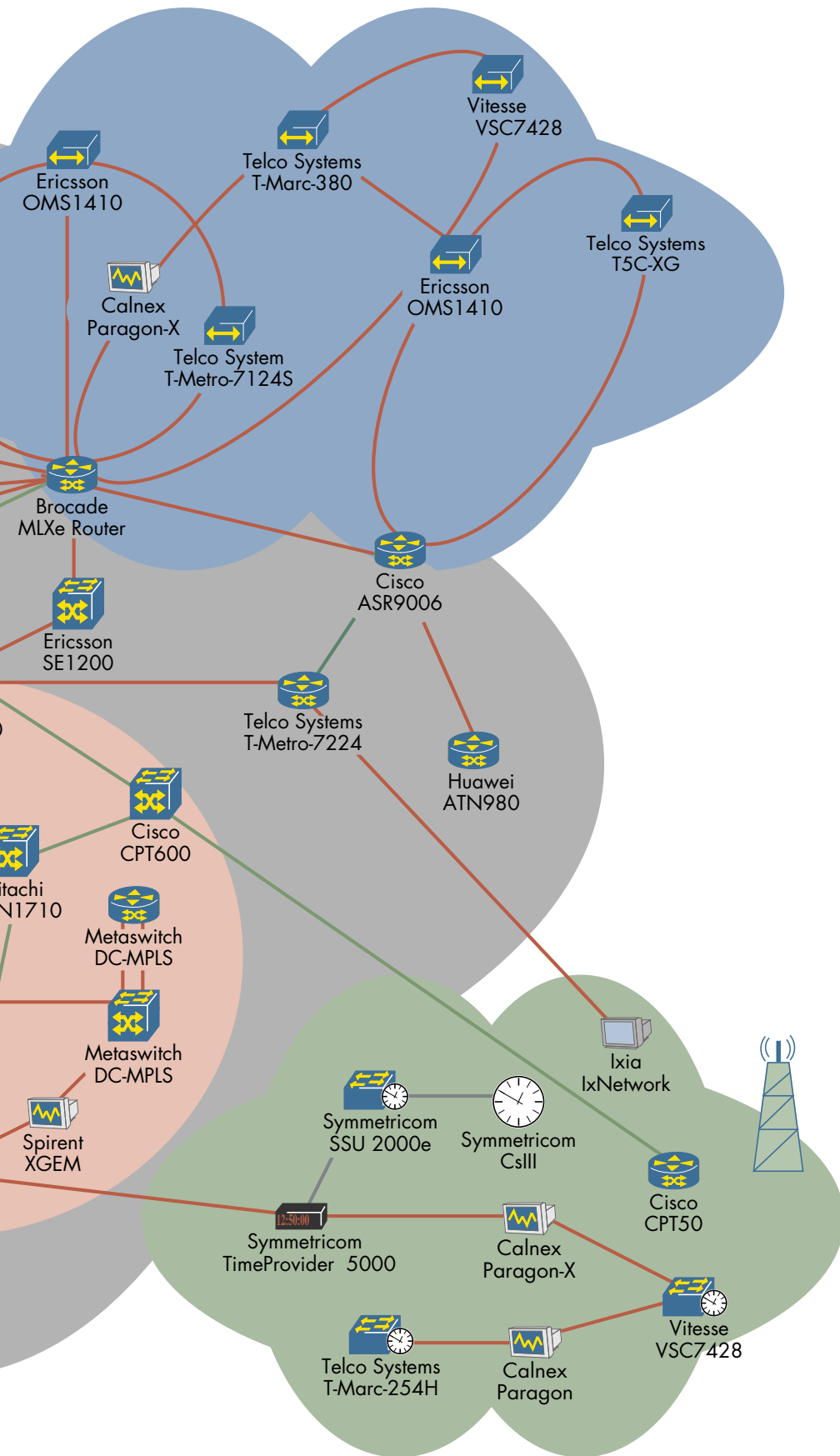
Management Systems













Cisco
Prime Network








Network Topology & Interoperability Event 2011









Device Types

-  Provider Edge device
-  Provider router
-  Optical Aggregation device
-  Access device
-  IEEE1588-2008 client
-  Analyzer/Emulator
-  Impairment tool
-  Reference Clock
-  Management system
-  IEEE1588-2008 Grandmaster

Connection Types

-  1 Gigabit Ethernet link
-  10 Gigabit Ethernet link
-  100 Gigabit Ethernet link
-  DWDM link
-  Clock link

Network Areas

-  IP/MPLS network
-  MPLS-TP network
-  Ethernet Ring Aggregation
-  Optical network
-  Radio access
-  Data Center

We performed the Transparent Clock test in two parts. First, the slave was allowed to synchronize to the Grandmaster without impairment. A Telco Systems T5C-XG was used to distribute the PTP frames to the different Ethernet interfaces of the TC devices. In the second part, a second Ethernet Virtual Connection (EVC) was created for data traffic. Traffic was transmitted downstream by Ixia IxNetwork on this EVC at 50% of line rate through the TC in the direction from the Grandmaster through an intermediate non-PTP switch and egressing from the slave.

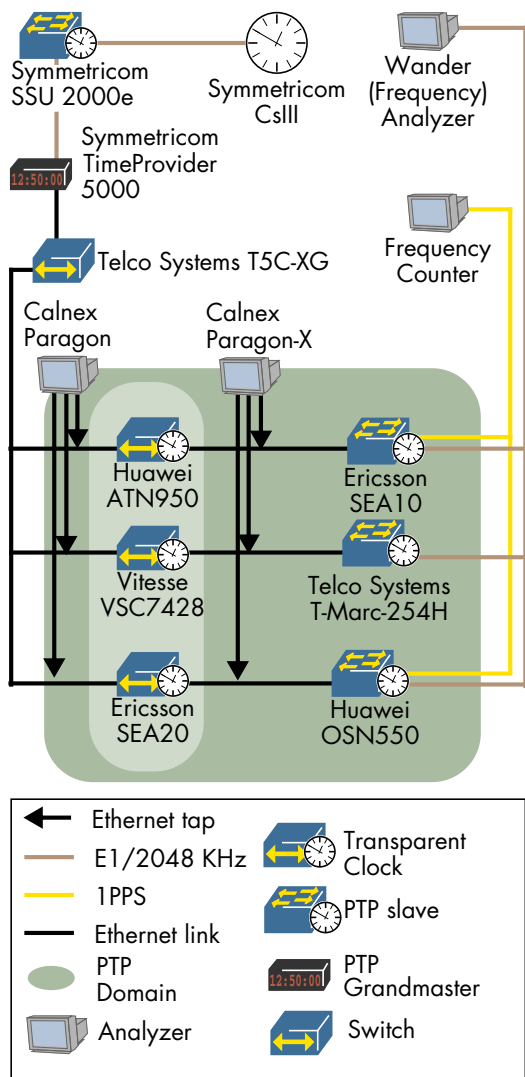


Figure 9: Transparent Clock Tests

In both parts, to ensure that the Correction Field offset had a positive impact on synchronization, the slave device was connected to a wander analyzer (frequency, 2048MHz or E1) and/or a frequency counter to measure phase (time of day) deviation.

Successful test pairs under these requirements included: Ericsson SEA20, Huawei ATN950 and Vitesse VSC7428 as Transparent Clocks with Huawei OSN550, Ericsson SEA10 and Telco Systems T-Marc-254H as slave clocks. In addition, following this end-to-end setup, Ixia IxNetwork was used in some cases to take a closer look at the accuracy of the Correction Field by connecting to the

TC, emulating the Grandmaster and the slave clock and comparing the Correction Field with the actual delay of the PTP frame. Accuracy was found to be within the 10 µs range.

IEEE 1588 Boundary Clocks

In addition to Transparent Clocks, the IEEE 1588-2008 standard also specifies another type of PTP-aware role in the network - the Boundary Clock. The idea is to break up a single PTP session into a chain of clocks - multiple masters and slaves rather than a single pairing between the Grandmaster and a slave across a large network. The test method here is a bit more straight forward than that of the Transparent Clock - the question is, can the slave-only Ordinary Clock still synchronize to the primary reference by locking to the intermediate PTP master sitting within the Boundary Clock, which in turn locks to the PTP Grandmaster.

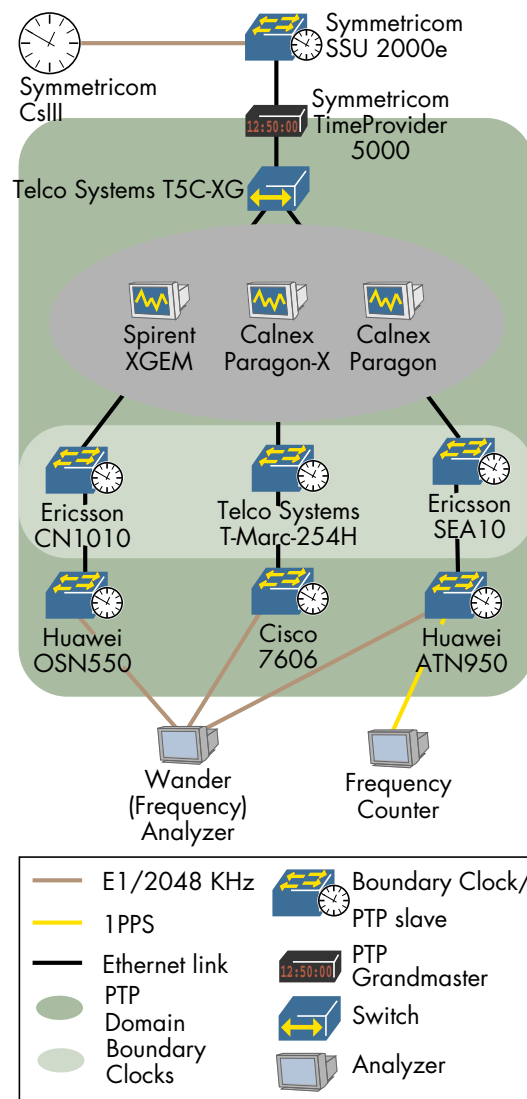


Figure 10: Boundary Clock Tests

Once again, we ran the wander analyzer and frequency counter for a minimum of four hours for a test to pass. In this case, vendors had to pass the ITU-T G.823 SEC MTIE mask for frequency synchroni-

zation and a maximum time error of 3 μs for the time of day deviation. Again, a Telco Systems T5C-XG was used to distribute the clock to the devices. The following devices successfully participated in the test as Boundary Clocks: Ericsson CN1010, Ericsson SEA10, and Telco Systems T-Marc254H. These Boundary Clocks were connected to the following Ordinary Clocks respectively in order: Huawei OSN550, Cisco 7606, and Huawei ATN950. All test pairs successfully tested frequency synchronization. The Ericsson SEA10 and the Huawei ATN950 test pair passed the phase requirements test as well.

Ethernet Synchronous Messaging Channel (ESMC)

When one mentions Synchronous Ethernet, this in fact simply refers to a physical Ethernet port which is capable of either providing or locking onto a high precision clock for frequency synchronization. ESMC has been designed to expand the capability of such interfaces with the ability to communicate the state of directly attached clocks, and the Quality Level (QL) of those clocks. This allows other SyncE devices to have redundant clocks to choose from, and to decide which of the redundant clocks to lock onto.

The test setup was comprised of three different SyncE nodes. Two of them, on opposite sides of the topology, were connected to external clock sources and allowed to advertise their Quality Level over their SyncE interfaces. One of these external clocks was connected to a Primary Reference Clock (PRC), and the other to a Synchronization Supply Unit (SSU). These two devices are denoted PRC and SSU for simplicity. The SyncE node in the middle of the topology was not connected to an external clock source, and advertised the quality level of its internal clock. This internal clock, if not advertised as DNU (Do Not Use), was expected to advertise SDH Equipment Clock (SEC) since many internal clocks comply to SEC. This device is therefore referred to as SEC.

We conducted the ESMC tests by connecting, disconnecting, and reconnecting the external clock sources to the PRC and SSU nodes. One device on each link was configured to mirror both transmit and receive traffic to an Ixia IxNetwork or Spirent TestCenter port to observe the SSM exchange.

The following devices participated successfully in the ESMC test scenarios as depicted by Figure 11: ESMC Tests on page 13: Cisco 7606, Cisco ASR9006, Huawei ATN980, Huawei ATN990, Huawei CX600-2, Telco Systems T-Marc-3208SH, and Telco Systems T-Metro-7124S-2. In addition, Ixia's IxNetwork and Spirent's TestCenter both demonstrated their ability to successfully transmit the required Synchronization Status Messages (SSMs) for all parts of this test case and have them recognized by the devices.

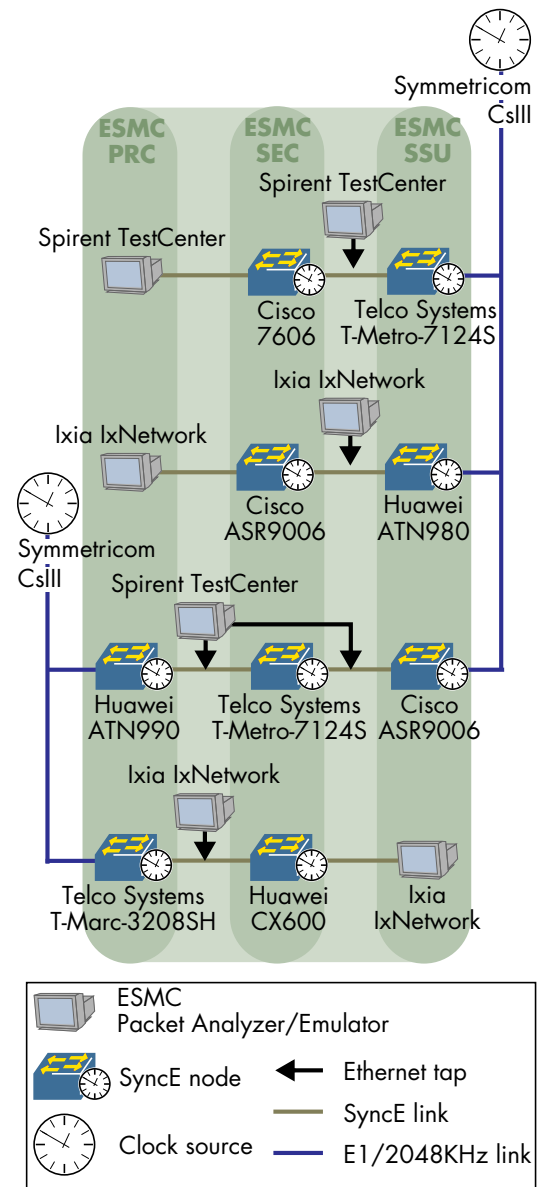


Figure 11: ESMC Tests

RESILIENCY TESTS

Throughout the forest of IEEE standards, ITU-T recommendations, and IETF RFCs, there is a wide breath of resiliency, redundancy, and protection mechanisms and protocols to cover virtually every part of the network. Our tests in this area focused on two more up-and-coming standards: MPLS-TP Protection, and ERPS.

Ethernet Ring Protection Switching (ERPS)

The second edition of ITU-T recommendation G.8032 was ratified and published just last year, and extends the protocol by adding more features like ring interconnection, where several rings are interconnected to extend their coverage, and administrative commands. Our tests focused on the advanced features introduced by the newer version of the standard.

The standard describes how a Ring Protection Link (RPL) "owner" is designated to block a port on the RPL in normal operation to guarantee a loop free topology. The ring also uses a "Service Channel" to transport data frames (customer traffic), and a Ring Automatic Protection Switching (R-APS) channel - a special in-band multicast channel used for the transmission of R-APS messages. R-APS is the protocol defined by the G.8032 standard to coordinate events and ring state across all Ethernet Ring Protection (ERP) nodes.

Regardless of the topology we conducted all tests by first sending traffic to verify the proper operation of the ring. We then used a Calnex Paragon to impair all traffic in one direction between two ring nodes and verified that the RPL owner removed the blocked port. Restoration was tested by disabling impairment, and repeating the test on the same link, in the opposite direction. In some cases however we resorted to pulling a cable to trigger the failover.

To start off, we conducted a successful test of a single ring topology with the following devices: Cisco ASR9006, Ericsson OMS-1410, and Telco Systems T5C-XG. Cisco ASR 9006 was the RPL Owner. R-APS and Service Channels were configured over VLAN interfaces. Connectivity Fault Management (CFM) was used to monitor the link liveness between nodes, running Continuity Check Messages at an interval of 100 ms. Proper failover and recovery was successfully tested. We recorded an asymmetric failover time, in that the direction where the traffic was dropped took far less time to failover than the other direction.

In all following scenarios two rings (a major ring and a sub ring) were interconnected via a shared link. To perform these tests we defined two parameters profiles to accommodate differences in vendors' implementations. The first profile, "Profile 1", made use of R-APS VIDs for the disambiguation of the R-APS channel in both rings. This profile was also specified for the ring interconnection model using a Virtual Channel to enable interconnection nodes to exchange R-APS messages for the Sub-ring. The second profile, "Profile 2", was intended for all implementations supporting [RingID] as the last octet of the MAC addresses used for the G.8032 R-APS messages communication thus using the same R-APS VID in both rings. This profile did not specify the use of a "Virtual Channel". In each ring one node was provisioned as RPL owner. In some tests the RPL owner neighbor in the major ring was configured to block the port on the RPL link.

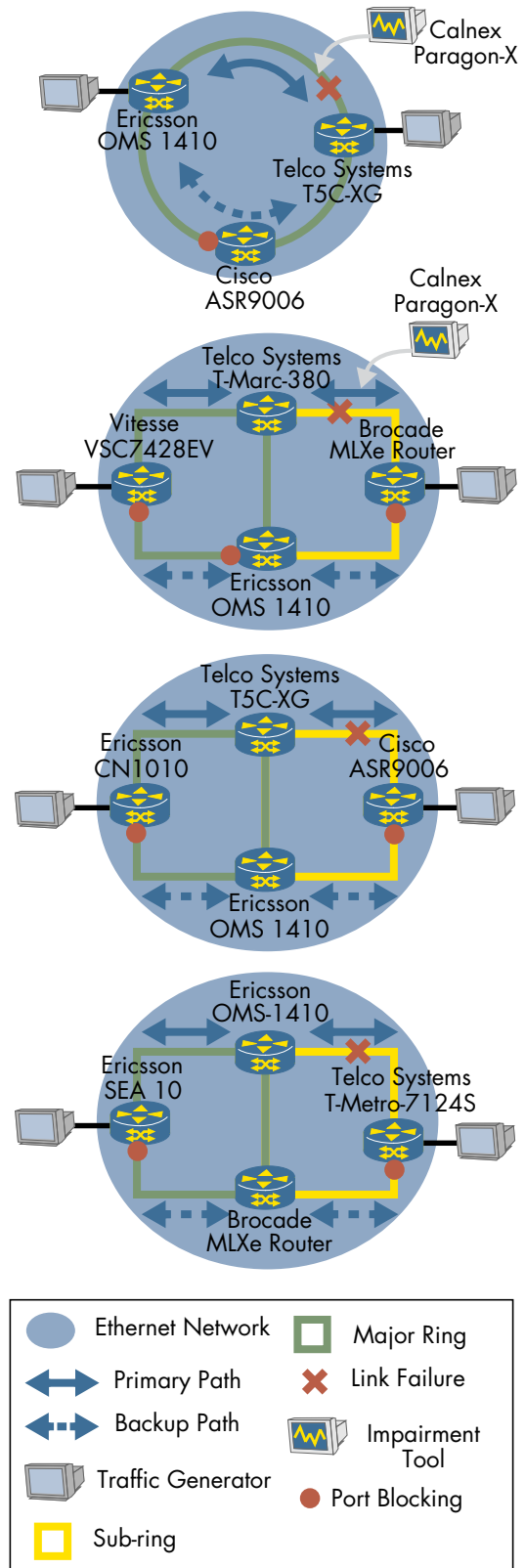


Figure 12: ERPS Tests

As we started the test according to "Profile 1" we noticed that vendors take two different approaches to implement the virtual channel: One using the same R-APS VID as in the sub-ring and the other using a different RAPS VID from the sub-ring. As a consequence there was no common method between those vendors who implemented virtual channel, so they configured the setup without the use

of a Virtual Channel.

We successfully performed two tests using "Profile 1". The first consisted of Brocade MLXe Router, Ericsson OMS-1410, Telco Systems T-Marc-380 and Vitesse VSC7428E where a CCM interval of 3.33 milliseconds was used. The second consisted of Cisco ASR 9006, Ericsson CN1010, Ericsson OMS 1410, and Telco Systems T5C-XG where a CCM interval of 100 ms was used.

The "Profile 2" run was successfully performed on the Brocade MLXe Router, Ericsson OMS 1410, Ericsson SEA 10, and Telco Systems T-Metro-7412. After emulating the failure in the sub-ring via LoS, we successfully verified that the protection switching unblocked the RPL port of the sub-ring.

In all test scenarios we successfully tested administrative commands including "Manual" and "Forced Switch", which move the blocked port as specified. The failover time ranged from 14 to 70 milliseconds. Both commands were also successfully removed by "Clear" command, resulting in a restoration time between 14 and 25 milliseconds. All test scenarios showed an asymmetric failover time ranging from 17 to 372 milliseconds. We also recorded a restoration time from no loss at all, to 108 milliseconds.

While we have seen tremendous progress in ERPS implementations, we also observed some interesting differences amongst them. In one case when a single direction was impaired by the impairment tool, the respective device properly sent an ETH-RDI to its peer after experiencing a loss of three CCMs. Its peer, in this case, did not then trigger failover and continued to use that link. Also, because of different implementation of the ETHDi/ETH_A function that 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. This resulted in even more issues. Some vendors were able to configure only one maintenance association (MA) per VLAN. Other vendors were not able to configure more than two MEPs per MA. Because of this it took a long time for vendors to find out which configuration to use in order to interoperate with other vendors. We also observed that some vendors could not run CCM on two interfaces simultaneously. These issues were resolved by disabling CCMs on one interface.

MPLS-TP LSP Protection

RFCs 5654 and 5921 discuss the need for a packet-aware transport technology with the reliability and simplicity that operators are used to from SONET/SDH, and thus respectively define the requirements and framework for an MPLS Transport Profile (MPLS-TP) to achieve just that. OAM tools, being one of the key aspects to accomplishing this goal, has been a main focus for standardization. In this test, we verified the ability of such tools to detect failure and of the network to react accordingly.

To perform the tests two separate Label Switched Paths (LSPs) - primary and backup LSP - were statically configured between the participating devices.

In some cases ARP was used to obtain the MAC address of the next hop peer, in other cases MAC addresses were statically configured. A single bidirectional BFD session was then configured on each LSP to run over Generic Associated Channel (G-ACh). BFD was running at 100 milliseconds (ms) intervals to monitor the liveness of each LSP.



Figure 13: MPLS-TP LSP Protection Tests

Once both LSPs were up we ran a baseline test by sending traffic and verifying that traffic was indeed forwarding only over the primary path. To verify that a loss of BFD messages would cause the edge nodes to switch over to the backup LSP, a Calnex Paragon impairment device was used to unidirectionally drop all traffic on the working path, thus emulating a unidirectional failure within a transmission network. Physical Ethernet link on all devices remained up.

The following devices successfully took part in the test: Cisco ASR9010, Cisco 7604, Ericsson SE1200, and Metaswitch DC-MPLS. In all test scenarios the traffic reverted to the backup path after the impairment took place. The failover time ranged from 270 to 310 ms. Additionally, upon removing the impairment, traffic successfully reverted to the primary LSP. The observed restoration time was no

more than 60 milliseconds.

In some additional test scenarios the restoration of the LSP was triggered manually because of a difference in G-ACh/GAL implementations when using BFD to protect an LSP or Pseudowire. Some vendors' implementations required that both LSPs and Pseudowires are protected, thus running BFD on both LSP and PW, whereas some other vendors could only protect at the LSP level - even though these LSPs indeed transported PWs. Ixia setup two LSPs on a single physical connection to a Huawei CX600-X3 which was redundantly connected to an Ericsson SE1200 using two disjoint LSPs, as shown in the diagram.

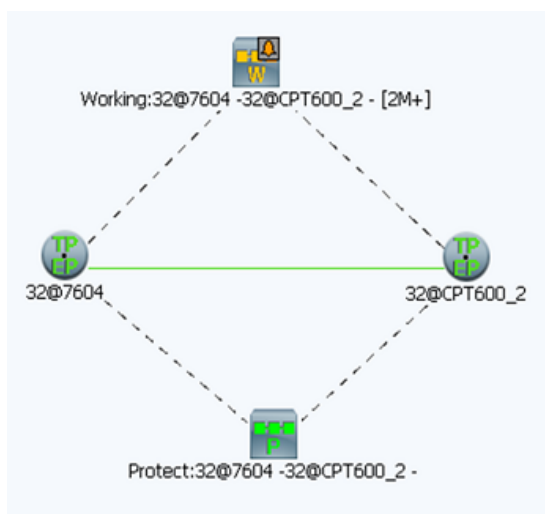


Figure 14: Cisco Prime Network Screenshot

An additional MPLS-TP Protection scenario was demonstrated between a Cisco CPT600 and 7600, with an Ericsson SE1200 as a mid-point. Cisco Prime Network was able to automatically discover the physical topology which included both Cisco and Ericsson devices and the MPLS-TP LSP. Additionally, when the working LSP was brought down through command initiated by the Prime Network, the management system reported the primary LSP as down by displaying an alarm and highlighting the backup path as being the newly active one. The switch was also verified by executing MPLS Pings by management system.

ACKNOWLEDGEMENTS

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

Editors. This report has been edited by Kari Asheim, Terje Krogdahl, Jonathan Morin, Stephen Murphy, Ronsard Pene and Xiao Tai Yu.

ACRONYMS

Term	Definition
AC	Access Circuit
APS	Automatic Protection Switching
AS	Autonomous System
ASN	Autonomous System Number
BC	Boundary Clock
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
C-VLAN	Customer VLAN
CC	Continuity Check
CCM	Continuity Check Message
CE	Customer Edge
CFM	Connectivity Fault Management
CoS	Class of Service
CV	Connectivity Verification
CVID	Customer VLAN ID
DMM	Delay Measurement Message
DMR	Delay Measurement Reply
DNU	Do Not Use
DSLAM	Digital Subscriber Line Access Multiplexer
DWDM	Dense Wave Division Multiplexing
eBGP	External BGP
EEC	Ethernet Equipment Slave Clock
EFM	Ethernet in the First Mile
ENNI	External Network-Network Interface
ERPS	Ethernet Ring Protection Switching
ESMC	Ethernet Synchronization Messaging Channel
EVC	Ethernet Virtual Connection
FDB	Filtering DataBase
G-ACh	Generic Associated Channel
GAL	G-ACh Label
GMPLS	Generalized MPLS
IANA	Internet Assigned Numbers Authority
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IP	Internet Protocol
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
LAN	Local Area Network
LDP	Label Distribution Protocol
LMM	Loss Measurement Message
LMR	Loss Measurement Reply
LOS	Loss of Signal

Term	Definition
LSP	Label Switched Path
MA	Maintenance Association
MD	Maintenance Domain
MED	MULTI_EXIT_DISC
MEP	Maintenance Entity Point
MIP	Maintenance Association Intermediate Point
MPLS	Multi-protocol Label Switching
MPLS-TP	MPLS Transport Profile
MTIE	Maximum Time Interval Error
OAM	Operations, Administration and Management
OC	Ordinary Clock
OSPF	Open Shortest Path First
P	Provider device
PE	Provider Edge device
PRC	Primary Reference Clock
PRS	Primary Reference Source
PSN	Packet Switched Networks
PTP	Precision Time Protocol
PW	Pseudowire
QL	Quality Level
R-APS	Ring Automatic Protection Switching
R-APS(NR)	ring automatic protection switching no request
R-APS(SF)	ring automatic protection switching signal failure
RPL	Ring Protection Link
RSVP-TE	Resource reSerVation Protocol Traffic Engineering
S-VLAN	Service VLAN
SEC	SONET/SDH Equipment Clock
SLA	Service Level Agreement
SSM	Synchronization Status Messages
SSU	Synchronization Supply Unit
SVID	Service VLAN ID
SyncE	Synchronous Ethernet
T-LDP	Targeted LDP
TC	Transparent Clock
TDEV	Time DEVIation
TIE	Time Interval Error
TLV	Type Length Value
TOD	Time of Day
TPID	Tag Protocol ID
UNI	User Network Interface
VCCV	Virtual Circuit Connectivity Verification
VLAN	Virtual LAN
WTR	Wait to Restore timer

REFERENCES

- Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description, RFC3471
- "Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)", RFC 3813
- "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812
- "Multiprotocol Label Switching (MPLS) Forwarding Equivalence Class To Next Hop Label Forwarding Entry (FEC-To-NHLFE) Management Information Base (MIB)", RFC 3814
- Generalized Multi-Protocol Label Switching (GMPLS) Architecture, RFC3945
- "Traffic Engineering Link Management Information Base", RFC 4220
- "BGP Extended Communities Attribute", RFC4360
- "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC4364
- "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447
- "BGP Support for Four-octet AS Number Space", RFC4893
- "MPLS Generic Associated Channel", RFC5586
- "MPLS Transport Profile Data Plane Architecture", RFC5960
- "Proactive Connection Verification, Continuity Check and Remote Defect indication for MPLS Transport Profile", draft-ietf-mpls-tp-cc-cv-rdi, work in progress
- "MPLS On-demand Connectivity Verification and Route Tracing", draft-ietf-mpls-tp-on-demand-cv, work in progress
- "Using the Generic Associated Channel Label for Pseudowire in MPLS-TP", draft-ietf-pwe3-mpls-tp-gal-in-pw, work in progress
- "MPLS Transport Profile Data Plane Architecture", draft-ietf-mpls-tp-data-plane, work in progress
- "A Framework for MPLS in Transport Networks", draft-ietf-mpls-tp-framework, work in progress
- "MPLS Fault Management OAM", draft-ietf-mpls-tp-fault, work in progress
- "Pseudowire Status for Static Pseudowires", draft-ietf-pwe3-static-pw-status, work in progress
- "Ethernet Pseudowire (PW) Management Information Base (MIB)", draft-ietf-pwe3-enet-mib, work in progress
- "Media Access Control Parameters, Physical Layers and Management Parameters for 40 Gb/s and 100 Gb/s Operation", IEEE P802.3ba 2010
- "Precision Time Protocol (PTP)", IEEE 1588-2008
- "Spectral grids for WDM applications: DWDM frequency grid", ITU-T G.694.1
- "Interfaces for the Optical Transport Network (OTN)", ITU-T G.709
- "Synchronization layer functions", ITU-T G.781
- "OAM Functions and Mechanisms for Ethernet Based Networks", ITU-T Y.1731
- "Ethernet ring protection switching", ITU-T G.8032
- "Timing and Synchronization Aspects in Packet Networks", ITU-T G.8261/Y.1361
- "Precision Time Protocol Telecom Profile for frequency synchronization", ITU-T G.8265.1
- "Distribution of timing information through packet networks", ITU-T G.8264
- "Mobile Backhaul Implementation Agreement Phase 2", MEF technical specification, work in progress



EANTC AG
European Advanced Networking Test Center

Einsteinufer 17
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>

This report is copyright © 2011 EANTC AG. While every reasonable effort has been made to ensure accuracy and completeness of this publication, the authors assume no responsibility for the use of any information contained herein.

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

20110131 v1.0

