EANTC

Interoperability & Feasibility Showcase 2015

White Paper

MPLS SDN IP WORLD CONGRESS

NFV & SDN SUMMIT 2015

V6World CONGRESS 2015

Editor's Note

Carsten Rossenhövel
Managing Director, EANTC

Welcome to the 2015 edition of the EANTC showcase! This year's white paper deals with multi-vendor testing of SDN and NFV topics, and there are leap-frog advancements which we are very proud and happy to share.

But the core theme of our interoperability test has in fact been *evolution*, or more precisely: How existing transport networks can be smoothly upgraded using a range of technologies such as Segment Routing, PCE, OpenFlow, and orchestrated across multi-vendor scenarios.

19 leading Western and Eastern vendors joined our test for which the EANTC team created a test plan resembling a playground: We offered test areas such as next-generation core, aggregation, access, data center transport; controller and orchestrator interaction in software-defined networks; and virtualized network functions running across the whole scenario. It was up to the vendors to select areas and implement their individual (yet standards- or open source-based) ideas — much more than in the past. Consequently, we decided to change the name of the showcase to Interoperability and *Feasibility* Event.

Some network equipment manufacturers were still a bit shy to join the testing in this new wild world: We are no longer in school where pencils are taken out, and standard blueprints carefully copied. This is the sandbox where molds are grabbed and sand castles are built together. The most innovative ideas will come from teams with open minds and open source tools. EANTC's job is now to keep the playground organized in all its chaotic glory and to continue encourage multi-vendor interoperability as the industry transitions.

For the first time, we conducted testing at two sites in parallel, Berlin and Beijing. Facilitated through EU-China FIRE research project, OpenFlow tests were carried out across the two sites successfully, including a number of first-time Chinese vendors. In an open-source future of software-defined networks, white-labeled products will likely be more important.

Meanwhile, the packet clock synchronization world faces massive growth in mobile networks — more (small) cell sites, more bandwidth, a wider range of application scenarios, and LTE-Advanced being deployed. Vendors of grandmaster, boundary and slave clock solutions demonstrated readiness; they tested very advanced phase sync concepts for multi-vendor interoperability at the EANTC lab, including scale to confirm design feasibility for very large networks.

Our team and I hope that this year's white paper — with 24 pages one of our most extensive interoperability event documentations so far — will provide useful insights. We are definitely open for your comments, feedback and ideas for the next event. Please get in touch!

Table of Contents

Participants and Devices 3
MPLS, Ethernet & Data Center Interconnect..... 4
Software Defined Networking..... 9
Topology..... 12
Clock Synchronization 15
Network Functions Virtualization Feasibility Test 21
Demonstration Scenarios..... 22

Introduction

EANTC preparation for this interoperability event started much earlier than in previous years. We invited vendors who repeatedly supported the showcase in the past to a strategy call already in June 2014. We outlined our vision and asked for feedback and a reality check. The support and commitment we received on that call was great. The participants of the call, including Alcatel-Lucent, Cisco, Ericsson, Juniper Networks, confirmed that this year the agenda is testable and is very much in line with the interest they received from their service provider customers.

Technical discussions ensued and over the next months we discussed the details with interested vendors. This year we opened the interoperability program to vendors who could not easily travel to our lab in Berlin. This part of the test was co-organized by EU-China FIRE project (ECIAO) and hosted by Beijing University of Posts and Telecommunications (BUPT)'s Network Information Centre. Functional OpenFlow testing across vendors in the two locations was carried out over a VPN tunnel.

Looking back at the fantastic results we collected in the two weeks hot staging, we can split the tests to three distinct areas.

Service Provider SDN. Alongside OpenFlow tests, an area we have been exploring for the third time now, we also welcomed several implementations of Path Computation Element Communication Protocol (PCEP) as well as Netconf with Yang models. Both protocols empower service providers to use software tools to define their network services – exactly the point behind SDN.

Significant Updates to Core Networking. As a lab that has been testing IP/MPLS since the start of the century, it was great to see so many updates supported by our vendor customers to core networking. We verified Ethernet VPNs interoperability between major vendors, tested Segment Routing for the first time in an EANTC event, and still had the pleasure of providing new RSVP-TE implementations with an Interoperability platform.

Advanced Clock Synchronization. Our interoperability showcases are the most well attended multi-vendor packet clock synchronization platform in the industry. As such we are always challenged by the participants to expend and extend the test scope. We saw an increase in GrandMaster functions (three implementations attended). We also tested more complicated conditions such as path asymmetry and path rearrangement for packet clocks.

As usual, this white paper documents positive results (passed test combinations) individually with vendor and device names. Failed test combinations are not mentioned in diagrams; they are referenced anonymously to describe the state of the industry. Our experience shows that participating vendors quickly proceed to resolve interoperability issues after our test so there is no point to punish vendors. Confidentiality is a main cornerstone to encourage manufacturers to participate with their latest (beta) solutions.

Participants and Devices

Vendor	Devices
Albis Technologies	ACCEED 2104
Alcatel-Lucent	7750 SR-7 Nuage 7850 VSG
BTI Systems	BTI 7800
Calnex	Paragon-X
Cisco	ASR 9001 UCS-C220 M3 WAN Automation Engine (WAE) Nexus 7004 Nexus 9396
Digital China Networks (DCN)	DCRS-7604E
Ericsson	MINI-LINK PT 2020 MINI-LINK TN OpenDaylight (ODL) SP 415 SP 420 SSR 8004 SSR 8010 Virtual Router (EVR)
Greenet	GNFlush
Huawei	VNE 1000
Ixia	Anue 3500 IxNetwork NetOptics xStream RackSim
Juniper Networks	MX80
Meinberg	LANTIME M1000 LANTIME M4000
Microsemi	TimeProvider 2300 TimeProvider 2700 TimeProvider 5000
Open Source	OpenDaylight Helium
RAD	ETX-2 MiCLK MiNID RADview Management with D-NFV Orchestrator
Ruijie	Ruijie-ONC
Spirent Communications	TestCenter TestCenter Virtual
Tail-f Systems	NETCONF-Console
ZTE	ZXCTN9000-2E10 ZXCTN9000-3E ZXCTN9000-8E

Interoperability Test Results

The following sections of the white paper describe each of the test areas and results.

Terminology. We use the term *tested* when reporting on multi-vendor interoperability tests. The term *demonstrated* refers to scenarios where a service or protocol was evaluated with equipment from a single vendor only. In any case, demonstrations were permitted only when the topic had been covered in the previously agreed test plan; sometimes vendors ended up with demonstrations because there was no partner to test with, or because multi-vendor combinations failed so that they could not be reported.

Test Equipment. With the help of participating test equipment vendors, we generated and measured Ethernet and MPLS traffic, emulated and analyzed control and management protocols and performed clock synchronization analysis. We thank Calnex Solutions, Ixia, and Spirent Communications for their test equipment and support throughout the hot-staging.

MPLS, Ethernet & Data Center Interconnect

MPLS has been widely adopted by many service providers across the world. Layer-2 multi-point services have often been implemented using Virtual Private LAN Services (VPLS). As customer applications evolves and as the adoption of cloud-based and virtualized infrastructure increases, the task of interconnecting multiple data centers, potentially with large number of devices (virtual and physical), could be facing scalability and efficiency issues. To provide another, perhaps more efficient layer-2 mutli-point connectivity solution, the IETF has been working on defining Ethernet VPNs (EVPN). We tested both EVPN and PBB-EVPN in various setups.

We also tested, for the first time, Segment Routing as another core-backbone emerging technology. Segment Routing extends the Interior Gateway Protocol (IGP) to provide an alternative to RSVP-TE.

From a network convergence and recovery perspectives, we covered multiple Fast Reroute (FRR) scenarios in both Segment Routing and RSVP-TE spaces. We also tested Segment Routing and LDP interworking as well as BGP Prefix Independent Convergence (BGP PIC).

Ethernet VPNs

EVPN is regarded by many as the next-generation VPN solution. It uses Multi-Protocol Border Gateway Protocol (MP-BGP) as a control plane for MAC and IP address learning/advertisement over an IP core. EVPN combined with Provider Backbone Bridging (PBB-EVPN) provides mechanism to reduce the number of MAC advertisements via aggregation. Both EVPN and PBB-EVPN are currently under standardization by IETF BGP Enabled Services (bess) Working Group. Both offer separation between the data plane and control plane, which allows the use of different encapsulation mechanisms in the data plane. Both MPLS and Network Virtualization Overlay (NVO), an example of which is Virtual Extensible LAN (VXLAN), are defined as data plane options. To enable interoperability between EVPN/PBB-EVPN implementations, a new BGP Network Layer reachability Information (NLRI) has been defined.

The EVPN/PBB-EVPN specifications introduce different route types and communities to achieve the following functions: MAC address reachability and withdrawal, split-horizon filtering, aliasing, endpoint discovery, redundancy group discovery and designated forwarder election. As interoperability has always been a key condition for any successful multi-vendor deployment, we included many of these functions in the tests.

Ethernet VPN: Single-Homing. While in previous Interoperability tests we used MPLS data plane to test multipoint services, in this year's event, vendors were interested in using VXLAN as data plane for EVPN in the single-homing scenario.

We built a multi-vendor test setup, where five PE nodes were connected to a Route Reflector (RR). As the first step for this test, vendors interconnected all participating PEs using overlay VXLAN tunnels. They then configured a common EVPN instance on each PE. In this setup, Ixia IxNetwork was used to emulate Customer Edges (CEs) devices, each of which was attached to a single Provider Edge (PE) in a single-homing setup. Each CE was configured with three VLANs, with each VLAN mapped to each of the VXLAN Network Identifier (VNI) associated with the EVPN instance. In the core network, vendors agreed to use OSPF as the IGP protocol.

Once OSPF was up and running in the core network, we enabled MP-BGP between all PEs and route reflector (RR). We first verified that BGP EVPN NLRI was properly negotiated between all BGP peers. The next step was to assure that each EVPN PE node received Inclusive Multicast Ethernet Tag routes (BGP route type 3) from all other PEs. We then started generating traffic between all emulated CEs, and verified that EVPN PEs learned the Customer MAC (C-MAC) on the local segment in the data plane according to the normal bridging operation. Furthermore we checked that the previously learned MAC addresses were received on the remote EVPN PE through BGP NLRI using BGP MAC Advertisement route (BGP route type 2). In the last step of this extensive test, we generated bidirectional known traffic between all CEs using Ixia IxNetwork. We did not observe traffic loss for the configured services.

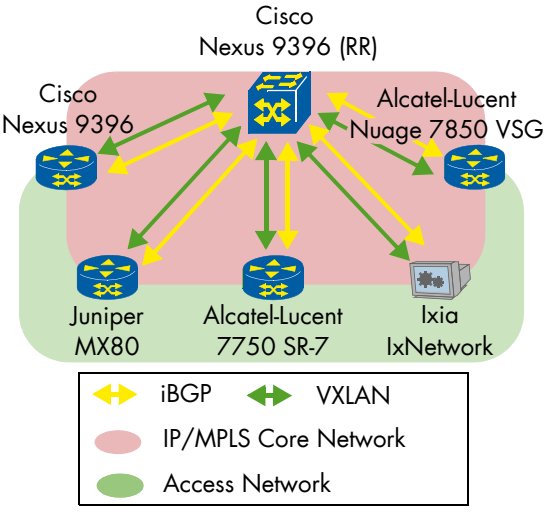

Figure 1: EVPN: Single-Homed

Five vendors successfully participated in the test as EVPN PE: Cisco Nexus 9396, Juniper MX80, Alcatel-Lucent 7750 SR-7, Alcatel-Lucent Nuage 7850 VSG and Ixia IxNetwork. An additional Cisco Nexus 9396 device functioned as the Route Reflector.

While sending multicast traffic, we observed an inconsistent behavior among vendors. Some implementations correctly replicated multicast traffic. We also observed that other implementations, although exchanging the Inclusive Multicast Ethernet Tag Route correctly, did not forward multicast traffic to the other EVPN PEs.

During the preparation phase, vendors recognized that they have different versions of the EVPN Overlay draft. Some vendors required the VXLAN

Network Identifier (VNI), a 24-bit identifier inside the VXLAN frame header used to designate individual connection, to be encoded in the Ethernet Tag. Other vendors used MPLS label to encode the VNI value. One vendor volunteered to modify their implementation before the hot staging. With this modification, we achieved interworking between all participants.

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

There are two broad approaches for IRB described in the draft document: Asymmetric and Symmetric. As their names indicate, these modes have different congruency behaviors for bidirectional flows as well as different host's MAC/IP learning requirements on VTEP switches. In the Asymmetric IRB scenario, both Layer2 and Layer3 lookup associated with the inter–subnet forwarding are performed in the ingress PE, whereas the egress PE performs Layer2 lookup only. In the Symmetric IRB scenario both ingress and egress PEs perform Layer2 and Layer3 lookup. Since both forwarding modes are not interoperable, we created two setups for the tests.

In our first EVPN inter-subnet forwarding setup, we tested symmetric mode. We built a spine-leaf topology, where the two leaves were interconnected over a VXLAN overlay tunnel through a spine node, acting as Route Reflector (RR). On the emulated Tenant System (TS) to leaf segment, two VLAN services (bridge domain) were provisioned and were attached with the common EVPN configured on each leaf. Each bridge domain was configured with an IRB interface. We configured an IP Virtual Route Forwarding (IP-VRF) instance on each leaf node and assigned the EVPN to it.

In our second inter-subnet forwarding setup, we tested asymmetric mode. The setup interconnected two EVPN PE nodes over a VXLAN overlay tunnel. Each EVPN PE was configured with two EVPN instances, these in turn were assigned to a single IP-VRF. Two Tenant Systems (TS) were connected to each EVPN PE. Each TS was associated with a single EVPN. The two TSes connected to each EVPN PE were considered part of the same IP subnet.

In both setups, we enabled BGP sessions between leaves and RR and verified that BGP EVPN NLRI was properly negotiated between all BGP peers. We then successfully verified intra-subnet forwarding by sending traffic between two endpoints in the same subnet. During the inter-subnet forwarding test we generated traffic between different subnets attached using Ixia IxNetwork, emulating the TSs. We first verified the exchange of BGP MAC/IP advertisement (BGP route type 2) carrying the relevant parameters. We observed no traffic loss.

In the symmetric forwarding mode, we successfully verified that traffic between the two subnets located on both leaves was forwarded in a symmetric way through the VNI dedicated to VXLAN routing for both ingress and egress traffic flows. Afterwards we emulated IP-subnets behind the TS attached to each of the leaf node. We first sent traffic between IP-subnets, then between the TS and the IP-subnet, emulating a common scenario, where the TS requires connectivity to external prefixes. We observed the exchange of BGP IP Prefixes (route type 5) carrying subnet prefixes with the IP of the TS as the next hop and other relevant parameters. We also validated the exchange of MAC/IP advertisement route (BGP route type 2) along with the associated Route Targets and Extended Communities.

In the asymmetric forwarding mode test, we successfully verified that traffic between the two subnets located at each leaf node was forwarded, where packets bypassed the IP-VRF processing on the egress node.

We also tested PIM-SM to implement multicast in the first setup and multicast ingress replication in the second setup. In both setups, when we sent multicast traffic between TSes, we observed no packet loss.

The following devices successfully participated in the symmetric forwarding mode demonstration: Cisco Nexus 9396 acting as leaf. Cisco Nexus 7004 as both spine node and route reflector.

The following devices successfully participated in the asymmetric forwarding mode test: Alcatel-Lucent 7750 SR-7 and Juniper MX80 as EVPN PE node. In both scenarios Ixia IxNetwork was used to emulate the TS, the IP prefixes as well as to generate traffic.

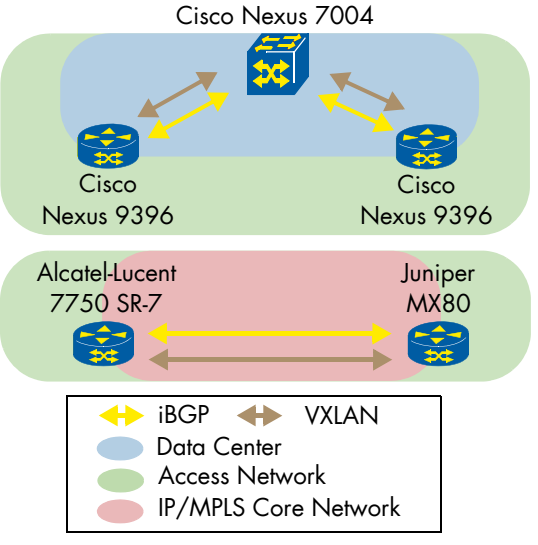

Figure 2: EVPN Inter-Subnet Forwarding

PBB-EVPN: Single-Homing. PBB-EVPN combines EVPN with Provider Backbone Bridging (PBB) MAC summarization. PBB-EVPN aggregates multiple Client MAC addresses (C-MACs) on the PE aggregation ports into a single Backbone MAC address (B-MAC). Only B-MAC addresses get advertised over the core.

Our setup interconnected three PBB-EVPN PE nodes over an MPLS data plane. For the MPLS control plane we used LDP as a signaling protocol and OSPF as Interior Gateway (IGP) protocol. We configured each PE with a single common PBB-EVPN instance. On each PBB-EVPN PE, we manually configured a single B-MAC for the attached CE. Ixia IxNetwork was used to emulate the CEs. On each CE-PE segment we configured a single CE-VLAN ID, that was mapped to a Provider Backbone Bridges Service Instance Identifier (PBB I-SID) associated with the PBB-EVPN instance at PE node.

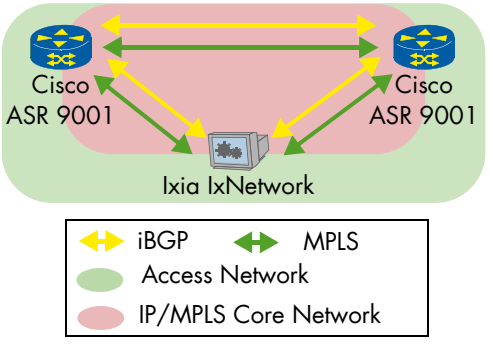

Figure 3: PBB-EVPN: Single-Homing

The vendors enabled MP-BGP on all PBB-EVN PE nodes and we verified that BGP EVPN NLRI was properly negotiated between all BGP peers. We also confirmed that each PE node advertised and installed multicast labels via Inclusive Multicast routes (BGP route type 3). We also confirmed that the B-MAC was successfully exchanged using the BGP MAC Advertisement routes (BGP route type 2). We used a traffic generator to send test traffic between all emulated CEs, and made sure that the C-MAC and B-MAC binding were correct. We successfully transmitted multicast traffic using Ingress replication without traffic loss.

Two Cisco ASR 9001 as well as Ixia IxNetwork successfully participated in the test as PBB-EVPN PE devices.

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

In this setup one of the Cisco ASR 9001 was configured as CE device and was dual-homed to two PEs — two Cisco ASR 9001 — using Link Aggregation Control Protocol (LACP). Ixia IxNetwork was used to generate traffic. Both Cisco ASR 9001 were configured to act as PBB-EVPN PEs, and to operate in all-active redundancy mode. One of the Cisco ASR 9001 had a single connection to the remote Ixia IxNetwork, acting as PBB-EVPN Provider Edge node. We tested two methods to identify customer sites: manual assignment of the B-MAC and a second method of auto-derivation of Ethernet Segment Identifier (ESI) using LACP.

Once OSPF and LDP sessions were established in the core, we enabled full-meshed MP-BGP sessions between PBB-EVPN PEs. Then we verified that all MP-BGP sessions were established and that BGP peers exchanged the correct BGP EVPN LNLRIs. Additionally, we verified that both Cisco ASR 9001 devices, acting as PBB-EVPN PEs auto-discovered each other, by advertising and importing the attached Ethernet segment using Ethernet Segment route (BGP route type 4), with the ES-Import Extended Community value. Afterwards, we verified that one Cisco ASR 9001 device was elected as the Designated Forwarder (DF) for the provisioned service on the multi-homed segment. Next, we first validated that each PE received and installed the multicast label via Inclusive Multicast routes (BGP route type 3). We then verified the exchange of BGP MAC Advertisement route (BGP route type 2) carrying the B-MAC. When we sent BUM traffic using the traffic generator connected to the Cisco ASR9001, which act as CE, we successfully tested that the Layer 2 Broadcast, Unknown Unicast and Multicast (BUM) traffic was not forwarded back to the origin Ethernet segment, and the split-horizon filtering was done based on B-MAC match. We observed no loss. We also sent unicast traffic between the CEs attached to the PEs. We verified that traffic originated from traffic generator attached to the dual-homed CE reached the remote PE over both attached PEs (with equal distribution). We also verified aliasing (i.e. load-balancing from one PE to multiple remote PEs) by configuring Ixia IxNetwork on the remote PE with manual flows using the EVPN unicast MPLS label of each the ASR 9001 dual homed PEs.

We also tested failure and recovery scenario by reconnecting the link between the CE and the designated forwarder. We validated MAC and Ethernet segment withdrawal using BGP MAC Advertisement (BGP route type 2) in the case of the link failure.

In a recovery scenario, we tested Ethernet Segment route (BGP route type 4) re-advertisement.

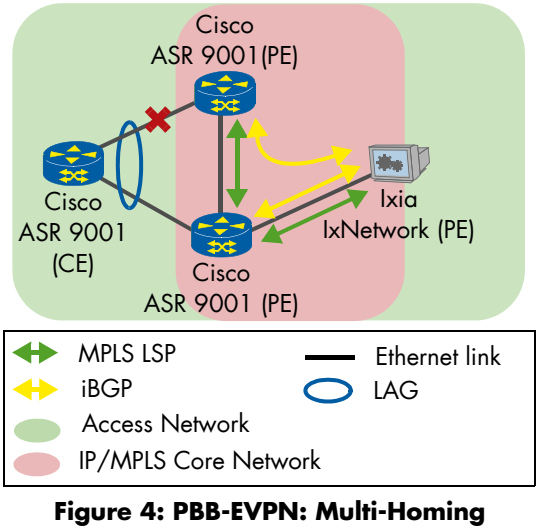

Figure 4: PBB-EVPN: Multi-Homing

Segment Routing

Segment Routing (SR) is beginning to emerge as an approach that could provide flexible forwarding behaviors with minimal network state. In Segment Routing, paths are encoded in the packet itself as a list of segment identifiers.

Segment Routing control plane utilizes the existing IGP protocols by adding an extension to OSPF and IS-IS. The IETF draft states that Segment Routing can be used with MPLS and IPv6 data planes and can offer multi-service capabilities of MPLS such as VPN, VPWS and VPLS. We started by testing the distribution of segment identifiers in an MPLS network using IPv4 and IPv6 control planes. Then, we tested the integration with MPLS control and data plane for Layer-2 and Layer-3 services. We used two configuration profiles, based on participating vendors' support.

Profile	IGP	Data Plane	Control Plane
1	IS-IS	MPLS	IPv4
2	IS-IS	MPLS	IPv6

After the Segment Routing (SR) nodes established IS-IS sessions, we verified the distribution of node and adjacency segments.

We then configured MPLS L3VPN service between the edge SR nodes. We verified that the VPNv4 and VPNv6 prefixes/labels were exchanged between the service end point nodes.

We also verified correct encoding of the data path, sending IPv4 and IPv6 traffic in the L3VPN service. Traffic was forwarded between service end-points using the labels corresponding to the previously advertised node segment identifiers. All traffic followed shortest path route as expected.

The figure below depicts the SR network, composed of three groups of vendors that passed the test. We verified each group separately.

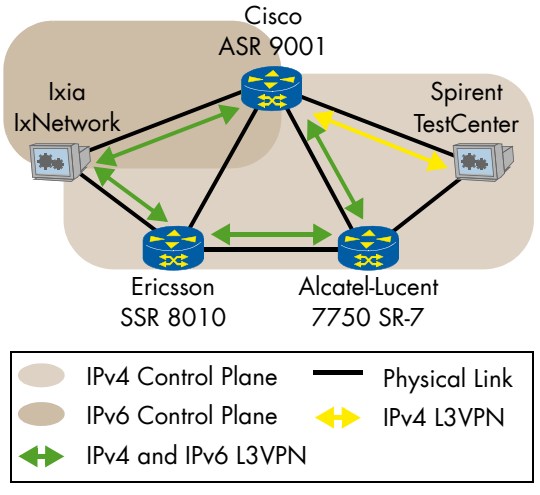

Figure 5: Segment Routing

We observed that one of the SR nodes did not consider the P-Flag (signal for penultimate hop popping) when it encoded the MPLS label stack.

Segment Routing: Fast Reroute (FRR)

To be a serious core-backbone technology, Segment Routing aims to enable sub-50ms protection against network failures. Segment Routing utilizes the local repair properties of IP Fast Reroute (IP FRR) in conjunction with explicit routing to protect and maintain end-to-end connectivity without requiring additional signaling upon failure.

In this test, we focused on two FRR approaches: Loop Free Alternate (LFA) and Topology Independent LFA (TI-LFA). These approaches are defined in the IETF documents: RFC5286 and draft-francois-spring-segment-routing-ti-lfa

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

Topology Independent Loop Free Alternate. By design, Segment Routing eliminates the need to use Targeted LDP (TLDP) sessions to remote nodes as in Remote LFA (RLFA) or Directed Forwarding LFA (DLFA). This, combined with SR's explicit routing, minimizes the number of repair nodes and simplifies network recovery through a list of optimized detour paths for each destination.

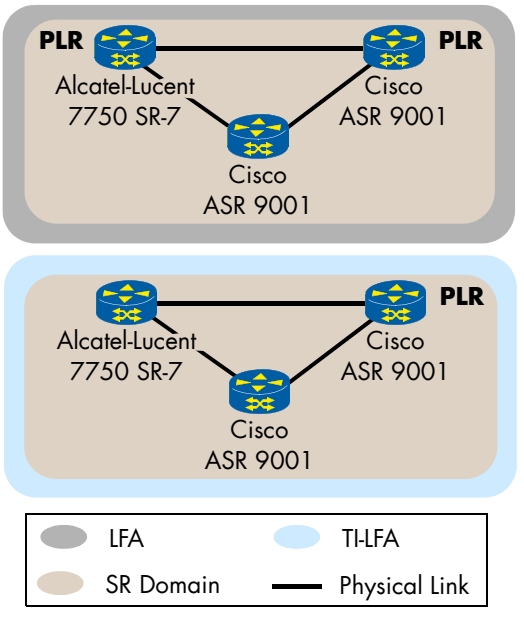

Figure 6: Segment Routing FRR

We started this test by verifying the distribution of node and adjacency segment IDs on each of the nodes. We also verified the Forwarding Information Base (FIB) when the LFA was not configured. As expected, there was a single forwarding entry for each of node segment IDs. We performed baseline measurement for the packet loss using bidirectional traffic from a traffic generator.

Then we configured LFA on the nodes and verified that the nodes installed backup forwarding entry in FIB. While the traffic was running via the primary path we disconnected the link and measured the service interruption time based on the packet loss. After we verified that the traffic was taking the backup path, we restarted the traffic generator

statistics and reconnected the previous disconnected link to measure service interruption time during recovery. The measured service interruption time was between 0.6 ms and 33.4 ms for LFA and between 23.2 ms and 37.8 ms for TI-LFA.

We continued the test with increasing the link metric over one of the links and make sure that there was no LFA backup path installed in FIB. We then configured TI-LFA on Cisco ASR 9001 device and verified the backup path. The device's CLI showed the label stack of three labels including the adjacency segment ID from the mid node. We repeated the LFA test procedure for TI-LFA.

Both, Cisco ASR 9001 and Alcatel-Lucent 7750 SR-7 successfully participated in the LFA part of the test. In addition, Cisco ASR 9001 successfully participated in the TI-LFA part.

Segment Routing and LDP Interworking

Since MPLS technology is widely deployed in the service provider networks and green field deployment is not always feasible, the Segment Routing architecture needs to provide interworking mechanism for seamless interconnection with existing MPLS networks.

In this test we looked at SR and LDP interworking. The architecture defines two functional blocks, SR mapping server and mapping client. SR mapping server provides interworking between SR and LDP networks. The mapping server advertises remote-binding Segment ID for prefixes from non-SR capable LDP nodes. SR mapping client uses this mapping to forward the traffic to the LDP nodes.

We started the test with the SR mapping server – the function that interconnects SR capable device and non-SR capable LDP device. We verified the advertised prefix corresponding to non-SR capable LDP device and its associated Segment ID.

We then verified that the SR capable device (mapping client) processed the mapping and programed its MPLS forwarding table accordingly.

We also verified the control and data plane operations of an Ethernet VPWS service between the SR and non-SR capable devices. We generated bidirectional traffic between the attachment circuits to verify the proper encoding of the data path. The following figure depicts the successful combination that we verified.

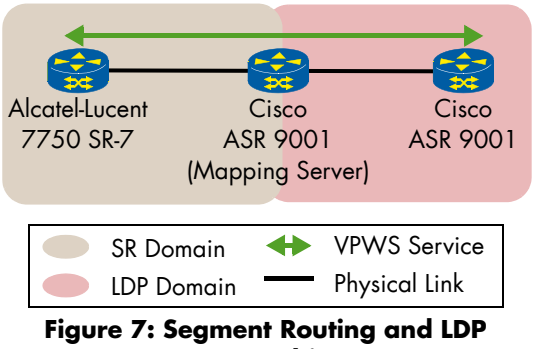

Figure 7: Segment Routing and LDP Interworking

Initially, we observed that mapping client used implicit null label as advertised in the mapping server prefix-SID sub-TLV instead of the explicit label value derived from the SR capability TLV. During the hot-staging the vendors corrected the code and successfully tested the requested functionality.

BGP Prefix Independent Convergence

The increase in the BGP routing table size and multi-path prefix reachability could be seen as challenges in large networks. The IETF draft draft-rtgwg-bgp-pic-02 introduces BGP Prefix Independent Convergence (BGP PIC) as a solution to address the classical flat FIB structure by organizing BGP prefixes into Path-List consisting of primary paths along with precomputed backup paths through FIB hierarchy.

BGP PIC attempts to keep network convergence time independent of the number of prefixes being restored. The IETF draft, in work-in-progress state, describes two failure scenarios: core (BGP PIC Core) and edge (BGP PIC Edge). BGP PIC Core describes a scenario where a link or node on the path to the BGP next-hop fails, but the next-hop remains reachable. BGP PIC Edge describes a scenario where an edge link or node fails, resulting in a change of the next-hop.

We performed both scenarios in this year's event. In all tests we used Ixia IxNetwork to emulate a Customer Edge (CE) device attached to two different upstream PEs, each provisioned with an L3VPN service terminating at a common downstream PE.

The participating vendors agreed to which the primary path by advertising a higher BGP local preference value to the downstream PE. Upstream PEs peered with the emulated dual-homed CE using eBGP as shown in the figure. We configured multi-hop BFD sessions to monitor BGP next-hop reachability between the upstream and downstream PEs. The multi-hop BFD was running at 20 ms intervals.

In all test scenarios, vendors enabled BGP PIC on the downstream PE device. We verified that each of the 50,000 VPNv4 and 50,000 VPNv6 prefixes advertised were learned over two different paths, and installed in the Forwarding Information Base (FIB) at the downstream PE as primary and backup path.

We ran a baseline test by sending unidirectional traffic on each of the VPN prefixes using Ixia IxNetwork and verified that all packets were indeed traversing the primary path and arrived at the destination without loss and duplication. In the case of BGP PIC Core, we then emulated a physical link failure on the primary path, while the next hop remained reachable and verified that traffic failover to the backup path. The failover time ranged from 60 ms to 120 ms.

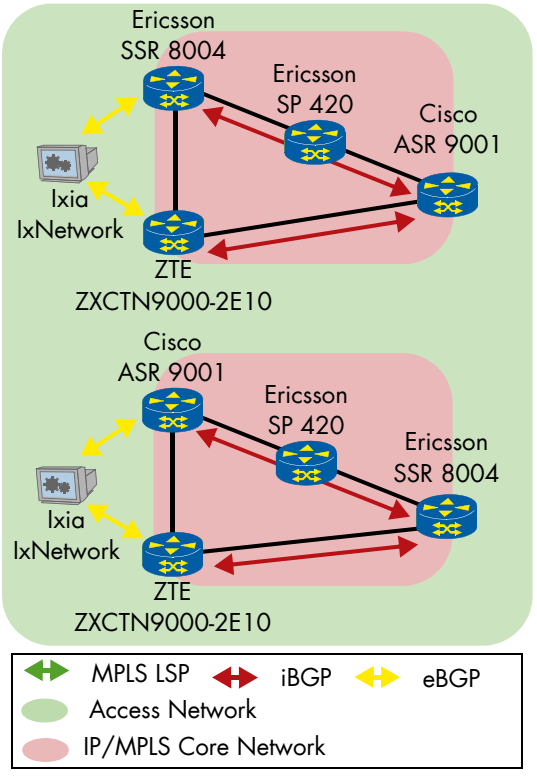

Figure 8: BGP PIC Core Test Results

In the case of BGP PIC Edge, we shut down all core-facing interfaces on the primary next-hop upstream PE, resulting in a failure of the primary next-hop. In all test pairs, we measured a failover time ranging from 64 to 216 ms. The measured failover time met our expectation given the number of advertised prefixes.

Finally we tested a recovery scenario by clearing the failure emulation. In this case we expected the traffic to be forwarded over the primary path. In all test scenarios we measured a recovery time of up to 0.2 ms. Some scenarios showed no traffic loss at all.

The following devices served as BGP-PIC nodes: Cisco ASR 9001 and Ericsson SSR 8004. Ixia IxNetwork functioned as emulated CE devices. Cisco ASR 9001, Ericsson SSR 8004 and ZTE ZXCTN9000-2E10 successfully participated in the test as IP/MPLS PE node. Additionally, in all test pairs Ericsson SP 420 acted as an IP/MPLS Provider (P) router.

RSVP-TE Fast Reroute

EANTC started MPLS Fast Reroute (FRR) interoperability tests in 2006 in a time that implementations were based on individual author drafts. Since then, MPLS FRR has been repeatedly tested at our interoperability events. In this year's event, vendors expressed new-found interest in testing RSVP-TE FRR. We created a test topology as shown the figure. We configured primary and backup Label Switched Paths (LSP) between the Point of Local Repair (PLR) and the Merge Point (MP), acting as ingress and egress router respectively. The participating vendors enabled RSVP-TE FRR facility backup for the LSP to protect against link connectivity failure on the primary path. L3VPN services were configured to run over the protected LSP.

Initially we generated bidirectional traffic at constant bi rate over the configured services, and verified that all packets were forwarded over the primary path and received at destination without loss. We generated Loss of Signal (LoS) by pulling the active link either from the PLR or the MP.

During this phase we simulated failure scenario and once successful, we followed by a recovery test in which the failed link was reconnected. We measured an out-of-service time less than 50 ms during the failover and less than 1 ms during the recovery.

Ericsson SP 415 successfully participated as a PLR. The merge router was ZTE's ZXCTN9000-2E10. BTI 7800 acted as an intermediate node and Ixia IxNetwork was used for traffic generation.

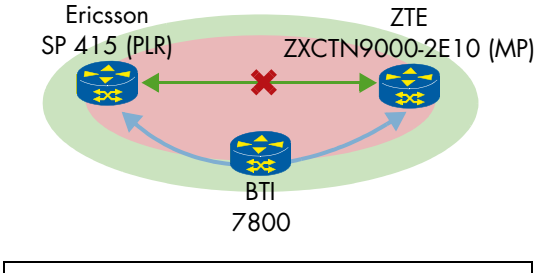

Figure 9: RSVP-TE Fast Reroute

Software Defined Networking

The Open Networking Foundation (ONF) defines Software Defined Networking (SDN) as the physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices. In this year's interoperability test, we showcased multiple examples of SDN which were not just limited to OpenFlow implementation. We extended our test to include the IETF-defined Path Computation Element Communication Protocol (PCEP) as well as NETCONF with Yang models.

OpenFlow: Rate Limiting

Maintaining service quality and SLAs is very crucial for service providers and network operators. OpenFlow implements traffic classification and rate limiting using meter tables and meter bands. Meter tables are used to store a list of meter bands, where each meter band specifies the rate at which the band applies, the burst size and the way packets should be processed.

The OpenFlow protocol defines two main meter band types, which specifies the action applied for matching meter bands: Band type DROP defines a simple rate that drops packets that exceed the band rate value, while band type DSCP REMARK defines a simple DIFFSERV policer that remarks the drop precedence of the DSCP field.

For readers of our 2014 white paper, this test will look familiar. What makes it different this year is the participation of many new vendors.

In this test, the participating vendors were combined in pairs, each pair consisted of an OpenFlow Controller and an OpenFlow Switch (Forwarder). We configured the OF controller with two band types, DSCP DROP for low, and DSCP REMARK for high traffic class as depicted in the below table.

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

Once we verified that the OF Controller installed the flow entries and their corresponding meters into the OF Switch tables, we generated bidirectional traffic for each traffic class at constant rates as shown in the table. We observed neither traffic loss nor DSCP value change at this stage as expected.

We then doubled the traffic rate for the high traffic class, then monitored that half of the high traffic was remarked with the DSCP 0 as expected.

In order to make sure that the low traffic class was also being metered, we also doubled the low traffic rate, and observed that half of the low traffic was dropped as expected.

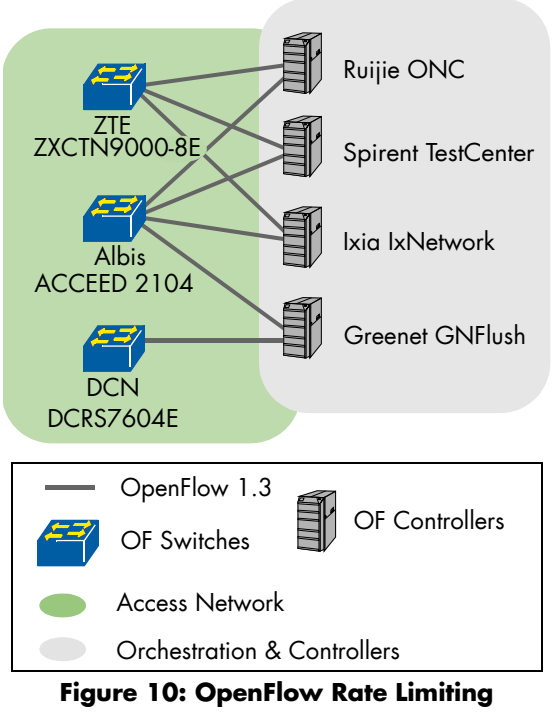

Figure 10: OpenFlow Rate Limiting

The following vendor devices successfully participated in this test: Albis ACCEED 2104, DCN DCRS-7604E, ZTE ZXCTN9000-8E, acting as OF Switch and Greenet GNFlush, Ixia IxNetwork, Ruijie ONC and Spirent TestCenter functioned as OF Controllers.During the test we discovered a section of the Openflow 1.3 specification, which was interpreted in different ways by the vendors. Section A.3.4.4 states that *"The prec_level field indicates by which amount the drop precedence of the packet should be reduced if the band is exceeded"*. This was interpreted in two ways:

1. If the band was exceeded, the DSCP value of the output packet is set to the configured prec_level value.
2. If the band was exceeded, the prec_level value is subtracted from the DSCP value of the incoming packet.

OpenFlow: Link Protection and ECMP Load Balancing

OpenFlow implements load balancing and link protection using Group tables described in the OpenFlow specification 1.3. A group type can be set to select load balancing or Fast Failover for link protection. The group will contain multiple buckets, each with a defined weight. The weight parameter is set to 0 in a Fast Failover scenario, while in a load balancing scenario the weight is set to 1 in ECMP load sharing. Also the weight can be set to a number different than 1 for unequal load sharing but for this test, we focused on equal load sharing.

In the **1:1 link protection scenario** the OpenFlow controller was configured to install primary and backup path into the OF switch using Group type Fast Failover and weight of 0. While we transmitted traffic using Ixia IxNetwork, we emulated a failure condition by pulling the link of the primary path. The recorded failover time was inconsistent among test implementations. While in some test pairs the failover time was below 50 ms, we recorded in other pairs failover time up to 1 second. The importance of the test, however was the interoperability between vendors. After we cleared the failure condition, the traffic reverted to the primary path. The recorded failover time was below 50 ms.

The following vendor devices successfully participated in this test: DCN DCRS-7604E, ZTE ZXCTN9000-8E and ZXCTN9000-3E, acting as OF Switch. Ruijie ONC, Greenet GNFlush and Ixia IxNetwork functioning as OF Controllers.

In an **Equal Cost Multi Path (ECMP) Setup** we configured the OpenFlow controller to enable ECMP on the OF switch using Group type select and weight of 1. We generated traffic using Ixia IxNetwork. We successfully verified that traffic was equally load balanced between all ECMP links. When we triggered failure condition on one of the links, traffic failed over to the second link. For some test pairs the failover time was less 50 ms. In other pairs we measured failover time up to 900 ms. After the failed link was recovered, traffic reverted and was equally load balanced between the links. The recovery time was below 50 ms. As in the case of 1:1 link protection, the importance of the test, however was the interoperability between vendors.

The topology for both tests, load-balancing and 1:1 protection were based on two OF Switches connected via two links. Both switches communicated with an OF Controller via the OF Channel.

The below diagram depicts the pair combinations between the OF Controllers and OF Switches.

The following vendor devices successfully participated in this test: DCN DCRS-7604E, ZTE ZXCTN9000-8E and ZXCTN9000-3E, acting as OF Switch. Ruijie ONC, Greenet GNFlush and Ixia IxNetwork functioning as OF Controllers.

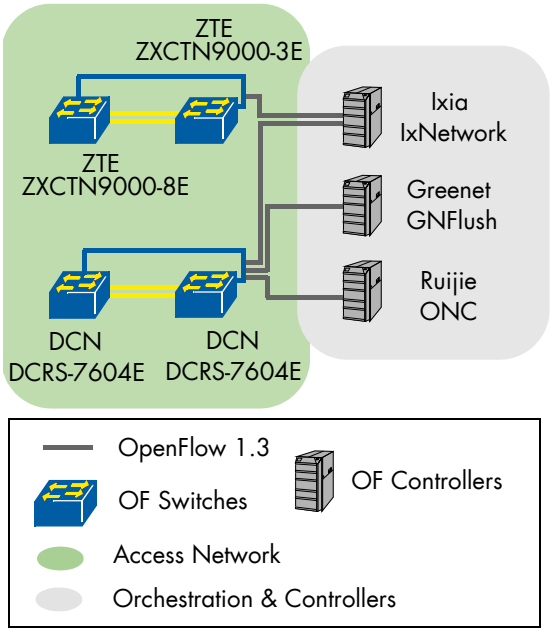

Figure 11: OpenFlow: Link Protection and Load Balancing

OpenFlow Bandwidth on Demand/ Bandwidth Guarantee

In an OpenFlow network, an OpenFlow controller may provide an interface to customer applications. Service attributes could be modified on demand or scheduled by the applications through this interface.

The test case explored two options for bandwidth on demand. The first option is used to provide a higher bandwidth upon user demand for a limited time period, and the second is used to provide a higher bandwidth until a pre-defined amount of data (quota) is consumed by the user. In both options, the original bandwidth rate is restored once the time or quota limit is reached.

In this test, Albis demonstrated an OpenFlow based network using Albis ACCEED 2104, acting as an OF switch connected to an OpenDaylight controller. The controller was connected on the northbound interface to an application provided by Albis. The application was programmed to dynamically change the meter band on the OF switch based on the amount of bandwidth consumed or added.

Initially we configured the OF controller to install flow entries for high and low class in the OF switch. We expected the OF Switch to install both flows into the flow table. We set the band rate for high class to 100 Mbit/s and for low class 200 Mbit/s.

We first configured the customer application to double the band rate for high class for 5 minutes.

We generated traffic for the high class at 200 Mbit/s for a duration of 5 minutes. We initially observed no loss. After the five minutes elapsed, we expected 50% loss of the high class traffic.

In the next step, we configured the customer application to decrease the meter band rate for low class if the consumed traffic exceeded the predefined quota of 750Mbytes.

We generated traffic for the low class at 100 Mbit/s for a duration of about 20 seconds. We first observed no loss. When the consumed quota exceeded 750 Mbytes, we expected the received traffic rate for low traffic to be 1 Mbit/s.

Stateful Path Computation Element (PCE)

In a stateful PCE model, a centralized controller maintains a global and detailed view of the network state Traffic Engineering Database (TED) and connection state Label Switch Path Database (LSPDB).

Both, TED and LSPDB, are used by the PCE controller as an input for optimal path computation. An active stateful PCE does not only consider the LSPDB as an input, but also control the state of the LSPs i.e. PCE can modify/re-optimize existing LSPs. Moreover, an instantiation capable PCE can setup new LSP and delete those LSPs.

In addition to the network abstraction, PCE provides open and well-defined Path Computation Element Protocol (PCEP) interface to the network nodes. This facilities PCE deployment in Software-Defined Network (SDN) architectures.

PCE-initiated RSVP-TE Paths. The initiation of paths by a PCE provides a mechanism for the programmability of MPLS networks. An application could request a path with certain constraints between two network nodes by contacting the PCE. The PCE computes a path that satisfies the constraints, and instructs Path Computation Client (PCC) network node to instantiate and signal the path. When the path is no longer required by the application, the PCE requests a tear down of it.

We started the test with the verification of PCEP sessions between the PCCs and the PCE controller and the verification of update and instantiation capabilities. We then verified the status of BGP Link State (BGP-LS) protocol used to synchronize TED between the network and PCE. We also verified the content of TED and LSPDB via requests to PCE.

We then sent a request to the PCE to instantiate RSVP-TE LSPs using its REST API. The LSPs were used as transport tunnels for a Virtual Private Wire Service (VPWS). We generated bidirectional traffic using traffic generator to verify that the PCC have successful signal the RSVP-TE tunnels.

Next, we verified LSP states synchronization between the PCCs and PCE. We started with disabling PCE network interface towards the PCCs, then we verified that the traffic was forwarded without loss using the PCE signaled tunnels. After we enabled the PCE interface we verified that the PCE has synchronized LSPDB.

Topology

(Topology Page 2)

Subsequently, we verified the ability of the PCE to update the LSP. Using its REST API, we sent an update request to the PCE with an explicit path. We observed that the PCE sent the update request to the corresponding PCC with the new path. We generated bidirectional traffic using traffic generator to verify that the PCC have successful signal the new RSVP-TE tunnels and performs switchover in make-before-break fashion.

As a final step, we sent a request to the PCE to delete the RSVP-TE LSP.

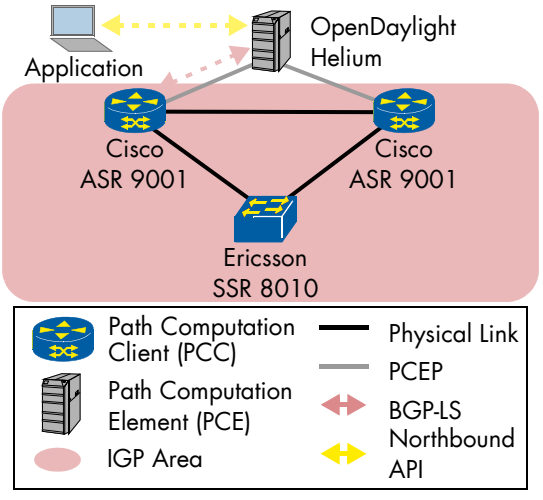

Figure 12: PCE-initiated RSVP-TE Paths

In this test, we tested OpenDaylight Helium as the PCE and Cisco ASR 9001 as the PCC.

PCE-initiated Segment Routing Paths. Internet Engineering Task Force (IETF) draft "PCEP extension for Segment Routing" brings new Segment Routed Explicit Route Object (SR-ERO) used to carry Segment Routing path.

Ericsson demonstrated a setup of explicit Segment Routing tunnel using Ericsson's version of OpenDaylight. We generated bidirectional IP traffic for L3VPN service that used SR tunnel to verify the data plane of both PCCs, Ericsson SSR 8010 and Ericsson Virtual Router (EVR).

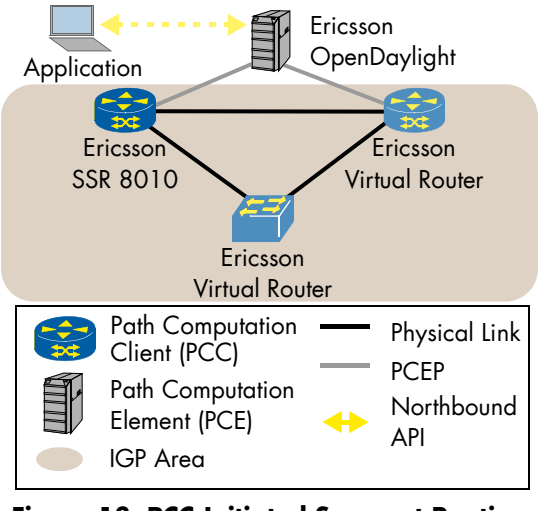

Figure 13: PCC-Initiated Segment Routing Paths

PCC-Initiated RSVP-TE Paths. In MPLS traffic engineering architecture the LSP state is available only locally on each edge LSR. Stateful PCE provides higher degree of network visibility, including the set of computed paths and reserved resources in use in the network.

Active stateful PCE may have control of a PCC's LSP and actively modify the LSP attributes considering operator-defined constrains and LSPDB as input for LSP optimization.

We started the test by verifying the PCEP sessions between the PCCs and the PCE controller and verification of update and, instantiation capabilities. We then verified the status of TED synchronization protocol used to synchronize TED between the network and PCE.

After we verified the TED and LSPDB, we configured RSVP-TE LSPs between a given pair of PCCs. Once we verified that LSP path was taking the direct link between PCCs, we delegated the LSPs to the PCE. We then verified that PCE has accepted the LSP delegation for both LSPs and sent a request to the PCE to update the LSP path using the PCE REST API. For some of the test pairs we observed that the switch-over was hitless.

We then sent a request to PCE to update the LSPs state to operational down.

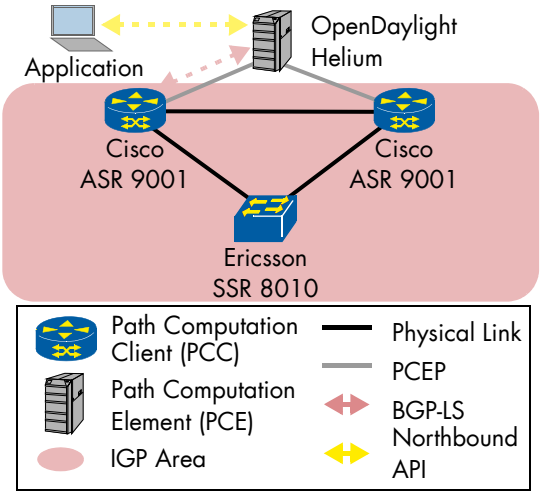

Figure 14: PCC-Initiated RSVP-TE Paths

After the verification of the LSP state on the PCCs, we revoked LSP delegation and the PCE was no longer permitted toF modify the LSPs.

The diagram depicts the test pair that passed the test case requirement for PCC-initiated LSPs. We tested OpenDaylight Helium as the PCE, and Cisco ASR 9001 as the PCC.

Network Provisioning with NETCONF

The NETCONF protocol aims to provide means to configure and maintain a multi-vendor network. The NETCONF RFC6241 describes a list of tools and functions - the latter referred to as "capabilities"- which could help with configuration sequencing, validation, profiling, and roll-back in case of errors.

NETCONF uses YANG, a data modeling language described in RFC 6020, to define the configuration and state parameters. YANG can represent complex data models such as lists and unions. YANG also supports nested data definitions.

A Client/Server model is used by NETCONF, where one or multiple clients would connect to a NETCONF server instance on the managed device.

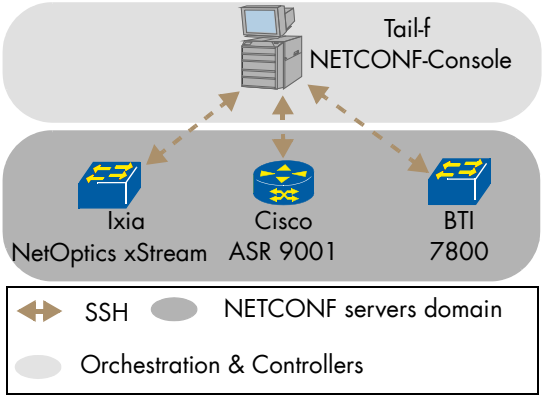

Figure 15: NETCONF Provisioning

In this test we used the NETCONF testing tool NETCONF–Console, from Tail-f to test the NETCONF server implementations in the following devices: BTI 7800, Cisco ASR 9001 and Ixia NetOptics xStream.

The test examined NETCONF session data exchange between the client and server. Then evaluated NETCONF functions like YANG model exchange, listing and modification of running and/or candidate configuration profiles, filtering queried configuration results, changing configuration on the device, then verifying the configuration updates.

We observed that while all participated vendors had passed the test, a few tweaks on the console were necessary to fully complete the test due to variation of configuration profile support between the vendors. In other words, some vendors supported only running configuration profiles, while other vendors supported candidate configuration only, but with limited capability of the candidate profile such as delayed commit.

Clock Synchronization

Since 2008, the year the second version of the Precision Time Protocol (PTP; IEEE standard 1588-2008) was published, we have been testing PTP in our interoperability events. The protocol has been developed further over this period of time – not just with the addition of new profiles (e.g., the telecom profile for phase/time), but also with specifications of qualitative aspects of nodes and the network limits. A work done within the scope of ITU-T study group 15 question 13.

Our focus in this event was on phase/time synchronization with full and partial support. PTP is designed to deliver accurate time synchronization over packet networks, where the non-deterministic behavior is the biggest hurdle. This is where full timing support comes in: it alleviate such effects by controlling the accuracy at each node from the clock source to the end application. Of course, deploying a network which has full timing support may not be an option for some operators. This is addressed by the set of standards dedicated to partial timing support in the network.

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

For frequency synchronization, we continued using the G.823 SEC mask as a requirement. The primary time reference clock was GPS using L1 antennas located on the roof of our lab.

Phase/Time Synchronization with Full Timing Support: T-BC Noise

Achieving a phase/time accuracy of ±1.1µs over a non-deterministic packet network is no small step. For this reason the network should provide full timing support for synchronization, where each intermediate device between the master clock and the end application functions as a boundary clock. Full timing support by itself still does not guarantee achieving the desired phase accuracy since each of these boundary clocks also generate noise as any oscillator does. The ITU-T recommendation G.8273.2 deals directly with the noise specifications for the boundary clocks and defines the maximum absolute constant and dynamic time error allowed for a T-BC. Constant time error (cTE) is set to ±50 ns for class A, ±20 ns for class B; dynamic time error (dTE) is set to 40 ns.

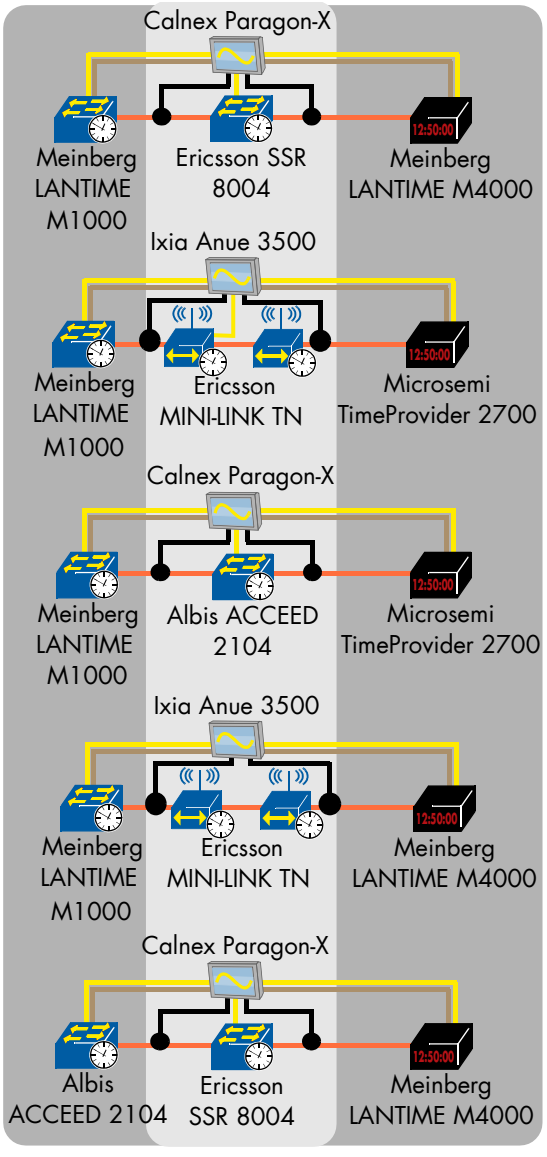

Figure 16: Phase/Time Synchronization with Full Timing Support: T-BC Noise, Setup 1

We measured the time error of PTP packets at the ingress of the boundary clock for the packets originated from the grandmaster to estimate the inbound constant and dynamic noise. At the same time we measured the time error at the egress of the boundary clock. As an additional control, we also measured the physical phase output via 1PPS interface.

We used two test setups, first we used master and slave devices and second we used master and slave emulation. The diagrams depict the results that pass the T-BC requirement of G.8273.2 for constant time error (cTE) and dynamic time error (dTE).

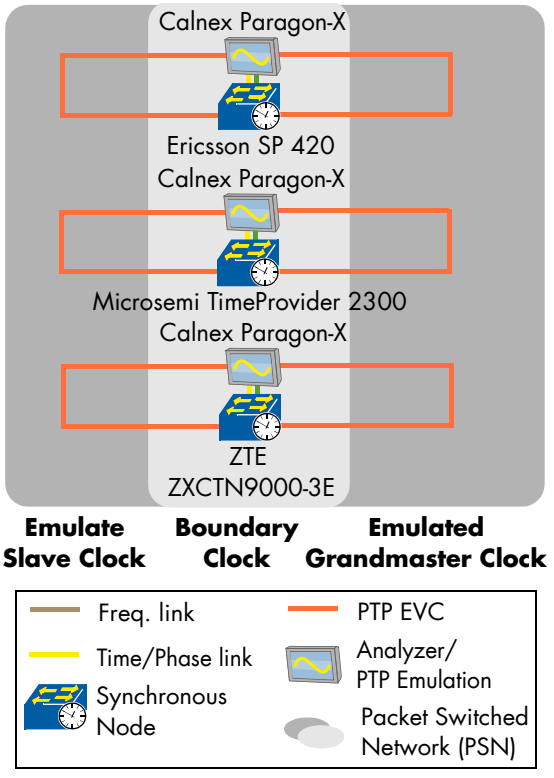

Figure 17: Phase/Time Synchronization with Full Timing Support: T-BC Noise, Setup 2

Albis ACCEED 2104, Ericsson SSR 8004, Ericsson MINI-LINK TN and Ericsson SP 400 used Ethernet encapsulation based on ITU-T G.8275.1 profile.

Microsemi TimeProvider 2300 and ZTE ZXCTN 9000-3E used IP encapsulation based on ITU-T G.8265.1 profile.

Phase/Time Synchronization with Full Timing Support: Non-Forwardable Multicast

The phase and time profile defined in ITU-T G.8275.1 recommendation supports forwardable and non-forwardable Ethernet multicast address. The advantage of using forwardable Ethernet address is the rapid configuration and deployment, since each node along the path may process PTP, but is not required to do so. However, it might be desirable to ensure that by using the non-forwardable address, network operations can ensure that each node from the master to the slave supports and participates in the clock synchronization distribution in order to satisfy the stringent phase requirements.

We started the test with configured non-forwardable address on grandmaster, boundary (T-BC) and slave clock, and allowed T-BC to lock to grand master and slave clock to lock to T-BC. We then did packet capture to verify Ethernet address used for PTP traffic.

We then disabled the PTP on T-BC and verified that the slave clock indicated that PTP reference was lost. We then configured the grandmaster and slave clock to use the forwardable address and

verified that slave clock indicated that PTP reference is being acquired and identify the grandmaster as parent. In the last step, we configured T-BC to use forwardable address and enabled PTP. We then verified that slave clock identified the T-BC as parent.

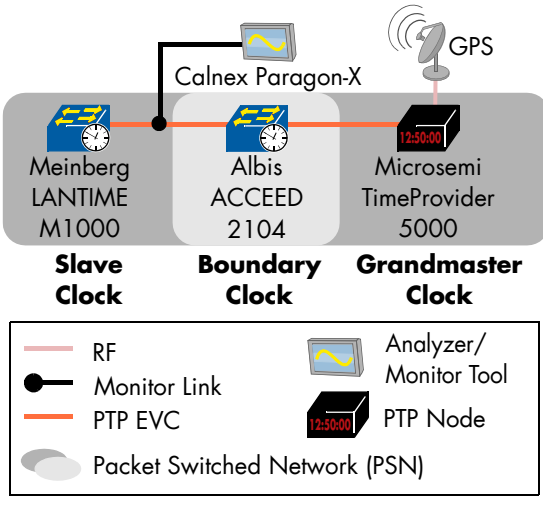

Figure 18: Phase/Time Synchronization with Full Timing Support: Non-Forwardable Multicast

The diagram above depicts the tested combination we executed.

Phase/Time Synchronization with Full Timing Support: Path Rearrangements

One of the key aspect in synchronization networks is the protection of the timing transport during network rearrangement. However, PTP rearrangement can cause accumulation of phase/time error.

In this test we emulated a path rearrangement and measure the resulting effect on the slave clock's phase and frequency synchronization.

We started the test with enabling an impairment profile that emulated a chain of five boundary clocks (T-BC) class A according to ITU-T recommendation G.8273.2 clause 7.1.

Initially, the slave clock was in free-running mode and we allowed it to lock onto the grandmaster clock. We then performed baseline measurements and verified that they passed the requirements. We then restarted the measurements and simulated a path rearrangement event by disconnecting the link for less than 15 seconds. We then enabled an impairment profile that emulated a chain of ten T-BCs.

We verified that the short-term phase transient response complied with the ITU-T G.813 requirements, the phase accuracy requirement of ±1.1 µs and the frequency accuracy requirements of G.823 SEC.

The following diagram depicts the tested combination we executed.

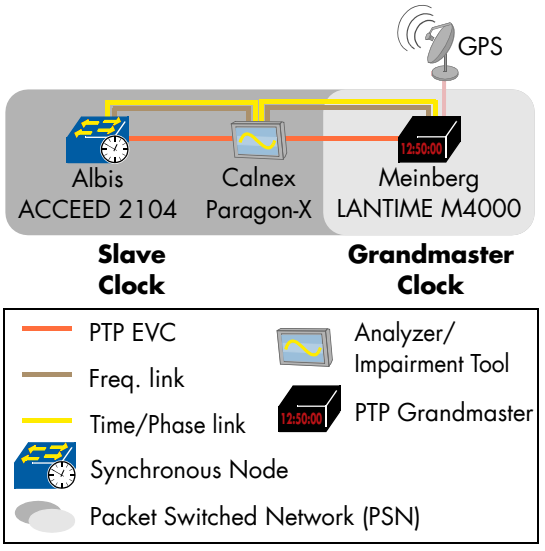

Figure 19: Phase/Time Synchronization with Full Timing Support: Path Rearrangements

Phase/Time Synchronization with Full Timing Support: Hold over Performance

Hold-over time is a crucial metric for mobile service providers. It is a major factor in the decision whether to send a field technician to a cell site to perform urgent maintenance or to delay it for more cost-effective scheduling of operations. In case of a prolonged outage, a slave clock in a radio controller which exceeds its hold-over period will most likely result in major failure in hand-over from (and to) neighboring cell sites. Vendors design their equipment frequency hold-over oscillator performance accordingly. But what about time/phase hold-over performance?

We started the test with the slave clock in free-running mode and allowed it to lock onto the grandmaster clock. We then performed baseline measurements. After passing the masks we set for the test, we restarted the measurements and used an impairment generator to drop all PTP packets, simulating a PTP outage. We then verified that the slave clock was in phase/time holdover mode and Synchronous Ethernet (SyncE) was used as physical reference to provide assistance during the holdover. We let the measurements run over night.

We verified that the short-term and long-term phase transient responses passed ITU-T G.813 masks, and phase accuracy requirement of ±1.1 µs and frequency requirements.

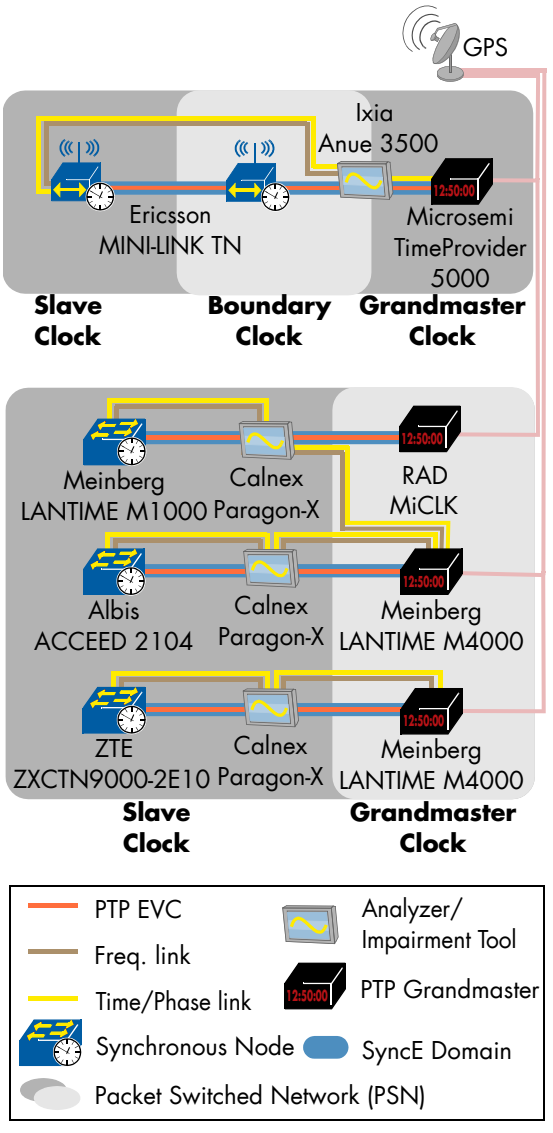

Figure 20: Phase/Time Synchronization with Full Timing Support: Hold over Performance

Albis ACCEED 2104, Ericsson MINI-LINK TN, Meinberg LANTIME M1000, and ZTE ZXCTN 9000-2E10 passed the phase accuracy requirement of ±1.1 µs and frequency requirements of G.823 SEC as slave clocks. The diagram depicts the tested combinations we executed.

Phase/Time Synchronization with Full Timing Support: Microwave Transport

In some deployment scenarios, such as rural areas access, a microwave transport could be a cost-effective solution for mobile backhaul. Microwave radios use Adaptive Coding and Modulation (ACM) to ensure continuous transport under severe weather conditions.

We designed this test case to verify that the accuracy of the phase synchronization does not degrade — in normal and emulated severe weather conditions. To emulate such severe weather conditions, we used an attenuator to reduce the RF signal to the lower modulation scheme.

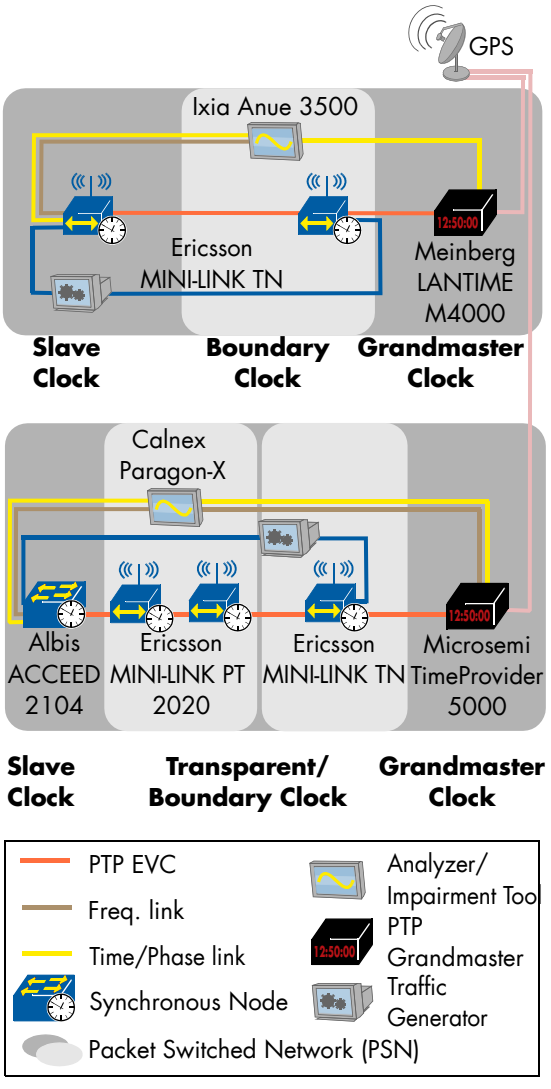

Figure 21: Phase/Time Synchronization with Full Timing Support: Microwave Transport

We started the test with the slave clock in free-running mode and generated traffic according to G.8261 VI2.2 at 80% of the maximum line rate for the maximum modulation scheme and expected no traffic loss. We took baseline measured for phase and frequency from the slave clock.

After passing the requirements, we restarted the measurements on the slave clock and attenuated the signal from 1024QAM to 16QAM on Ericsson MINI-LINK TN and from 1024QAM to 4QAM on Ericsson MINI-LINK PT 2020. Since the bandwidth decreased accordingly, we verified that data packets were dropped according to the available bandwidth. We evaluated the measurements on the slave clock with the requirements and compared them to the baseline measurements. Moreover, we analyzed the short-term phase transient response during the attenuation of the signal and verified it complies with ITU-T G.813.

We performed the test for two microwave systems, Ericsson MINI-LINK TN as boundary clock and Ericsson MINI-LINK PT 2020 as transparent clock.

Phase/Time Assisted Partial Timing Support: Delay Asymmetry

Assisted Partial Timing Support (APTS) was developed as a concept integrating the benefits of a Global Navigation Satellite System (GNSS), such as the Global Positioning System (GPS), and network-based timing technologies. The basis is to use the GPS as the primary time reference at cell sites (or at an aggregation point close to the cell sites), complemented by network-based timing distribution, to assist maintaining the time/phase accuracy during holdover periods when the GPS signal is unavailable.

One of the issues that may affect APTS is link asymmetry. The test started with both grandmaster and slave clocks locked onto GPS while PTP was also active. We then impaired PTP by dropping all PTP messages and verified that no transients occurred, indicating that GPS was the primary source. We also verified that the slave clock detected the PTP failure. We then re-enabled PTP packet flow and introduced packet delay variation (PDV) based on G.8261 Test Case 12 to simulate a network of 10 nodes without on-path support. Afterwards, we took baseline phase and frequency measurements from the slave clock.

We proceeded by restarting the wander measurements and then disconnected the GPS antenna from the slave, simulating a GPS outage.

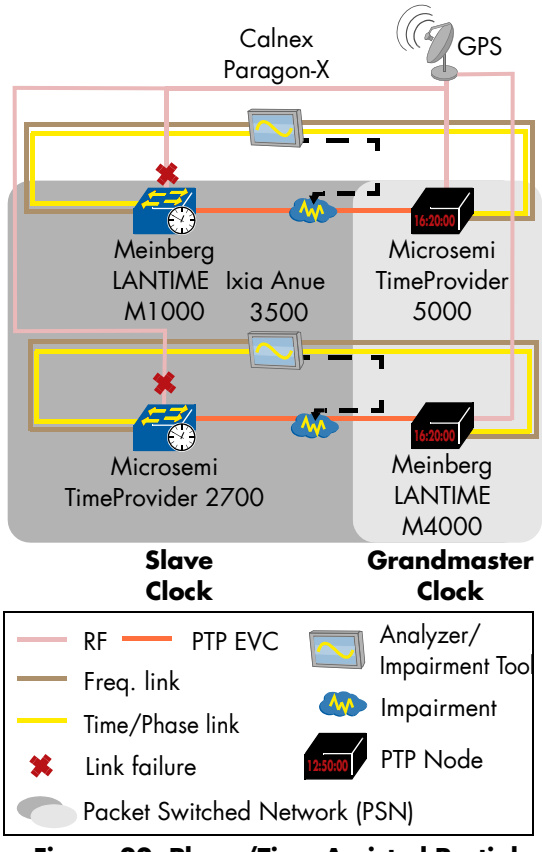

Figure 22: Phase/Time Assisted Partial Timing Support: Delay Asymmetry

While the GPS antenna was disconnected, we introduced a unidirectional delay of 250 µs in the direction towards the slave clock. We verified that the delay asymmetry was detected by the slave. We evaluated the results according to the phase requirement of ±1.1 µs and G.823 SEC MTIE mask.

The diagram depicts the results that passed the phase accuracy requirement of ±1.1 µs and frequency accuracy requirements of G.823 SEC.

We observed different slave implementation of delay asymmetry detection and correction. One slave implementation detected and calibrated the delay asymmetry while the GPS was disconnected.

Phase/Time Assisted Partial Timing Support: Hold over Performance

GNSS such as GPS is an optimal choice for phase synchronization as it can deliver —under normal working conditions — a maximum absolute time error in the range of ±0.1µs. This allows deployment of accurate phase distribution. However GPS is subject to jamming, which could bring severe operational risks. Since GPS provides phase, frequency and time of day information, currently the only protocol that could serve as an alternative to delivering this information is the IEEE 1588-2008 or Precision Time Protocol (PTP).

The test started with both grandmaster and slave clocks locked onto GPS and PTPv2 was active for the first setup. For the second test setup, the boundary clock was locked to distributed grandmaster. We then impaired PTP by dropping all PTP messages and verified that no transients occurred, indicating that GPS was the primary source, while also verifying that the slave clock detected PTP failure. We then re-enabled PTP packet flow and introduced packet delay variation (PDV) based on G.8261 Test Case 12 to simulate a network without on-path support. Following this, we took baseline phase and frequency measurements from the slave clock.

We proceeded by restarting the measurements and disconnected the GPS antenna, simulating an outage. We evaluated the results according to the phase requirement of ±1.1 µs and G.823 SEC MTIE mask. We also evaluated the results according to G.813 for short-term and long-term phase transient response.

As expected, we measured higher phase accuracy when the GPS was available. We measured maximum time error of 77 ns, which is below our measurement threshold. We still managed to measure maximum time error of 127 ns with PTP — which is also below our set goals.

The diagram depicts the two different setups that passed the phase accuracy requirement of ±1.1 µs and frequency accuracy requirements of G.823 SEC. We did overnight measurement for one of the test setups.

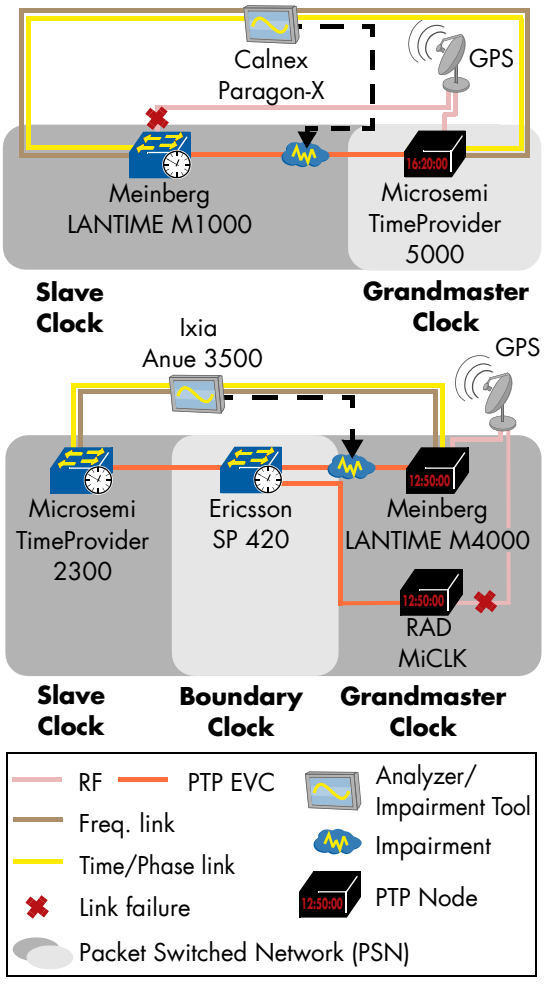

Figure 23: Phase/Time Assisted Partial Timing Support: Hold Over Performance

Phase/Time Synchronization: Distributed (Edge) Grandmaster with Physical Clock Reference

The Distributed (Edge) Grandmaster approach brings the PTP grandmaster closer to the base stations. This solution reduces the number of T-BCs on the synchronization path and thus reduces the amount of noise, which removes some of the engineering complexity needed to ensure the required phase accuracy is maintained.

We started the test with the slave clock locked to the distributed grandmaster clock (SFP plugged into boundary clock) which was locked to its internal GPS receiver. In addition, we provided physical reference, a Synchronous Ethernet (SyncE) from Primary Reference Clock (PRC) to the grandmaster. We then started wander measurements and verified that no phase transients occurred when we disconnected and reconnected the physical reference from the grandmaster.

We then performed baseline measurements. After passing the masks we set for the test, we restarted the measurements and disconnected GPS antenna, simulating GPS signal outage. We verified that the short-term and long-term phase transient responses passed ITU-T G.813 masks, and phase accuracy requirement of ±1.1 µs and frequency requirements.

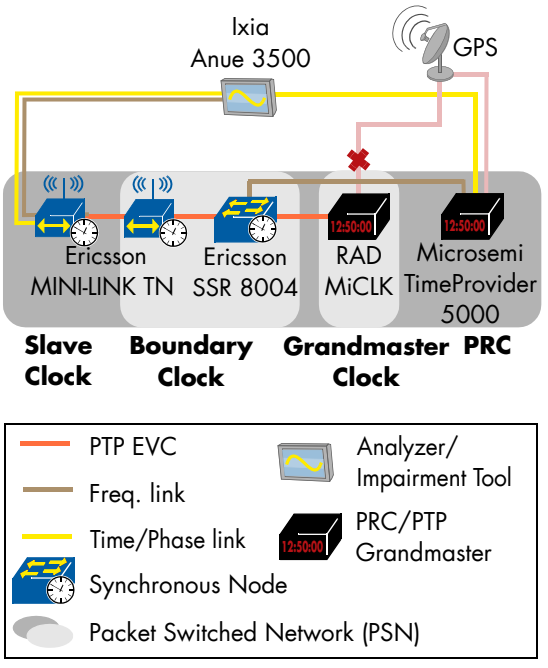

Figure 24: Phase/Time Synchronization: Distributed (Edge) Grandmaster with Physical Clock Reference

The diagram depicts the test setup and results. Ericsson SSR 8004 hosted the RAD MiCLK SFP as distributed grandmaster.

Phase/Time Synchronization: Master Clock Scalability

An important characteristic of a PTP grandmaster is the amount of clients it can support. We designed a test to verify that with the maximum client utilization, PTP accuracy quality meets the requirements for phase.

We started with two slave clocks in free-running mode and allowed them to lock to the grandmaster clock via a transparent clock. We then performed baseline measurements. After passing the requirements, we restarted the measurements and started the emulated clients.

The number of emulated clients was set according to the vendor's specifications of supported clients, to match the maximum clients of the non-emulated slave clocks. We verified that no transients occurred when we started the emulated clients. We then evaluated the results of the slave clock according to the phase and frequency requirements and also compared it with the baseline measurements.

We configured all emulated clients with a message rate of 128 packets per second (sync, delay request and delay response). The following devices were tested for their PTP client scalability:

Meinberg LANTIME M4000 with 2048 clients, Microsemi TimeProvider 5000 with 500 clients and RAD MiCLK with 20 clients.

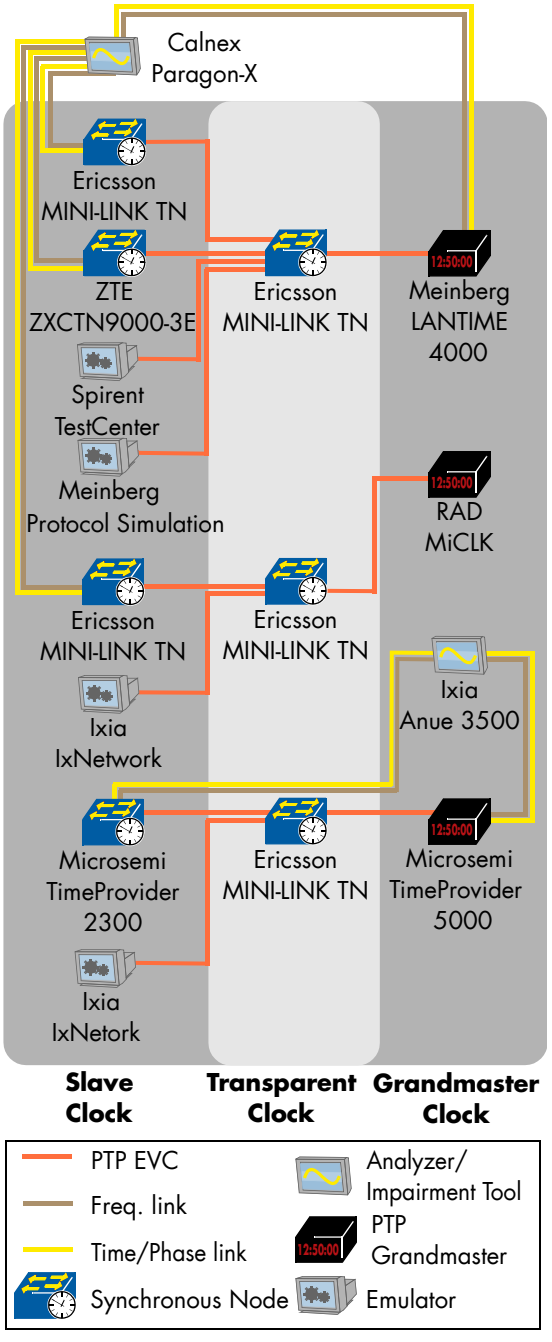

Figure 25: Phase/Time Synchronization: Master Clock Scalability

Network Functions Virtualization Feasibility Test

Huawei and RAD demonstrated a joint solution of VPN service using virtual CPE (vCPE) with two implementation options: One option places virtual functionality at the network edge and the second places it at the customer site. The vCPE combines the virtualized capabilities of Huawei's virtualNetEngine (VNE 1000) with physical capabilities of RAD's smart demarcation devices (ETX-2 and the miniature MiNID). Huawei's VNEs were placed at two locations – at the network edge running on a COTS server, and at the customer site running on the x86 engine built into RAD's ETX-2. RAD's RADview D-NFV Orchestrator configured both physical and virtual networking, while Huawei's web portal managed both VNEs and served as a customer portal, allowing enterprise customers to self-manage their services.

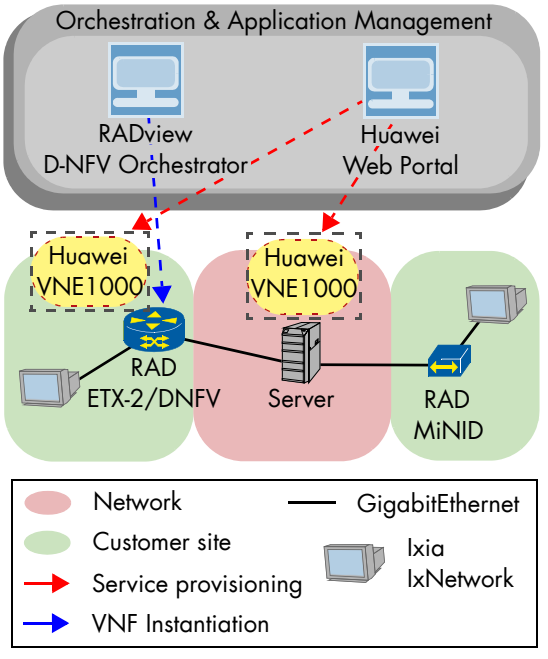

Figure 26: NFV Feasibility Demo

The feasibility test was performed in three phases:

Phase 1. Set up and verification of physical connectivity

The connectivity between sites was set up and verified using L3 IP TWAMP tool of the RAD ETX-2 and MiNID devices. RAD's RADview management system configured and activated the test.

Phase 2. Virtual networking setup and VNF instantiation

RAD's RADview D-NFV Orchestrator performed instantiation of the Huawei VNE virtualized network function (VNF). First, the VNF image was on-boarded to the RADview VNF repository, and then the VNF orchestrator downloaded it to the ETX-2. The D-NFV Orchestrator also configured relevant physical components and internal virtual networking (Open vSwitch) in the ETX-2. Upon Phase 2 completion, the underlay physical network was properly configured and VNFs were instantiated and ready for services creation.

Phase 3. Service creation and management

In the test, we acted as an enterprise customer, using Huawei's web portal, without touching the CLI, to create a VPN service. We than verified that the service was established and the configured bandwidth was actually enforced. Next, we modified the service bandwidth using Huawei's web portal. This updated the configuration both in the VNE 1000 running in the RAD ETX-2, as well as in the COTS server. As a result the customer service capacity was updated without changing the underlay network configuration.

This scenario demonstrated a new mean to VPN service rollout in a matter of minutes. As the test demonstrated, the enterprise IT manager can roll the service by him or herself.

Demonstration Scenarios

Clock Synchronization

In the packet clock synchronization area, we constructed a network composed of two segments: Assisted Partial Timing Support (APTS) and full on-path timing support.

Microsemi TimeProvider 5000 was integrated as grandmaster located in one of the data centers. Ixia Anue 3500 was simulating a network without on-path support, which introduced packet delay variation (PDV) based on G.8261 Test Case 12. Meinberg LANTIME M1000 acted as an edge grandmaster, connecting the APTS and full on-path timing network segments. Calnex Paragon-X simulated multiple T-BCs in the full on-path timing network. Albis ACCEED 2104 acted a slave clock. The devices in the full on-path timing segment used ITU-T G.8275.1 profile.

Stateful Path Computation Element (PCE)

In the PCE area, we integrated two scenarios. In the first scenario, we had a use case of PCE in IP/MPLS network where RSVP-TE was used as LSP signaling protocol. We demonstrated setup of PCE/PCC-initiated RSVP-TE LSPs.

The second PCE scenario was part of the SR network. We had a use case of PCE in SR network. We demonstrated setup of explicit PCE-initiated Segment Routing paths between data centers.

Segment Routing

Beside the IP/MPLS core network, we created a second network domain composed of successful Segment Routing (SR) results. We used SR to transport layer 2 and layer 3 services between the data centers.

We interconnected the IP/MPLS and SR domains based on SR and LDP interworking test results. Alcatel-Lucent 7750 SR-7 acted as SR mapping client and Cisco ASR 9001 acted as SR mapping server between the SR and LDP domains.

VM Mobility Between Data Centers over EVPN VXLAN

Virtualization has revolutionized and transformed today's data centers. However, one of the challenges of virtualization is resource distribution and re-distribution (VM Mobility). We at EANTC have witnessed this last year when we tested VM mobility over the MPLS data plane, but the concept was not successful due to limited support of IP localization.

This year, however, we continue on this concept by testing VM mobility between two multi-tenant data centers inter-connected via EVPN with VXLAN in the data plane.

The participating vendors chose two topologies to interconnect between the data centers. The first was a spine-leaf topology using two Cisco Nexus 9396 (leaves), one in each data center, interconnected through a Cisco Nexus 7000 (spine) as Route Reflector (RR). The second topology directly connected one Alcatel-lucent 7750SR in one data center to one Juniper MX80 in the second data center.

We used Ixia RackSIM to emulate the virtual machines (VMs) and the forward and reverse migration of these VMs between the data centers. In one data center, we had two virtual machines running on two separate hypervisors and in the second data center we had one hypervisor without any virtual machine. The second data center also hosted Ixia's VM Manager which is responsible for VMs migration. RackSIM was connected to each of the leaf devices in the first topology and to each of the devices in the second topology.

We started by sending traffic between the two virtual machines and while sending the traffic, we ran a script to move one virtual machine in one data center to the second data center. We observed that the virtual machine was migrated to the second data center and the traffic was sent over the EVPN network to the second data center.

OpenFlow Bandwidth on Demand

Albis Technologies successfully demonstrated a bandwidth on demand setup using ACCEED 2104 as an OpenFlow switch connected to an OpenFlow controller (OpenDaylight Helium). Albis provided an application that connected to the REST Northbound API of the controller and kept track of quota and time for two different traffic profiles. The demonstration setup was able to increase or limit traffic rates based on predefined time and quota limitations. Also applied the correct policies to each of the traffic profiles according to their defined traffic rates.

Abbreviations

- Adaptive Coding and Modulation (ACM)
- Assisted Partial Timing Support (APTS)
- Backbone MAC Address (B-MAC)
- Broadcast, Unknown Unicast and Multicast (BUM)
- Customer Edge (CE)
- Customer MAC Address (C-MAC)
- Directed LFA (DLFA)
- Designated Forwarder (DF)
- Equal Cost Multi Path (ECMP)
- Ethernet Segment Identifier (ESI)
- Ethernet Virtual Private Network (EVPN)
- Fast Reroute (FRR)
- Forwarding Information Base (FIB)
- Global Navigation Satellite System (GNSS)
- Global Positioning System (GPS)
- Integrated Routing and Bridging (IRB)
- Interior Gateway Protocol (IGP)
- Label Switch Path Database (LSPDB)
- Link Aggregation Control Protocol (LACP)
- Multi-Protocol Border Gateway Protocol (MP-BGP)
- Network Layer Reachability Information (NLRI)
- Network Virtualization Overlay (NVO)
- Packet Delay Variation (PDV)
- Path Computation Element Protocol (PCEP)
- PIM Sparse Mode (PIM-SM)
- Precision Time Protocol (PTP)
- Provider Backbone Bridges Service Instance Identifier (PBB I-SID)
- Provider Backbone Bridging (PBB)
- Provider Backbone Bridging - Ethernet VPN (PBB-EVPN)
- Quality of Service (QoS)
- Resource Reservation Protocol - Traffic Engineering (RSVP-TE)
- Route Reflector (RR)
- Software Defined Networking (SDN)
- Subsequent Address Family Identifiers (SAFI)
- Synchronous Ethernet (SyncE)
- Targeted LDP (TLDP)
- Tenant System (TS)
- Virtual Extensible LAN (VXLAN)
- Virtualized Networking Function (VNF)
- Virtual Private LAN Service (VPLS)
- Virtual Private Wire Service (VPWS)

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

20150309 v0.5