29-04-2013

Deliverable DJ1.1.3: Transport Technologies and Operations

Deliverable DJ1.1.3

Contractual Date:	31-03-2013
Actual Date:	29-04-2013
Grant Agreement No.:	238875
Activity:	JRA1
Task Item:	Task 1
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	NORDUnet
Document Code:	GN3-13-109
Authors:	Kurosh Bozorgebrahimi (UNINETT), Alberto Colmenero (NORDUnet), Marcin Garstka (PSNC), Michal Hažlinský (CESNET), Victor Olifer (Janet), Jan Radil (CESNET), Josef Verich (CESNET), Henrik Wessing (DTU)

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 2007–2013) under Grant Agreement No. 238875 (GÉANT).

Abstract

Following on from the theoretical research into and extensive testing of Carrier Class Transport Network Technologies (CCTNTs) performed by JRA1 Task 1 earlier in the project, this report describes the additional testing carried out in Year 4. The tests covered MPLS-TP, EoMPLS and time-sensitive transport technologies.

Table of Contents

Executive Summary 1

1 Introduction 3
- 1.1 Background 3
- 1.2 Purpose and Audience 3
- 1.3 In this Document 4
- 1.4 Disclaimer 5

2 MPLS-TP Testing 6
- 2.1 Overview 6
- 2.2 Introduction 6
- 2.3 Test Objective 7
- 2.4 Test Infrastructure 8
- 2.5 T-MPLS/MPLS-TP Architecture Test 10
  - 2.5.1 Test Objective 10
  - 2.5.2 Technology Briefing 10
  - 2.5.3 Test Setup 10
  - 2.5.4 Test Description 11
  - 2.5.5 Test Results 18
  - 2.5.6 Test Conclusions 18
- 2.6 Services 18
  - 2.6.1 Test Objective 18
  - 2.6.2 Technology Briefing 18
  - 2.6.3 Test Setup 19
  - 2.6.4 Test Description 19
  - 2.6.5 Test Results 31
  - 2.6.6 Test Conclusions 32
- 2.7 OAM Testing 32
  - 2.7.1 Test Objective 32
  - 2.7.2 Technology Briefing 32
  - 2.7.3 Test Setup 33
  - 2.7.4 Test Description 33

2.7.5 Test Results 36
2.7.6 Test Conclusions 36
2.8 LSP 1+1 Linear Protection Test 37
2.8.1 Test Objective 37
2.8.2 Technology Briefing 37
2.8.3 Test Setup 38
2.8.4 Test Description 38
2.8.5 Test Results 46
2.8.6 Test Conclusions 46
2.9 Conclusions 47

3 EoMPLS Interoperability Tests 48
3.1 Overview 48
3.2 Introduction 48
3.3 Test Objective 49
3.4 Test Infrastructure 50
3.4.1 Description 50
3.4.2 Configurations 50
3.4.3 Hardware and Software Versions 51
3.5 Test 1 – L2VPN (Pseudowire) Introductory Test 52
3.5.1 Test Setup 52
3.5.2 Configuration 52
3.5.3 Test Description 53
3.5.4 Expected Results 53
3.5.5 Test Results 53
3.5.6 Test Conclusions 54
3.6 Test 2 – Pseudowire OAM Test 55
3.6.1 Test Setup 55
3.6.2 Configuration 55
3.6.3 Test Description 56
3.6.4 Expected Results 56
3.6.5 Test Results 57
3.6.6 Test Conclusions 61
3.7 Test 3 – L2VPN Fast Reroute (Node Protection, Facility Backup) Test 62
3.7.1 Test Setup 62
3.7.2 Configuration 62
3.7.3 Test Description 63
3.7.4 Expected Results 63
3.7.5 Test Results 64

3.7.6 Test Conclusions 65
3.8 Test 4 – VPLS Introductory Test 65
3.8.1 Test Setup 65
3.8.2 Configuration 65
3.8.3 Test Description 66
3.8.4 Expected Results 66
3.8.5 Test Results 66
3.8.6 Test Conclusions 68
3.9 Test 5 – VPLS OAM Test 69
3.9.1 Test Setup 69
3.9.2 Configuration 69
3.9.3 Test Description 70
3.9.4 Expected Results 70
3.9.5 Test Results 70
3.9.6 Test Conclusions 77
3.10 Test 6 – VPLS Fast Reroute (Node Protection, Facility Backup) Test 78
3.10.1 Test Setup 78
3.10.2 Configuration 78
3.10.3 Test Description 78
3.10.4 Expected Results 79
3.10.5 Test Results 79
3.10.6 Test Conclusions 80
3.11 Conclusions 80

4 Time-Sensitive Transport Technologies 82
4.1 Introduction 82
4.2 Physical Media Dependent Layer 83
4.2.1 Technology Overview 83
4.2.2 Physical Media Latency Experiment 85
4.2.3 Conclusions for the Physical Layer 87
4.3 Optical Transport Network (OTN) 88
4.3.1 Technology Overview 88
4.3.2 OTN Delay Experiment 89
4.3.3 Conclusions for the OTN Layer 95
4.4 Packet Layer 95
4.4.1 Technology Overview 95
4.4.2 PBB-TE Delay Experiment 99
4.4.3 Conclusions for PBB-TE 104
4.5 Hybrid Solutions 104

4.5.1 Technology Overview 104
4.5.2 Hybrid Solution Experiments 107
4.5.3 Conclusions for Hybrid Networks 110
4.6 Conclusions 110

5 Final Conclusions 112

References 114

Glossary 116

Table of Figures

Figure 2.1: Alcatel-Lucent 1850 TSS-320 platform 8
Figure 2.2: EXFO FTB-500 tester 8
Figure 2.3: EXFO FTB-8130 NGE testing module 8
Figure 2.4: MPLS-TP test setup 9
Figure 2.5: MPLS-TP architecture 10
Figure 2.6: MPLS-TP sections 11
Figure 2.7: Tunnel/LSP configuration 12
Figure 2.8: Tunnel/LSP OAM configuration 13
Figure 2.9: Tunnel with 1+1 protection 13
Figure 2.10: Tunnel/LSP configuration via craft terminal 14
Figure 2.11: Tunnel representation from NMS 15
Figure 2.12: PW configuration window in NMS 16
Figure 2.13: First end of the PW configured via craft terminal 17
Figure 2.14: PW representation in NMS 17
Figure 2.15: MEF service types 19
Figure 2.16: EPL service configuration window 20
Figure 2.17: Traffic flows for one of the EVCs termination points configured via craft terminal 20
Figure 2.18: EPL service verification with EXFO tester 21
Figure 2.19: EVPL service configuration window 22
Figure 2.20: No traffic in VLAN 30 prior to service creation 23
Figure 2.21: Creation of EVPL for VLAN 30 24
Figure 2.22: Tester detecting traffic on VLAN 30 25
Figure 2.23: Outgoing and incoming traffic flows at the EVC terminations via craft terminal 26

Figure 2.24: EVP-LAN service configuration 27
Figure 2.25: EVP-LAN configuration window 28
Figure 2.26: Traffic flows for EVP-LAN service 29
Figure 2.27: EVP-Tree service configuration window 30
Figure 2.28: Port role configuration 30
Figure 2.29: Traffic flows configuration 31
Figure 2.30: OAM MPLS-TP label 33
Figure 2.31: CV configuration during LSP setup 34
Figure 2.32: Single- and dual-ended Delay Measurement concept 35
Figure 2.33: DM configuration window via craft terminal 35
Figure 2.34: DM results 36
Figure 2.35: 1+1 protection with OAM CV 37
Figure 2.36: APS protocol 38
Figure 2.37: Test setup 39
Figure 2.38: Current PWs in Tunnel 1 39
Figure 2.39: Traffic running in VLAN 10 40
Figure 2.40: Traffic running after fibre cut 41
Figure 2.41: NMS notification of protection status 41
Figure 2.42: 1+1 LSP protection 42
Figure 2.43: Protection Management window 43
Figure 2.44: Traffic switched back to main LSP 44
Figure 2.45: Pattern Loss when traffic switches to main LSP 44
Figure 2.46: Force switch activation window 45
Figure 2.47: Force switch process 45
Figure 2.48: Traffic running after force switch command 46
Figure 3.1: Topology of the testbed 50
Figure 3.2: The bypass LSPs for the Fast Reroute tests 63
Figure 4.1: DWDM transmission components 83
Figure 4.2: Field testbed scenarios. Latency over DWDM using transponders (A) and without transponders (B) 86
Figure 4.3: Loop in PSS2 - 2 times OTN processing 90
Figure 4.4: Looping through PSS2 to PSS1 – 4 times OTN processing 91
Figure 4.5: OTN triangle test 92
Figure 4.6: OTN measurements: delay in µsec as function of the number of loops for different frame sizes 94
Figure 4.7: Ethernet evolution from classic Ethernet to PBB 97
Figure 4.8: PBB-TE at DTU – all-in-one device virtualised into several switches 99
Figure 4.9: PBB-TE delay measurements showing delay in µsec as function of the number of PBB-TE switches 100

Figure 4.10: The Janet Ciena PBB-TE testbed 102
Figure 4.11: RTT of PBB-TE and VLAN-based frames for different frame sizes 104
Figure 4.12: Schematic diagram of the uni-directional transport through the fusion add/drop muxponder when using four of the 10 inputs 106
Figure 4.13: QoS and PVD in switches/routers vs fusion-based device (switch) 106
Figure 4.14: Setup for experiment 1 108
Figure 4.15: Setup for experiment 2 108

Table of Tables

Table 4.1: Latency based on the three different scenarios as shown in Figure 4.2 86
Table 4.2: Latency sources of DWDM path 87
Table 4.3: Internal loop in PSS1 to isolate the delay to the tester. All delays in µsec. 90
Table 4.4: Extra delay for 2 and 4 OTN processing functions. Delays in µsec. 91
Table 4.5: OTN measurement scenarios and delay factors 92
Table 4.6: OTN measurement scenarios and calculated delay 93
Table 4.7: Measured delay for OTN loops 93
Table 4.8: Zero-length Ethernet frame delay and difference from expected delays 94
Table 4.9: RTT times 103
Table 4.10: GST performance measurements 109
Table 4.11: SM performance measurements 109

Executive Summary

Following on from the theoretical research into and extensive testing of Carrier Class Transport Network Technologies (CCTNTs) performed by Joint Research Activity 1 Future Network, Task 1 Carrier Class Transport Network Technologies (JRA1 Task 1) during Years 1 to 3 of the GN3 project, this report describes the additional and/or continued testing carried out in Year 4, and presents the test results. The main objective of the tests was to evaluate the current status of implementation in the equipment used; gain practical, operational experience of the technologies; investigate the manageability of the technologies in a multi-domain, multi-vendor environment; and identify the particular benefits of the technologies for NRENs and GÉANT, and how they might use them. The purpose of the report in documenting the test activity and findings is to provide a reference document that will help the GÉANT NREN community, its primary audience, during the design, planning, procurement and implementation of next-generation transport network architectures.

The CCTNTs, relevant tools, technologies and transport network architectures discussed in this report are Multi-Protocol Label Switching Transport Profile (MPLS-TP), Ethernet over Multi-Protocol Label Switching (EoMPLS), and time-sensitive transport technologies.

The MPLS-TP testing was carried out at NORDUnet premises by NORDUnet engineers and over a longer period of time than the lab tests performed by providers earlier in the project. The main objective was to gain more experience with the operations of MPLS-TP and to get a better understanding of the possible benefits of using the Network Management System (NMS) for provisioning MPLS tunnels and pseudowires. The areas tested were: MPLS-TP architecture, services, OAM, and Label Switched Path (LSP) 1+1 protection. The test setup was based on an Alcatel-Lucent 1850 TSS-320; the software installed supported T-MPLS, an earlier version of MPLS-TP, but did not provide full support for MPLS-TP. However, the testing was very useful for understanding the current status of T-MPLS in the ALU 1850 TSS-320 and the usability of the NMS to configure services, LSPs and pseudowires, and for gaining experience with T-MPLS/MPLS-TP technology and becoming familiar with protection and OAM functionality in packet-based networks. The tests showed that the implementations need to become more mature, with the introduction of a control plane and the full set of OAM features. A more agile NMS in combination with a Command Line Interface would be the best solution for allowing operators with different skills to choose the best option for their profile and their needs. Finally, MPLS-TP needs to be fully interoperable with MPLS in order to achieve seamless connectivity of services across domains.

The EoMPLS tests focused on the basic interoperability of the service implementation, allowing the provisioning of Ethernet over MPLS services in a multi-vendor network; use of Ethernet OAM mechanisms (such as Continuity Check Protocol (CCP), loopback, linktrace) to monitor and troubleshoot an EoMPLS service in a multi-vendor network; and fast recovery of an EoMPLS service by using the Fast Reroute mechanisms

provided by the MPLS transport plane in a multi-vendor network. Six tests were performed to verify each of the three aspects of EoMPLS services listed above with regard to both the point-to-point Layer 2 Virtual Private Network (L2VPN) service and the multipoint Virtual Private LAN Service (VPLS) services. The tests were executed by PSNC and CESNET using equipment from Juniper (MX family) and Cisco (Catalyst 6500 and ME3600 families). The results of the tests show that the implementations of the tested services available on Juniper Networks and Cisco Systems routers are interoperable and that routers from the two vendors can be used together in a single network that provides EoMPLS services. One interoperability issue was encountered in the configuration of Border Gateway Protocol-signalled VPLS services. The service recovery tests proved that Fast Reroute mechanisms can be used to achieve sub-50 ms recovery times for EoMPLS services. The Ethernet OAM tests confirmed that CCP can successfully monitor the state of an EoMPLS service in both the provider and customer domains and provide information about loss and recovery of connectivity to remote devices. The other OAM protocols (linktrace and loopback) did not work correctly over the EoMPLS services and require further investigation. Some terminology differences between the terms used by Cisco Systems and Juniper Networks were noted.

For real-time and time-sensitive applications and packet-layer synchronisation, delay and packet delay variation (PDV) are key parameters that must be kept to a minimum. The time-sensitive transport technologies tests focused on the delay-contributing functions in the physical, transport and packet layers. The objective was to quantify the inherent delays in the respective technologies so that a user can estimate delays in a given transport network environment. The technologies evaluated were Dense Wavelength-Division Multiplexing (DWDM), Optical Transport Network (OTN), Provider Backbone Bridge Traffic Engineering (PBB-TE) and a hybrid packet-circuit solution. The main sources of latency in the physical layer of a DWDM transport network are the transmission fibre, fibre-based Dispersion Compensation Modules and transponders (and regenerators). Using three Alcatel 1830 PSS OTN switches, the OTN processing delay was measured in a lab scenario; the propagation delay was found to be insignificant. The delay in PBB-TE switches was found to be directly dependent on the frame length of the Ethernet frames due to the store and forward mechanism. The measurements were done with equipment from two different vendors. The overhead going from the Ethernet VLAN tagging to PBB-TE tagging was also measured and an additional delay of 10 µsec was found. Trials with Fusion Ethernet-based prototype nodes from TransPacket demonstrated a circuit quality of service performance for the guaranteed data: zero packet loss, a node delay much lower than the fibre transmission delay and with a packet delay variation in the nanosecond range. The circuit transport has a low processing overhead and enables efficient transparent router bypass.

With this report the work of JRA1 T1 in the GN3 project comes to an end. The Task's work has covered the study of the Carrier Class Transport Network Technologies most relevant to the NREN community and has provided information and test results that will help the GÉANT NREN community during the building of future networks.

1 Introduction

1.1 Background

For its study of transport network technologies, GN3 Joint Research Activity 1 Future Network, Task 1 Carrier Class Transport Network Technologies (JRA1 Task 1) has focused on Carrier Class technologies, since these provide the scale, functionality, reliability and performance required by the National Research and Education Networks (NRENs) who are the study's primary audience.

During the first phase of the project, the Task carried out theoretical research into CCTNTs, the results of which were documented in "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1]. In the second phase, the Task performed extensive testing, described in "Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing" [DJ1.1.2]. Originally due to finish at the end of Year 3, the Task was granted a one-year extension to continue its investigation of specific areas that were considered relevant for the NREN community or that were not fully completed during the second phase, namely:

- Multi-Protocol Label Switching Transport Profile (MPLS-TP).
- Ethernet over Multi-Protocol Label Switching (EoMPLS).
- Time-sensitive transport technologies.
- Service assurance (ITU-T Y.1564 standard).
- Ethernet Operation, Administration and Maintenance (OAM).

This report documents the work carried out in the first three areas; the second two were investigated in collaboration with JRA2 Multi-Domain Network Service Research, Task 3 Monitoring, and will be documented in a separate deliverable.

1.2 Purpose and Audience

Following on from the testing performed by the Task in the second phase of the project and documented in [DJ1.1.2], this report describes the additional and/or continued testing, and presents the test results. The main objective of the tests was to:

- Evaluate the current status of implementation of the technologies in the equipment used.

- Gain practical, operational experience of the technologies.
- Investigate the manageability of the technologies in a multi-domain, multi-vendor environment.
- Investigate the mutual impact of deploying the technologies in existing/legacy environments.
- Identify the particular advantages and benefits of the technologies for NRENs and GÉANT, and how they might use them.

The purpose of the report in documenting the test activity and findings is to provide a reference document that will help not only the GÉANT NREN community, its primary audience, during the design, planning, procurement and implementation of next-generation transport network architectures, but also all other NRENs, the industry and commercial service providers.

It is the intention of the authors of this report to present their work and its outcomes to external as well as internal audiences. External audiences with whom the knowledge will be shared include appropriate industry conferences and workshops, as well as standards bodies and journals. For example, JRA1 T1's work was presented to both the GÉANT NREN community and commercial vendors (Alcatel-Lucent, Ciena and Infinera) at a two-day workshop in November 2012, organised by NORDUnet (described in Issue 10 January 2013 of CONNECT magazine [CONNECT10Jan2013], while the Task's work on time-sensitive transport technologies was presented at the NORDUnet Conference in September 2012, which included participants from Cisco, Infinera and Juniper as well as from European NRENs and Internet2 [NORDUNet2012TSTT]. Within the project, the authors aim to collaborate further with those GN3 Activities and Tasks that consider the report content to be useful for their own work priorities.

These aims support the overall objective of JRA1, which is to bring innovation to the GÉANT infrastructure by investigating emerging technologies that enhance the network infrastructure and the corresponding portfolio of services offered by GÉANT and the NRENs [TechAnnex].

1.3 In this Document

This document is structured as follows:

- Chapter 2: MPLS-TP Testing:
- Chapter 3: EoMPLS Interoperability Tests:
- Chapter 4: Time-Sensitive Transport Technologies:

  In general, the sections in each chapter are as follows (the exact selection of sections in each chapter depends on the scope of the test being described):
  - Overview.
  - Introduction.
  - Technology Briefing.
  - Test Objectives.
  - Test Setup.
  - Test Description.
  - Test Results.

- Test Conclusions.

- Chapter 5 offers an assessment of the tests and recommendations.

1.4 Disclaimer

It is not the intention of this document to evaluate and compare different suppliers' implementations of the technologies. The objectives of the tests relate to the technologies themselves, as stated in Section 1.2.

2 MPLS-TP Testing

2.1 Overview

This chapter describes the Multi-Protocol Label Switching Transport Profile (MPLS-TP) testing extensions. MPLS-TP testing was extended because the testing carried out in Year 3 (reported in "Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing" [DJ1.1.2]) was executed by personnel from providers in their labs over a very short period of time. By contrast, the Year 4 testing has been carried out at NORDUnet premises by NORDUnet engineers and over a longer period of time. This was made possible by NORDUnet's purchase of an upgrade of their Alcatel-Lucent 1850 Transport Service Switch (TSS) 320 lab equipment to support MPLS-TP features.

This chapter of the report describes the tests and the experience gained during the testing period.

2.2 Introduction

Multi-Protocol Label Switching Transport Profile (MPLS-TP) is a new technology developed jointly by the ITU-T and the IETF. At least, this is what many claim. Others hold a different opinion: for those who have worked with and know MPLS, MPLS-TP is nothing new but just the well-known MPLS with some extra features (e.g. Operation, Administration and Maintenance (OAM), protection and control plane capabilities, all of which could be of potential interest to NRENs) and some features turned off. Despite the disputes around MPLS-TP, this technology has been a major focus for vendors (such as Alcatel-Lucent), who are putting significant effort into developing new hardware that supports it. Perhaps this is where the biggest innovation lies: the hardware being developed with MPLS-TP support is more reminiscent of legacy transport systems (such as Multi-Service Transport Platforms) than of the routers that have traditionally supported MPLS.

The MPLS-TP evolution and standardisation process is still ongoing; vendors are adding new features to their equipment as they are being standardised and in some cases even before. The standardisation process stalled because of disagreements between the IETF and the ITU-T around the subject of OAM [MPLS-TP_facts]. This dispute led to two different tracks of OAM standards: the IETF OAM standards based on MPLS tools such as Bi-directional Forward Detection (BFD) and Label Switched Path (LSP) ping, and the ITU-T OAM standards based on Y.1731. The question that remains is whether these two standards will coexist in the future, or whether the market will finally decide in favour of only one of them.

During the World Telecommunication Standardisation Assembly held last year in Dubai (WTSA-12), recommendations ITU-T G.8113.1, “Operations, Administration and Maintenance mechanism for MPLS-TP in Packet Transport Network (PTN)”, and ITU-T G.8113.2, “Operations, administration and maintenance mechanisms for MPLS-TP networks using the tools defined for MPLS”, were approved [ITU-T Newslog1, WTSA12Blog]. The first recommendation describes MPLS-TP OAM mechanisms built on Ethernet-based OAM protocols, while the second recommendation describes MPLS-TP OAM mechanisms built on IETF-based OAM protocols such as BFD and LSP ping. As required by the ITU-T, the Internet Assigned Numbers Authority (IANA) has assigned a new Ach Code point for Ethernet-based OAM protocols [IETF-RFC6671].

The OAM features tested during the testing process documented in this chapter are based on Ethernet OAM tools.

This chapter focuses on the following tests:

- MPLS-TP architecture.
- Services.
- OAM.
- LSP 1+1 protection.

For more information about the MPLS-TP technology, please refer to “Deliverable DJ1.1.1: Transport Network Technologies Study” [DJ1.1.1] and “Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing” [DJ1.1.2].

The testing during Year 4 was carried out at NORDUnet premises with equipment from Alcatel-Lucent (ALU) purchased by NORDUnet. The ALU 1850 TSS-320 was upgraded to release 3.4 in order to have the features required for the testing. This required the purchase of new controller, matrix and line cards. The configuration was achieved via Network Management System (NMS), which also needed to be upgraded, to release 9.4. No Command Line Interface (CLI) was available. The TSS software (release 3.4) has support for Transport MPLS (T-MPLS), which is an earlier version of the MPLS-TP, standardised by the ITU-T. The major difference between the two protocols concerns the OAM implementation. However, their basic features are very similar. In order to have full support for MPLS-TP, including control plane capabilities, the equipment should have been upgraded to release 4.0. However, this would have required version 9.5 of the NMS, which is a major software upgrade and was not included in the budget. The set of features available in release 3.4 of the ALU 1850 TSS-320 is quite limited and the testing was therefore not as extensive as expected. However, it was still possible to gain valuable experience with MPLS-TP architecture and services.

2.3 Test Objective

The Year 4 tests were a continuation of the testing carried out during Year 3 of the GN3 project. Due to the constraints described in Section 2.2 above, it was not possible to extend the testing as much as originally anticipated, and many of the tests were a repetition of those carried out in Year 3. However, in Year 4 the equipment was in-house and was available over a longer period. For this reason, the main objective was to gain more experience with the operations of MPLS-TP and to get a better understanding of the possible benefits of using the NMS for provisioning MPLS tunnels and pseudowires.

2.4 Test Infrastructure

The test setup was based on an Alcatel-Lucent 1850 TSS-320 (Figure 2.1). This equipment is a multi-service platform capable of delivering Ethernet, MPLS, MPLS-TP and Synchronous Digital Hierarchy (SDH) services. The nodes were equipped with non-redundant control and switching cards and one 10 x 1GE card. NORDUnet A/S purchased these cards for the purpose of these tests. The nodes were interconnected using GE interfaces. The client side was also 1 GE.

In order to verify the configuration, an EXFO FTB-500 tester with an FTB-8130 NGE module was used (Figure 2.2 and Figure 2.3).

Figure 2.1: Alcatel-Lucent 1850 TSS-320 platform

Figure 2.2: EXFO FTB-500 tester

Figure 2.3: EXFO FTB-8130 NGE testing module

The tester was used as termination test equipment, so only Ethernet packets could be inspected, not MPLS. This would have required pass-through functionally, which is not supported in the FTB-8130 NGE module.

The nodes were connected in a ring topology, as shown in Figure 2.4. The nodes were managed via an Alcatel-Lucent Optical Management System (OMS) 1350. The management server was connected directly to each node via Ethernet.

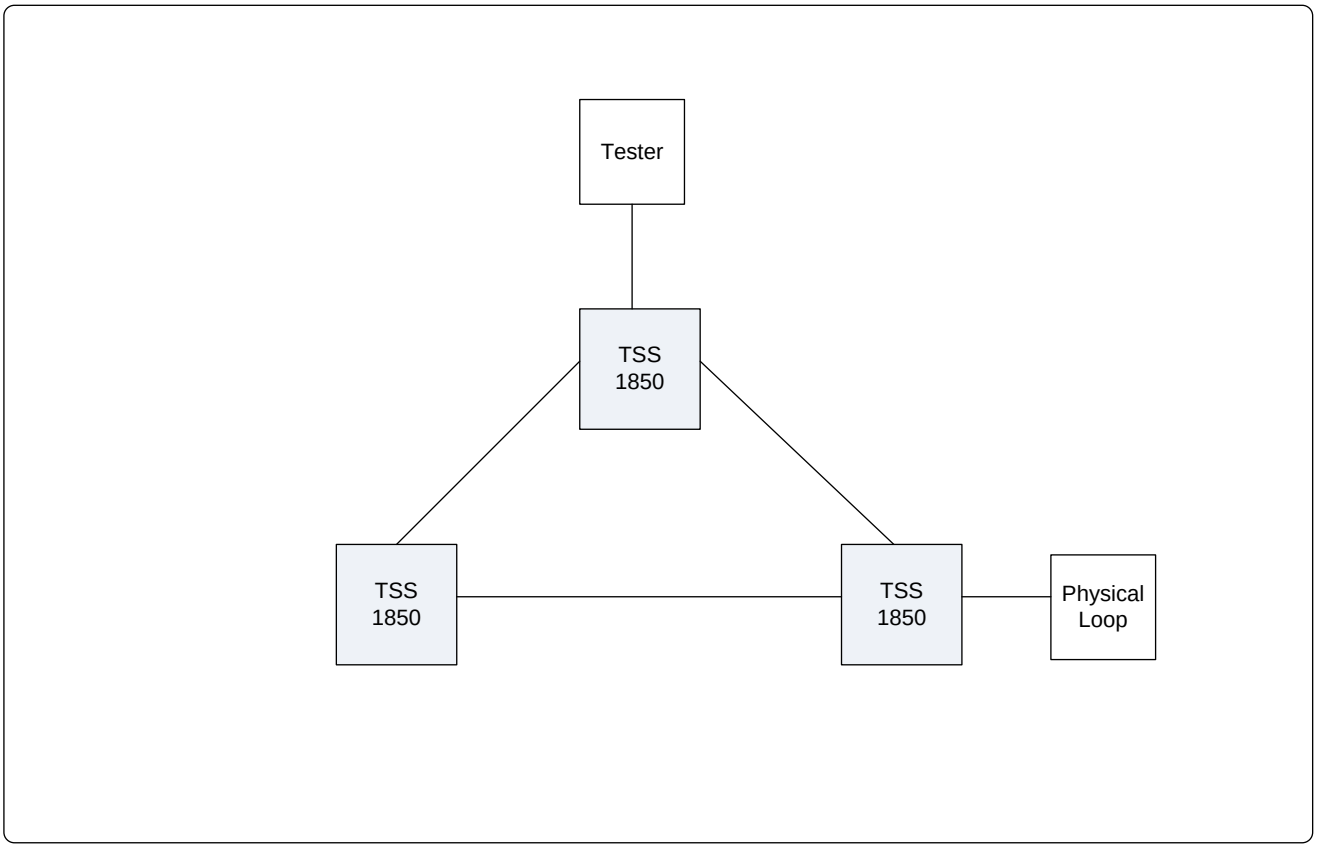


Figure 2.4: MPLS-TP test setup

The Alcatel-Lucent 1850 TSS-320 was run with firmware (FW) release 3.4; the NMS was release 9.4. This software release has basic T-MPLS functionality. For this reason, comprehensive testing of MPLS-TP technology was not possible; this would have required release 4.0. However, this in turn would have required release 9.5 of the NMS, a major software upgrade, which, due to time and budgetary constraints, was not possible.

The Alcatel-Lucent NMS with release 9.4 has support for:

- T-MPLS architecture.
- Ethernet services.
- 1+1 LSP protection.
- OAM Connectivity Verification (CV).

This implementation has no control plane support, which again would have required a newer release. All the configuration was therefore done via NMS. However, the MPLS-TP framework specifies the possibility of being able to operate with and without NMS or control plane.

This test setup is used throughout the whole testing with minor changes when needed.

2.5 T-MPLS/MPLS-TP Architecture Test

2.5.1 Test Objective

The main objective is to demonstrate the configuration and verification of sections, tunnels and pseudowires (PWs) and to understand the use of the NMS for this purpose. As previously mentioned, no CLI was supported to configure these components.

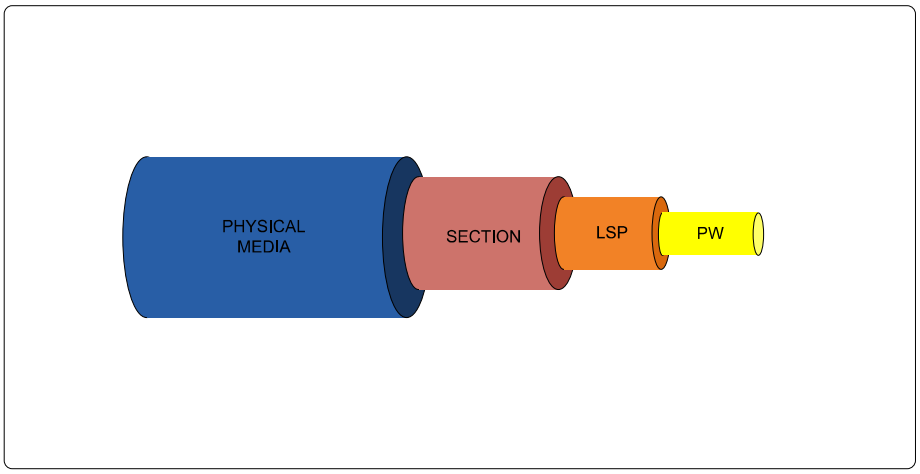


Figure 2.5: MPLS-TP architecture

2.5.2 Technology Briefing

MPLS-TP architecture and its components were extensively described in JRA1 Task 1's first GN3 deliverable, DJ1.1.1 [DJ1.1.1].

2.5.3 Test Setup

The test setup is the same as described in Section 2.4 above. In this case, the NMS and craft terminal software were used for configuration verification.

2.5.4 Test Description

2.5.4.1 *Section Configuration*

The MPLS-TP section represents the physical connection between the nodes. The sections are automatically defined when configuring the physical connections between the nodes in the NMS. Fibre connectivity needs to be in place between the nodes prior to the configuration, otherwise the NMS will report a Loss of Signal (LOS) alarm. Once the physical connection is properly established, the physical alarms will be cleared up. As shown in Figure 2.6, this type of link is defined as a network-to-network interface (NNI). These interfaces carry MPLS-TP traffic.

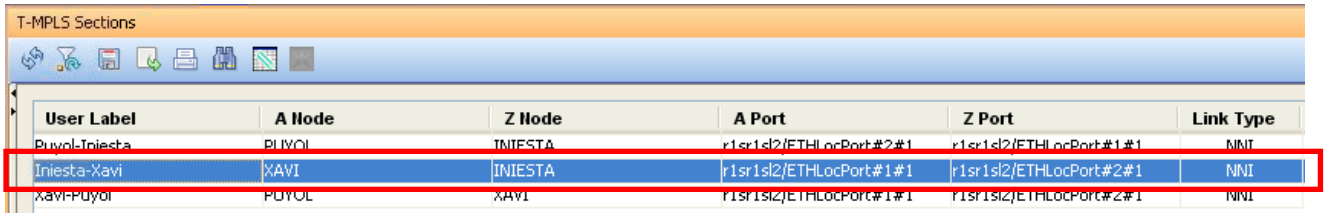


Figure 2.6: MPLS-TP sections

2.5.4.2 *Tunnel Configuration*

An MPLS-TP tunnel, also known as a Label Switched Path (LSP), is used to carry the MPLS-TP pseudowires. A tunnel is defined between two NNIs. During the tunnel configuration, the Quality of Service (QoS) parameters are defined. These parameters define the available bandwidth for the services carried inside this tunnel. As shown in Figure 2.7, the LSP can be configured with 1+1 protection.

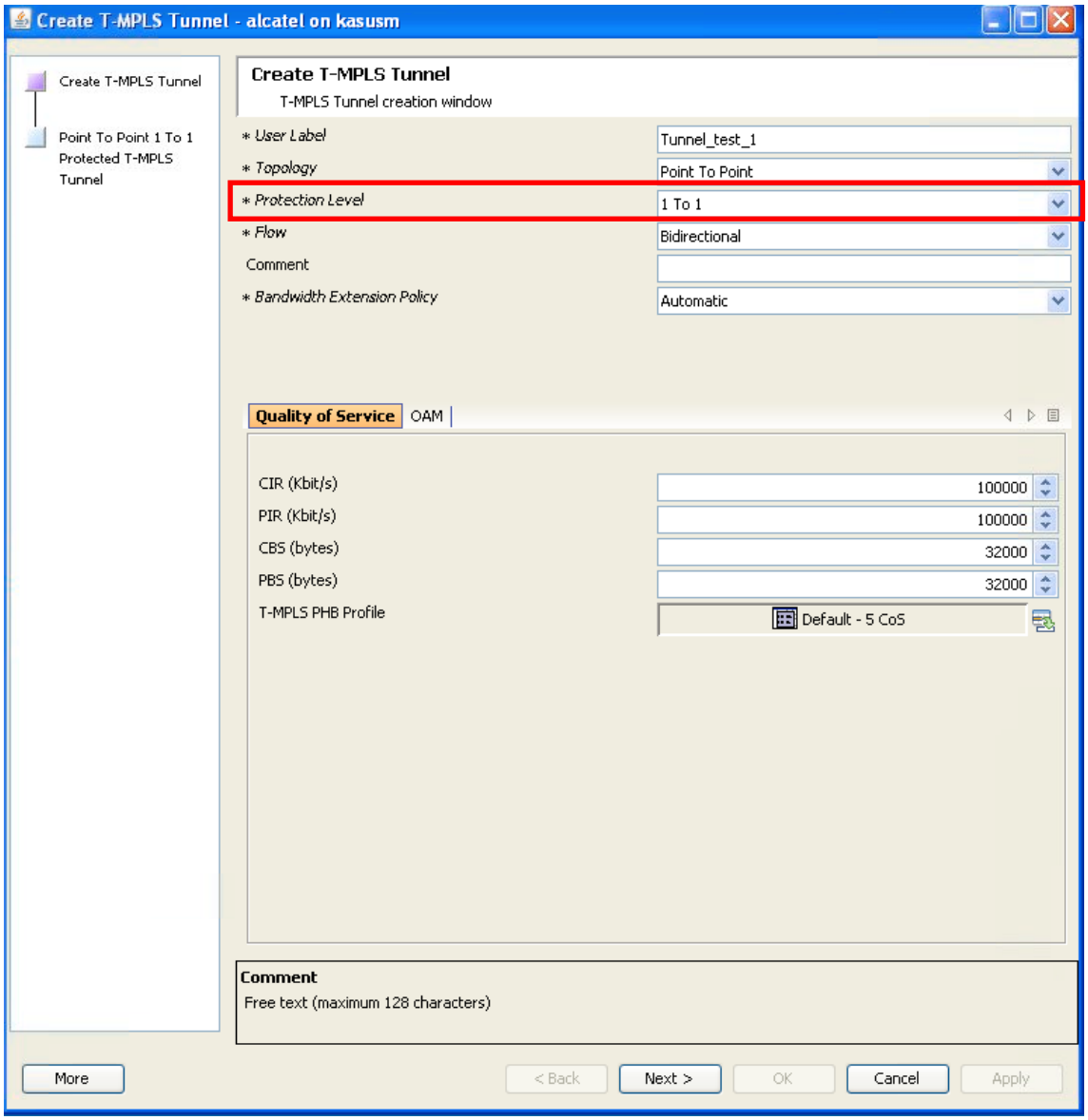


Figure 2.7: Tunnel/LSP configuration

During tunnel configuration, OAM Connectivity Verification (CV) can also be configured, as shown in Figure 2.8. CV is used to detect a failure condition in the network. CV has been tested in relation to 1+1 protection. This test is described in Section 2.8.

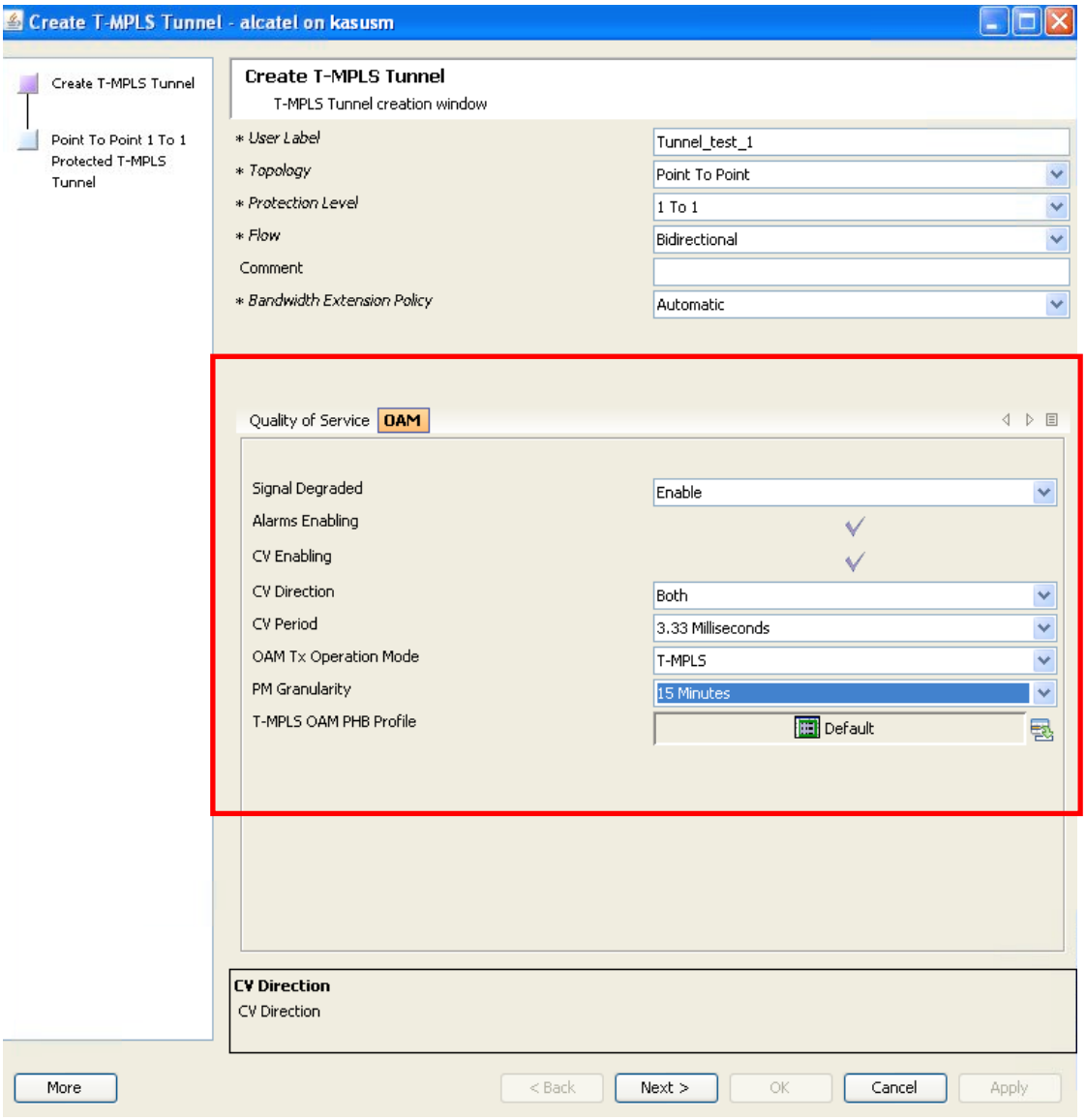


Figure 2.8: Tunnel/LSP OAM configuration

As shown in Figure 2.9, the tunnel has been configured with 1+1 protection. The active tunnel has the status "Working", while the standby tunnel has the status "Protecting".

User Label	Working S...	Last Action	Configuration State	Availability...	Administr...	Flow	Protection...	Topology	Protection...
Tunnel_test_1	Normal	Implementation	Implemented	Normal	On	Bidirectional	1 To 1	Point To Point	Working
Tunnel_test_1_spare	Normal	Implementation	Implemented	Normal	On	Bidirectional	1 To 1	Point To Point	Protecting

Figure 2.9: Tunnel with 1+1 protection

The tunnel configuration can also be verified via the equipment manager or craft terminal. This makes it possible to extract some extra information. Figure 2.10 shows the configuration of one end of the tunnel.

The end of a tunnel can be “head/tail” or “transit”. The end of the tunnel is configured as “head/tail” (as in Figure 2.10 – see the **Role** field) if it is one of the termination points of the tunnel and it is configured as “transit” when the node is only switching labels between segments of the same tunnel.

As shown in the configuration interface (Figure 2.10), the labels used for this tunnel are 18 and 19. This window also allows the configuration of OAM parameters such as CV period via the craft terminal.

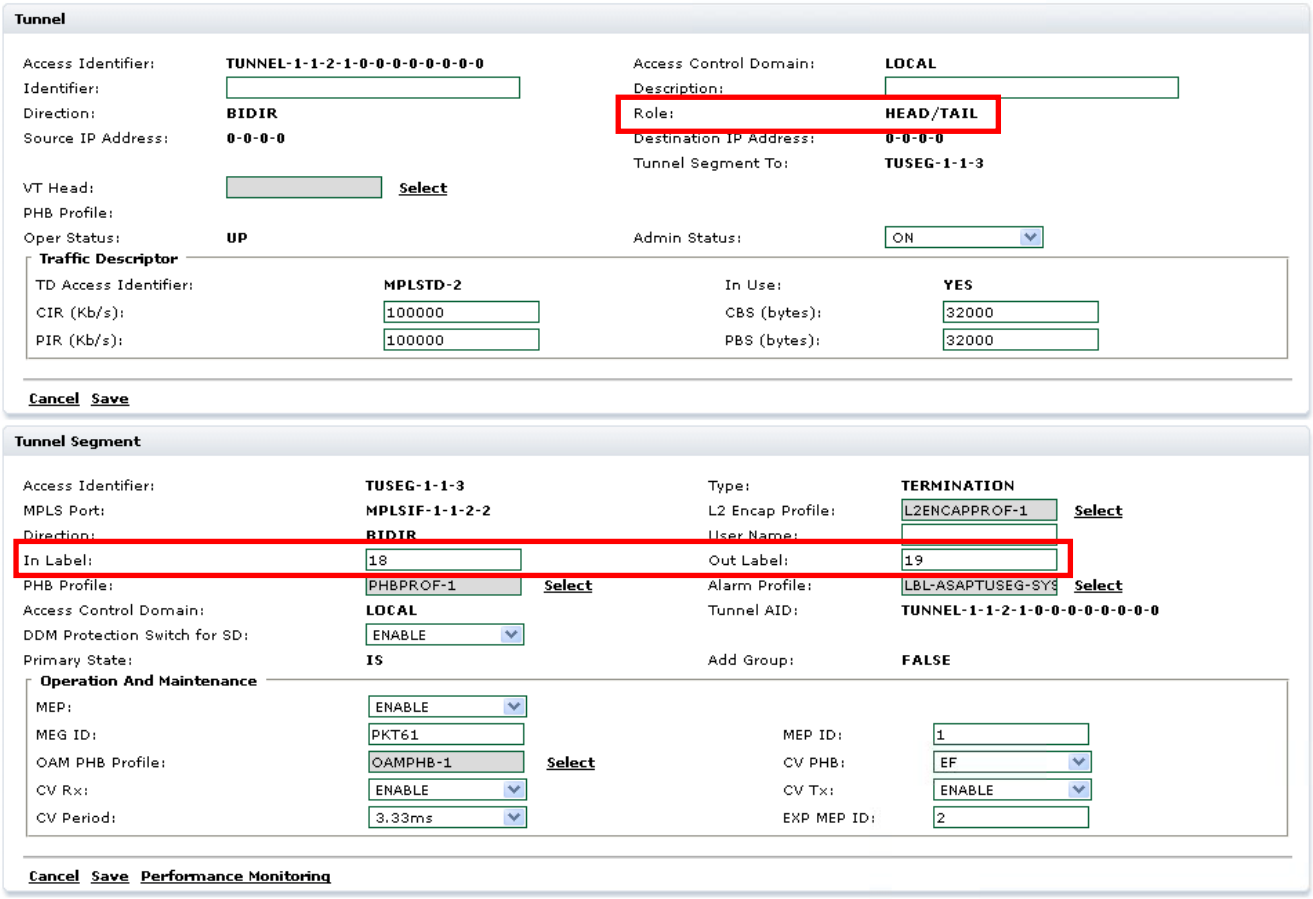


Figure 2.10: Tunnel/LSP configuration via craft terminal

Figure 2.11 shows the tunnel configuration with 1+1 protection, as represented by the Alcatel-Lucent NMS. (The capture is intended to illustrate how an NMS provides graphical representation of the different components (LSPs, PWs and so on) compared to a CLI, which does not offer the operator this option; the detail, in this instance, is not critical to an understanding of the figure.)

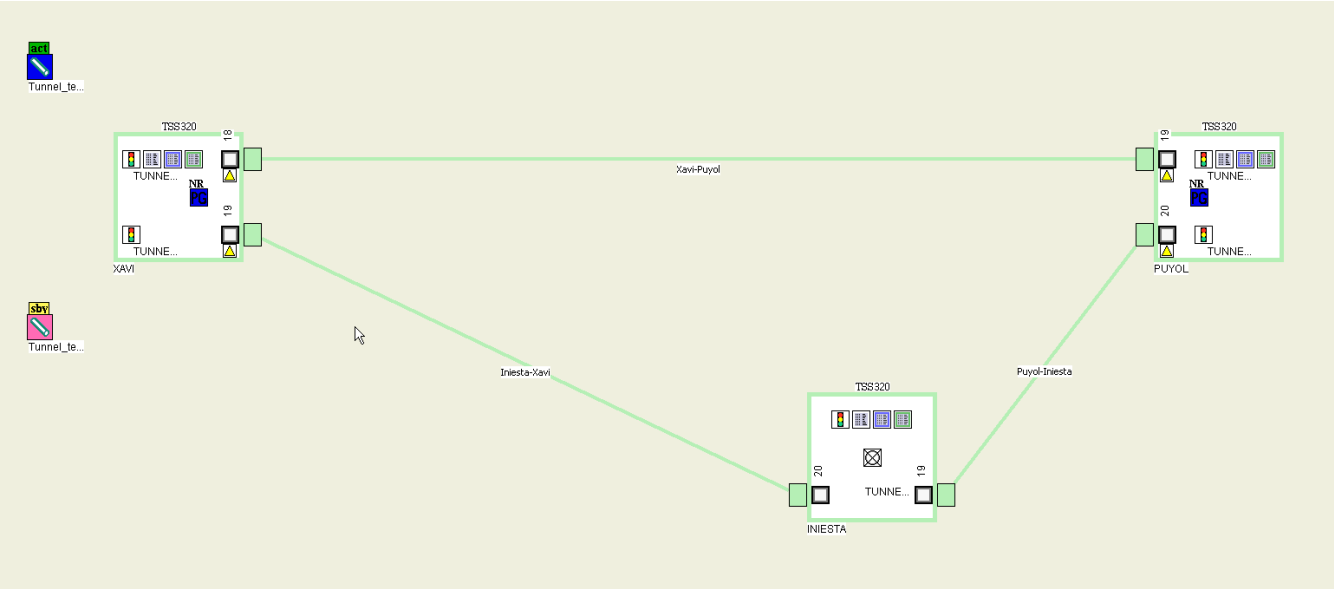


Figure 2.11: Tunnel representation from NMS

2.5.4.3 *Pseudowire Configuration*

In MPLS-TP protocol, pseudowires (PWs) are carried by LSPs. Several PWs can be mapped into the same LSP. The equipment under test does not support PW protection. However, the tunnel carrying the PW can be protected. The total traffic carried by the sum of the PWs cannot exceed the bandwidth allocated to the LSP, unless this traffic is configured as Best Effort, in which case oversubscription of the total traffic of the LSP is allowed.

Figure 2.12 describes the PW configuration. The PW can be configured to carry Ethernet, Asynchronous Transfer Mode (ATM) and Circuit Emulation Services. For the purpose of this testing, Ethernet service is chosen.

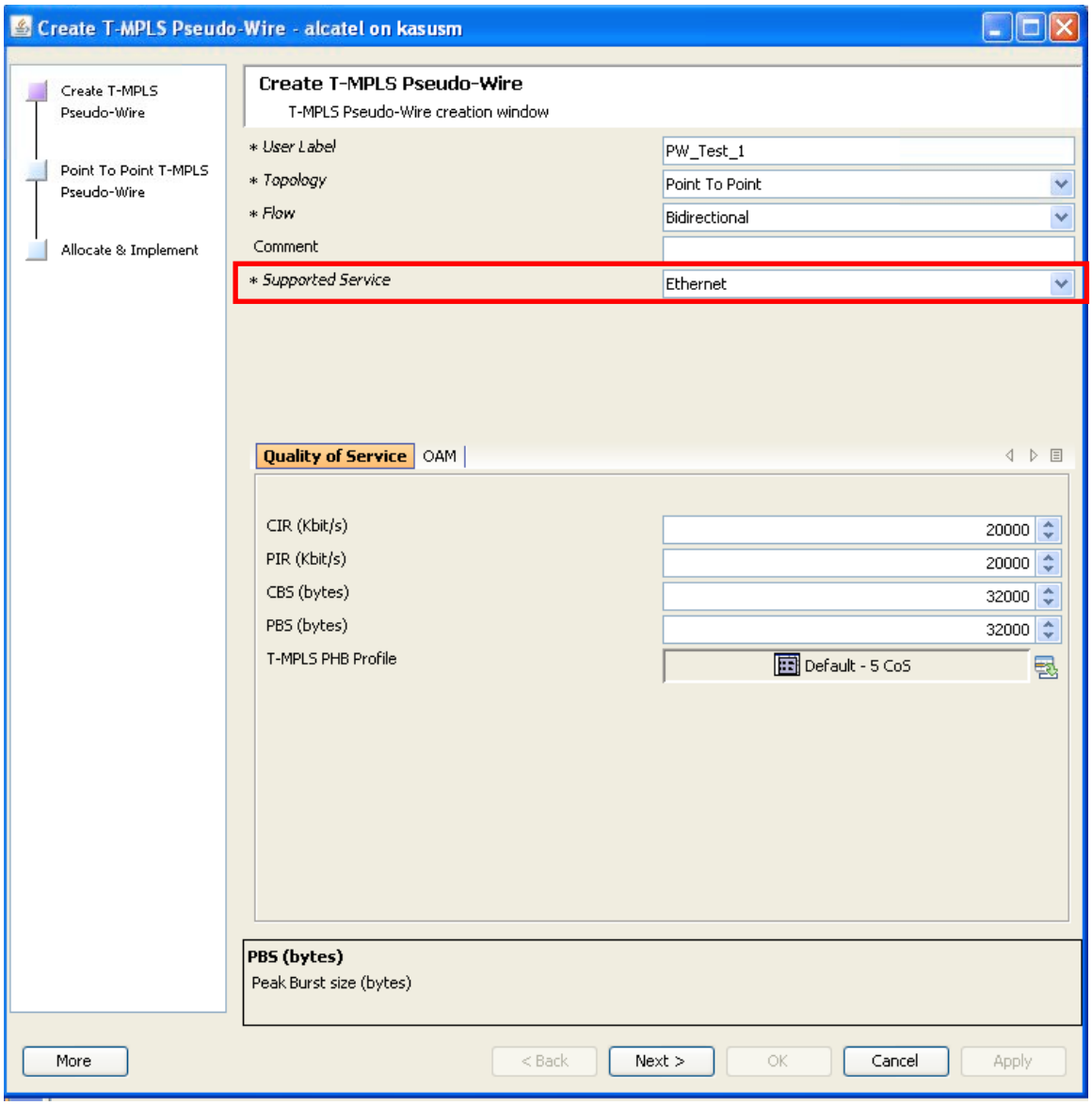


Figure 2.12: PW configuration window in NMS

As with the tunnel, it is possible to configure the PW via the craft terminal. As shown in Figure 2.13, this makes it possible to identify which labels are used for this specific PW, as well as the traffic parameters defining the allocated bandwidth.

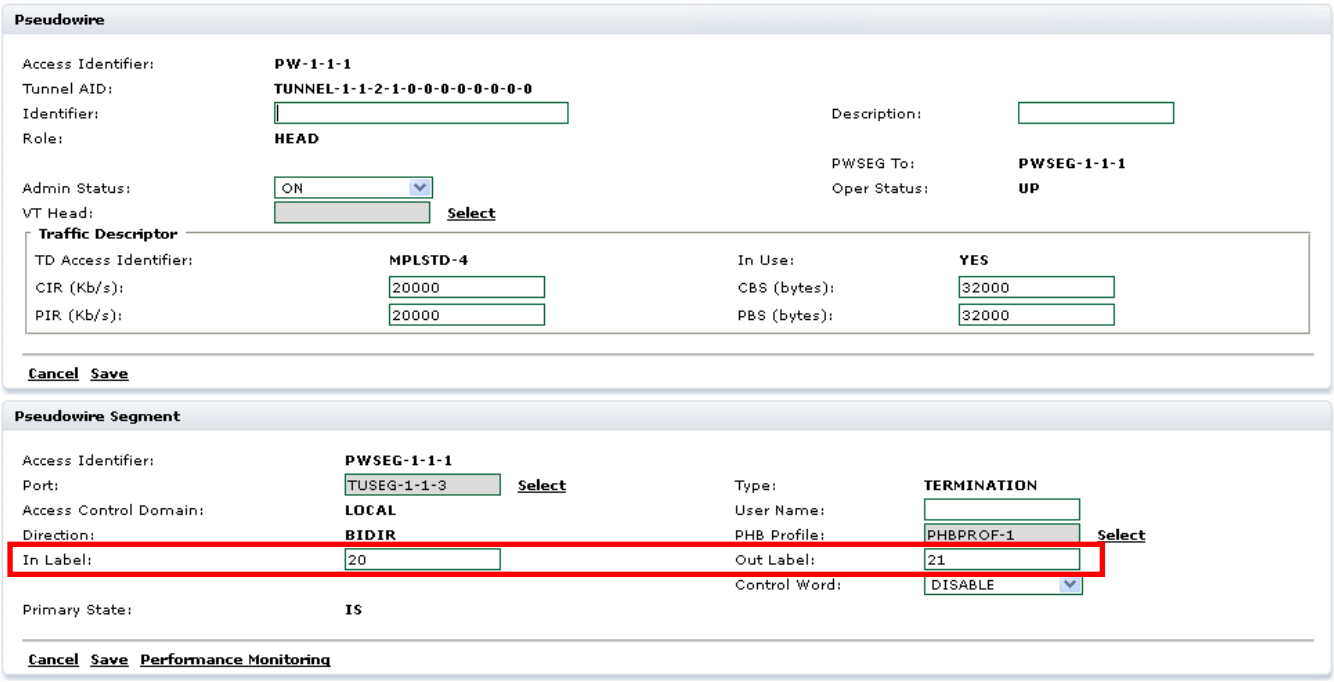


Figure 2.13: First end of the PW configured via craft terminal

The PW can be graphically represented in the NMS. Figure 2.14 shows how the tunnel named “Tunnel_test_1” carried PW called “PW_Test_1”.

The input and output labels can be configured by the operator or can be automatically chosen by the NMS itself.

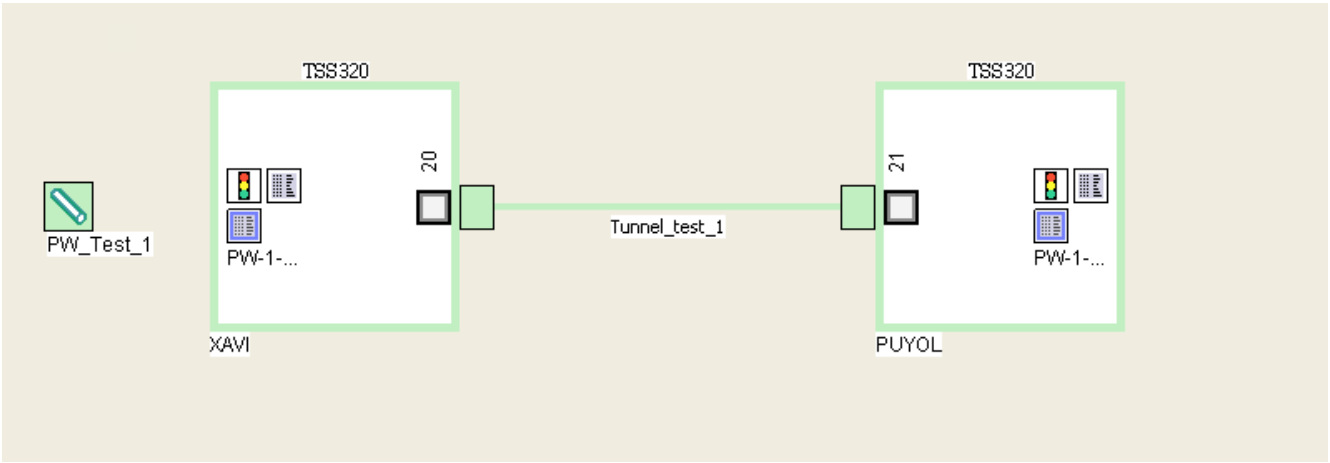


Figure 2.14: PW representation in NMS

2.5.5 Test Results

The result of this part of the testing was positive. It was possible to configure sections, LSPs and PWs. The configuration can be done via the NMS or the craft terminal. However, the latter is more time-consuming as it requires configuration of each node individually, while the NMS allows end-to-end configuration. The NMS configuration is also time-consuming due to the various configuration steps; however, it simplifies the process compared to the craft terminal.

During the testing process it was discovered that pseudowires cannot be individually protected. However, the tunnel carrying the PWs can be protected. This means that in the event of a failure at the LSP or physical layer, all PWs inside a tunnel would be rerouted. The MPLS-TP protocol defined PW 1+1 protection but this feature was not available in the equipment tested.

2.5.6 Test Conclusions

Configuration of sections, LSPs and PWs via NMS is time-consuming but allows the process to be easily understood. An advanced operator used to working via the CLI might find the NMS cumbersome. However, it definitely has some advantages for less experienced operators. For example, the NMS allows LSP and PW configuration without the need for a deep understanding of the underlying technology. It would have been interesting for the purpose of the test to compare the NMS with CLI configuration. However, for the reasons already explained this was not possible.

2.6 Services

2.6.1 Test Objective

The objective of the test was to verify the service configuration as defined by the Metro Ethernet Forum (MEF), to verify that the service works as expected and to obtain experience with the configuration of services in the Alcatel-Lucent platform.

2.6.2 Technology Briefing

MEF 6.1 specifies three different service types: E-Line, E-LAN and E-Tree. For each service type there are two different services depending on whether it is port-based or VLAN-based. In the port-based type the service is terminated in a port that is not shared with other services. In the VLAN-based service the port can be shared and the VLAN ID is used to identify the service. Figure 2.15 shows the three service types. For E-LAN and E-Tree services a Virtual Bridge (VB) function is needed. This function is in charge of doing Media Access Control (MAC) learning and updating the Forwarding Information Base (FIB) table. This function is present in all the nodes for an E-LAN service and only in the hub for an E-Tree service. (Further information about E-Line, E-LAN and E-Tree service types is provided in [DJ1.1.1 and DJ1.1.2].)

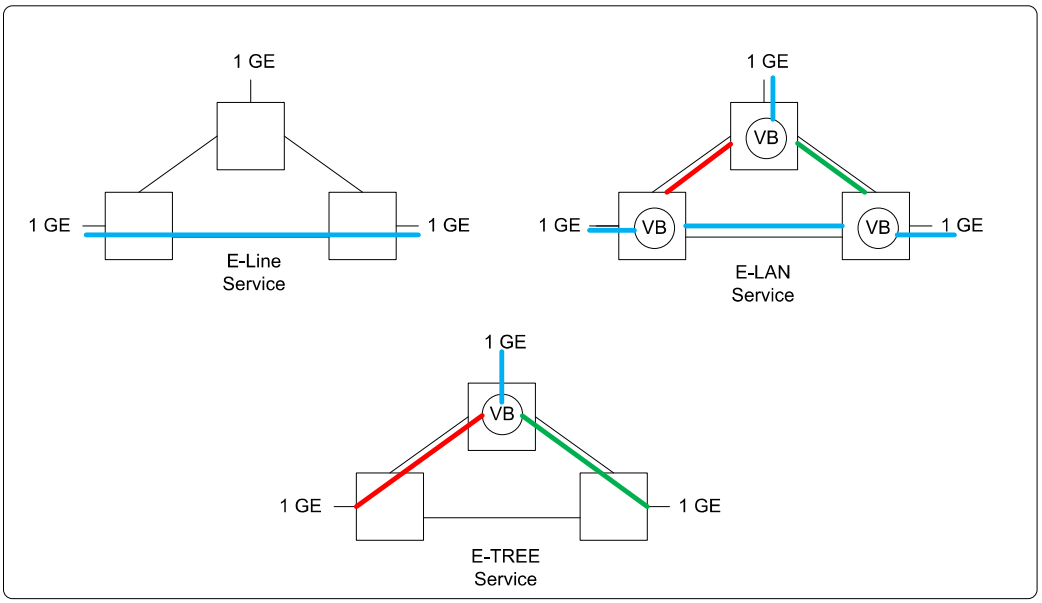


Figure 2.15: MEF service types

2.6.3 Test Setup

The test setup is the same as described in Section 2.4. The EXFO tester is used to verify the configuration and that the service is correctly working.

2.6.4 Test Description

2.6.4.1 *Ethernet Private Line Service*

The Ethernet Private Line (EPL) service is a transparent service. It is part of the E-Line service type as defined by the Metro Ethernet Forum specification MEF 6.1.1. In this service no filtering takes place and all frames are carried transparently. Figure 2.16 shows the EPL service configuration window. Notice that the **Resource sharing** field is set to “Private UNI”. This means that the resources in the specific client interface are fully allocated to the service and that service multiplexing is not allowed. (Service multiplexing allows various services in the same interface. Each service can be represented by one or a set of VLANs. In the case of Private UNI, only one service is allowed.

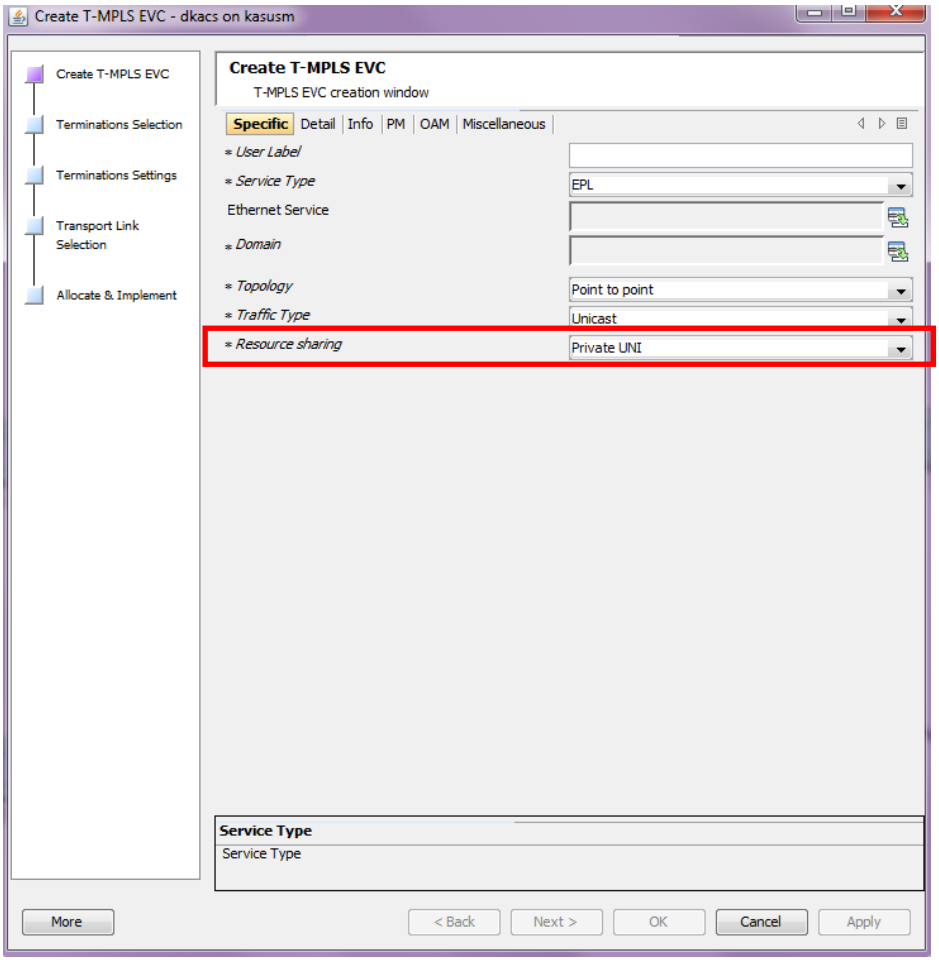


Figure 2.16: EPL service configuration window

Configuration via the craft terminal is shown in Figure 2.17 below. For each Ethernet Virtual Connection (EVC) termination point of the service it is necessary to create two flows from the User Network Interface (UNI) to NNI interface and two flows from the NNI to the UNI interface. In the configuration of these flows the filtering and the QoS parameters can be specified. In this example there is no filtering and the bandwidth allocated is the full bandwidth of the client interface.

	XC Id	From	To	Service Label	Topology
☐	EVC TERMINATION between LINK PW-1-1-1 and the edge PORT r1sr1sl2/ETHLocPort#3#1				
☐	2	ETSInFlow#1	ETSOutFlow#1	1354 BM#1	UNI-NNI
☐	3	ETSInFlow#2	ETSOutFlow#2	1354 BM#2	NNI-UNI

Figure 2.17: Traffic flows for one of the EVCs termination points configured via craft terminal

Once the service is configured, the tester is connected to one of the end points to verify that the traffic is flowing. A physical loop is placed at the other end of the service. Figure 2.18 shows that the service was running after the configuration was applied: the alarms are green, which means that they have been cleared.

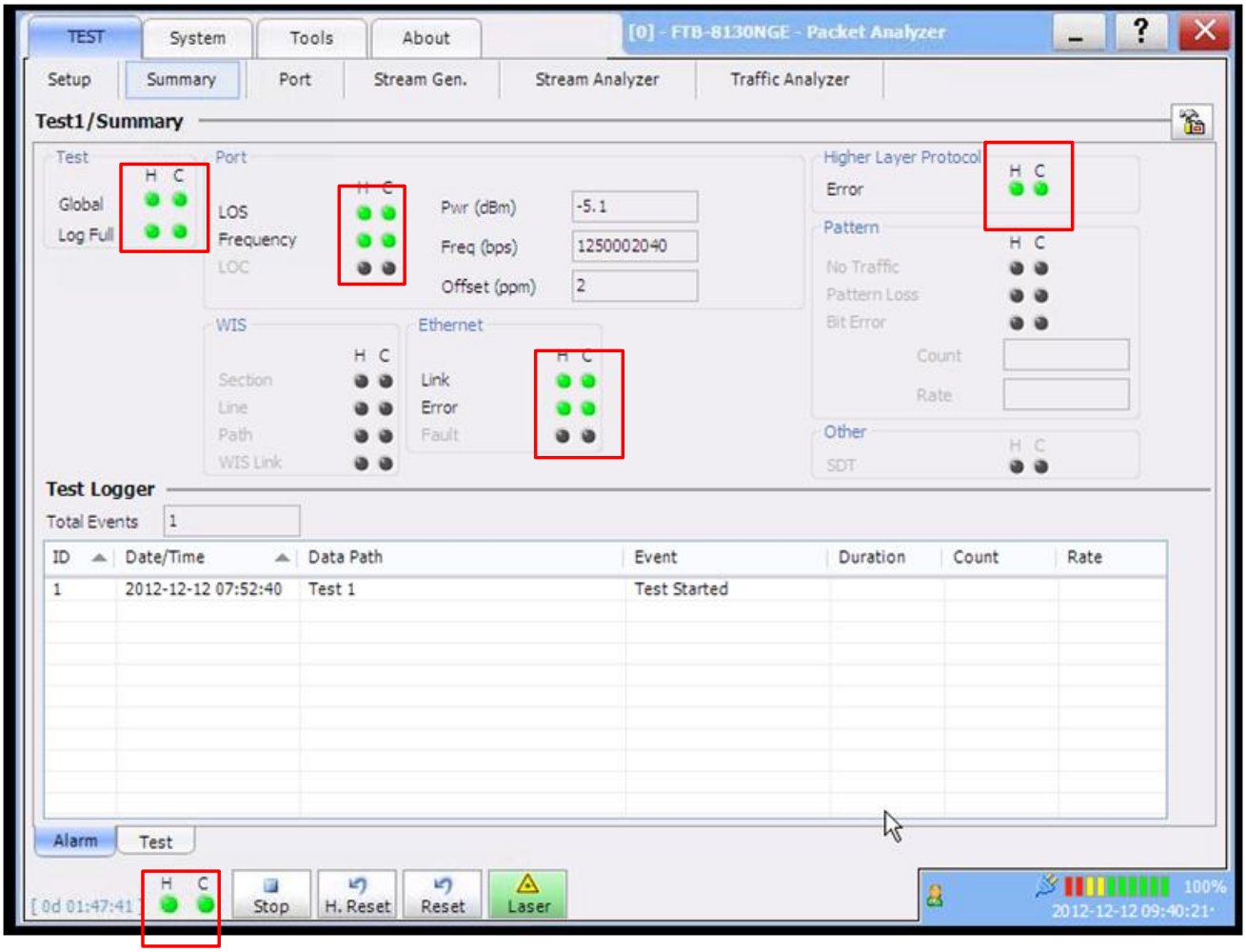


Figure 2.18: EPL service verification with EXFO tester

2.6.4.2 *Ethernet Virtual Private Line Service*

An Ethernet Virtual Private Line (EVPL) service is a point-to-point Ethernet service that carries a specific VLAN or group of VLANs between two points in the network. This is also an E-Line service according to the MEF specifications. In this case, the physical interface resources are shared between different services terminating in the same interface. This means that service multiplexing at the UNI is allowed. Services starting at the same interface can be terminated at different ones. The window in Figure 2.19 shows the configuration for an EVLP service. Notice that in this case the **Resource sharing** field is set to "Full". When configuring EVPL services it is necessary to specify the VLAN or group of VLANs that the service is carrying. This is done by defining a set of filters that are associated with the traffic flows of the interface.

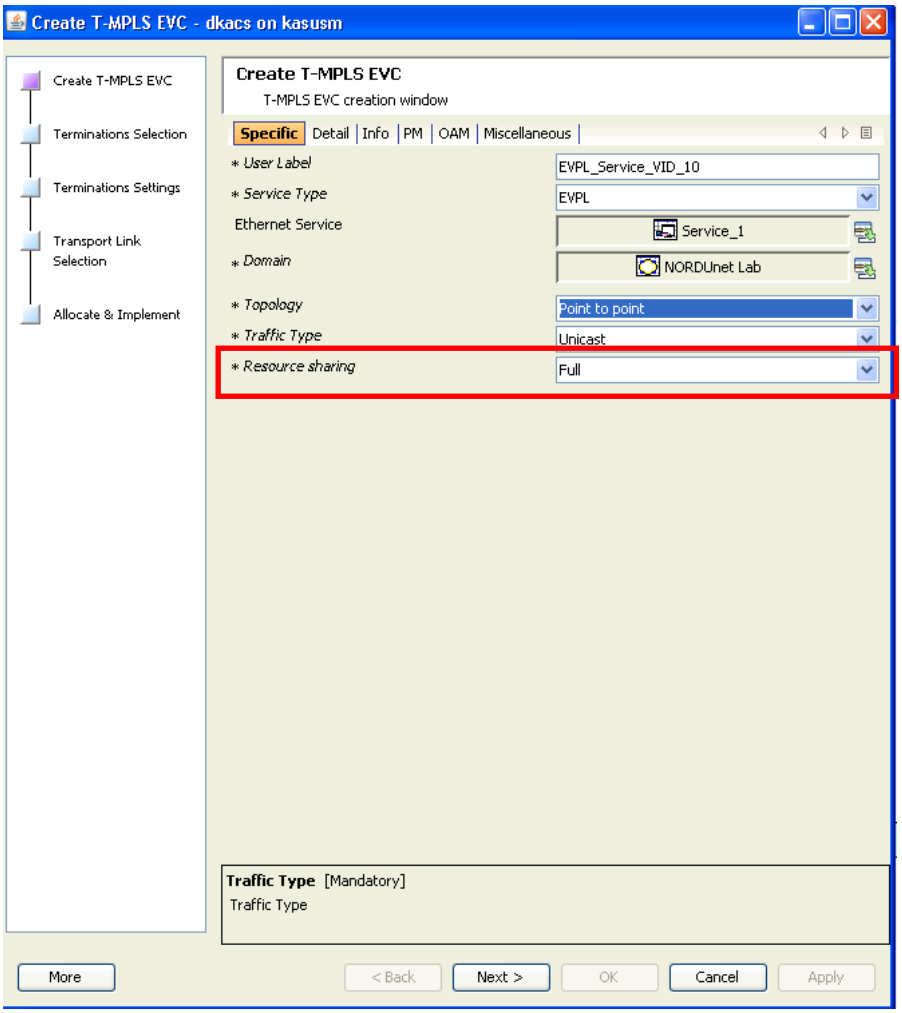


Figure 2.19: EVPL service configuration window

The tester is connected to the equipment and configured to send traffic tagged with VLAN ID 30. The service is not configured at this point and, as shown in Figure 2.20, the tester is not detecting any traffic. A physical loop is placed at the other termination point of the service.

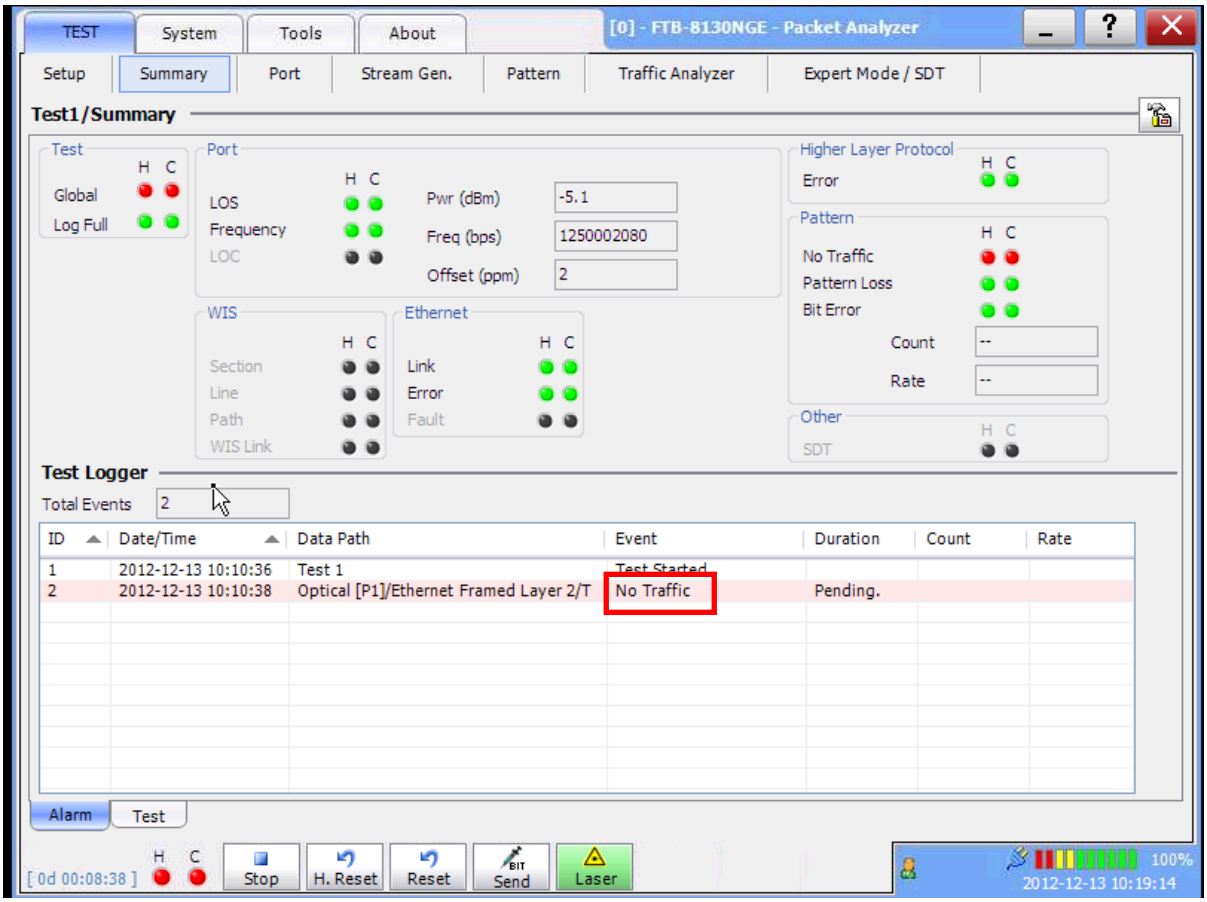


Figure 2.20: No traffic in VLAN 30 prior to service creation

Now, the EVPL service for VLAN 30 is configured as shown in Figure 2.21.

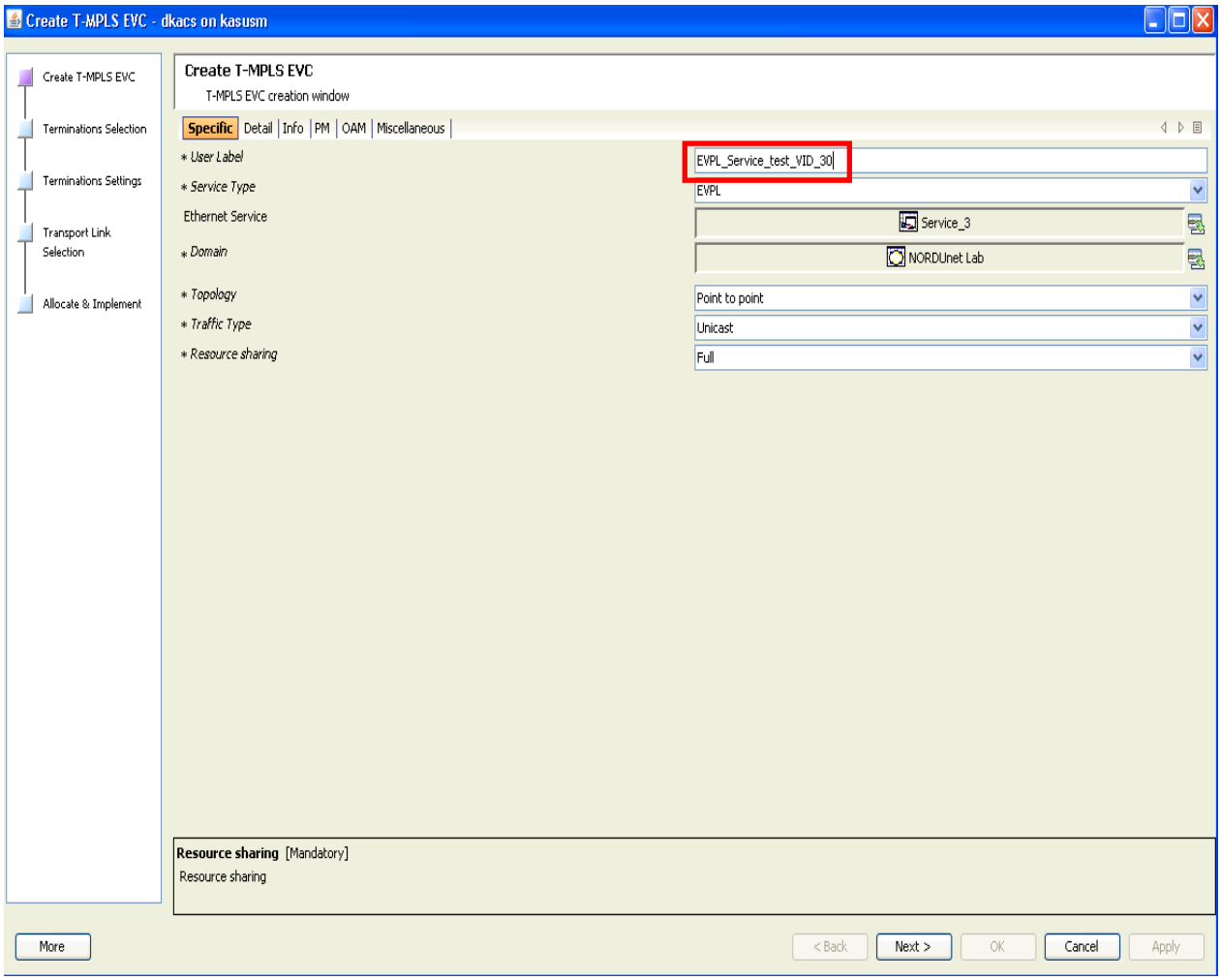


Figure 2.21: Creation of EVPL for VLAN 30

Once the creation of the service has taken place, the traffic starts flowing and the tester starts detecting frames, as shown in Figure 2.22. The detected frames are tagged with VLAN ID 30. If the tester was configured to send frames tagged with VID=10 the traffic would not flow, and the tester would report a “no traffic” event. This is because the configured EVC associated with the EVPL service only accepts frames with VID=30. Notice that the tester shows an alarm event that lasted 16:50 sec. This is the time that it took the operator to apply the configuration. The alarms (highlighted by the red boxes) are green, indicating that they were all cleared.

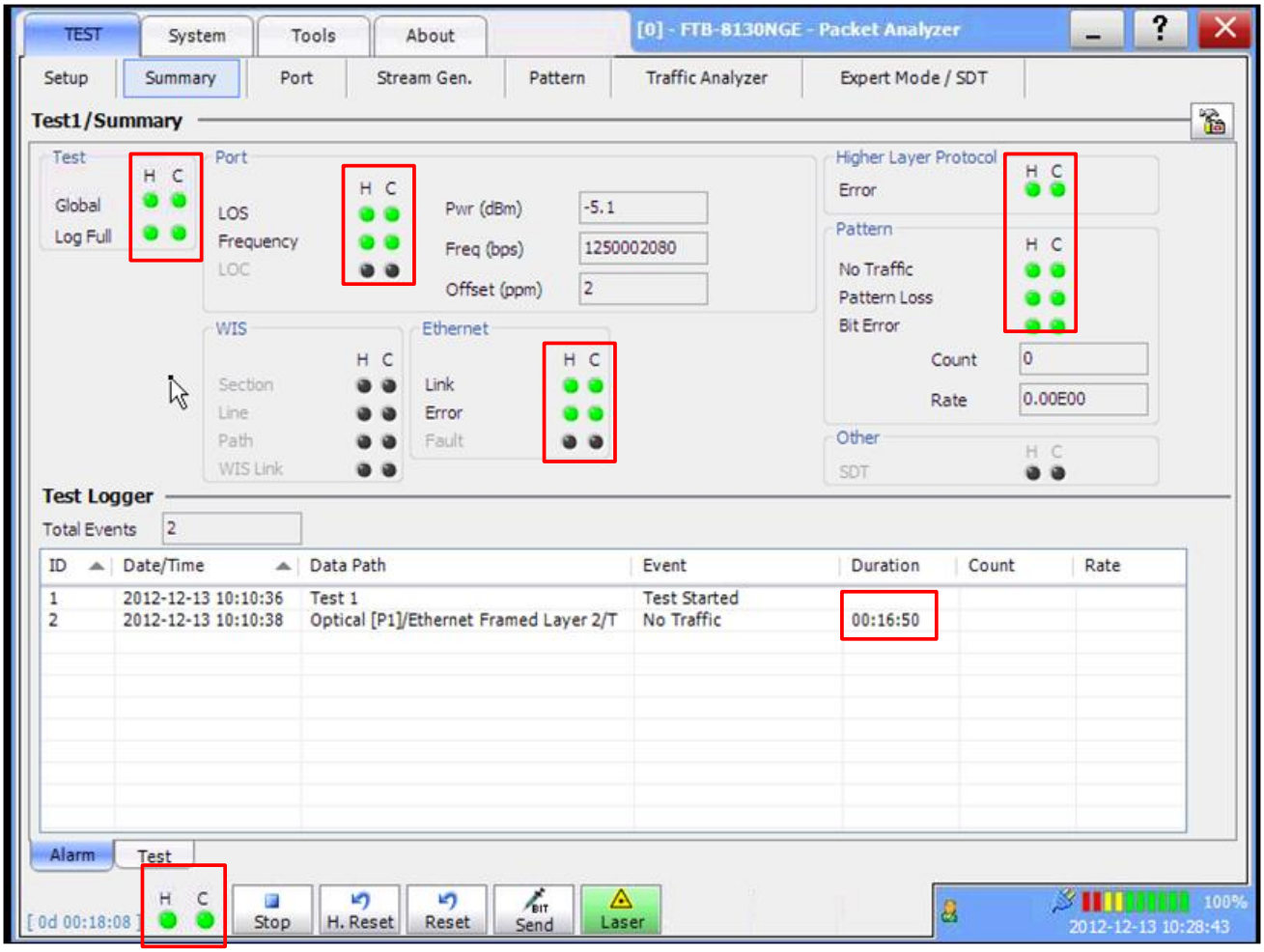


Figure 2.22: Tester detecting traffic on VLAN 30

Figure 2.23 shows the configuration of the traffic flows via the craft terminal. Although the configuration is done via the NMS, it is also possible to configure the equipment via the craft terminal. This process is more time-consuming. However, it is a good way to verify the configuration.

	XC Id	From	To	Service Label	Topology
☐	EVC TERMINATION between LINK PW-1-1-2 and the edge PORT r1sr1sl2/ETHLocPort#3#1				
☐	2	ETSInFlow#1	ETSOutFlow#1	1354 BM#1	NNI-UNI
☐	3	ETSInFlow#2	ETSOutFlow#2	1354 BM#2	UNI-NNI
☐	EVC TERMINATION between LINK PW-1-1-4 and the edge PORT r1sr1sl2/ETHLocPort#3#1				
☐	4	ETSInFlow#3	ETSOutFlow#3	1354 BM#3	UNI-NNI
☐	5	ETSInFlow#4	ETSOutFlow#4	1354 BM#4	NNI-UNI
☐	EVC TERMINATION between LINK PW-1-1-5 and the edge PORT r1sr1sl2/ETHLocPort#3#1				
☐	6	ETSInFlow#5	ETSOutFlow#5	1354 BM#5	UNI-NNI
☐	7	ETSInFlow#6	ETSOutFlow#6	1354 BM#6	NNI-UNI

☐	EVC TERMINATION between LINK PW-1-1-3 and the edge PORT r1sr1sl2/ETHLocPort#4#1				
☐	4	ETSInFlow#3	ETSOutFlow#3	1354 BM#3	NNI-UNI
☐	5	ETSInFlow#4	ETSOutFlow#4	1354 BM#4	UNI-NNI
☐	EVC TERMINATION between LINK PW-1-1-5 and the edge PORT r1sr1sl2/ETHLocPort#4#1				
☐	6	ETSInFlow#5	ETSOutFlow#5	1354 BM#5	UNI-NNI
☐	7	ETSInFlow#6	ETSOutFlow#6	1354 BM#6	NNI-UNI
☐	EVC TERMINATION between LINK PW-1-1-6 and the edge PORT r1sr1sl2/ETHLocPort#4#1				
☐	8	ETSInFlow#7	ETSOutFlow#7	1354 BM#7	UNI-NNI
☐	9	ETSInFlow#8	ETSOutFlow#8	1354 BM#8	NNI-UNI

Figure 2.23: Outgoing and incoming traffic flows at the EVC terminations via craft terminal

For EVPL services the traffic flows have a traffic profile and a filter profile associated. The traffic profile defines the bandwidth associated with the service and the QoS parameters for the specific service. The filter profile defines the conditions for traffic filtering. In this case the filter was defined to filter frames with VLAN ID 30.

2.6.4.3 *Ethernet Virtual Private LAN Service*

The MEF defines two types of E-LAN services in specification MEF 6.1.1: Ethernet Private LAN (EP-LAN), where the resources for each termination point are not shared, and Ethernet Virtual Private LAN (EVP-LAN), where the resources in each termination point are shared, meaning that another service can be configured in the same interface. In this case, only the EVP-LAN service was tested.

An EVP-LAN service is a multipoint-to-multipoint virtual service, meaning that the service can be terminated at various points of the network. The service can carry a single VLAN or a group of VLANs. In order to be able to configure an EVP-LAN service it is necessary to have the corresponding LSPs and PWs in place. Figure 2.24 shows the infrastructure used for this test. The necessary LSPs and PWs were configured between the three nodes as the EVP-LAN service will have a termination point at each network element.

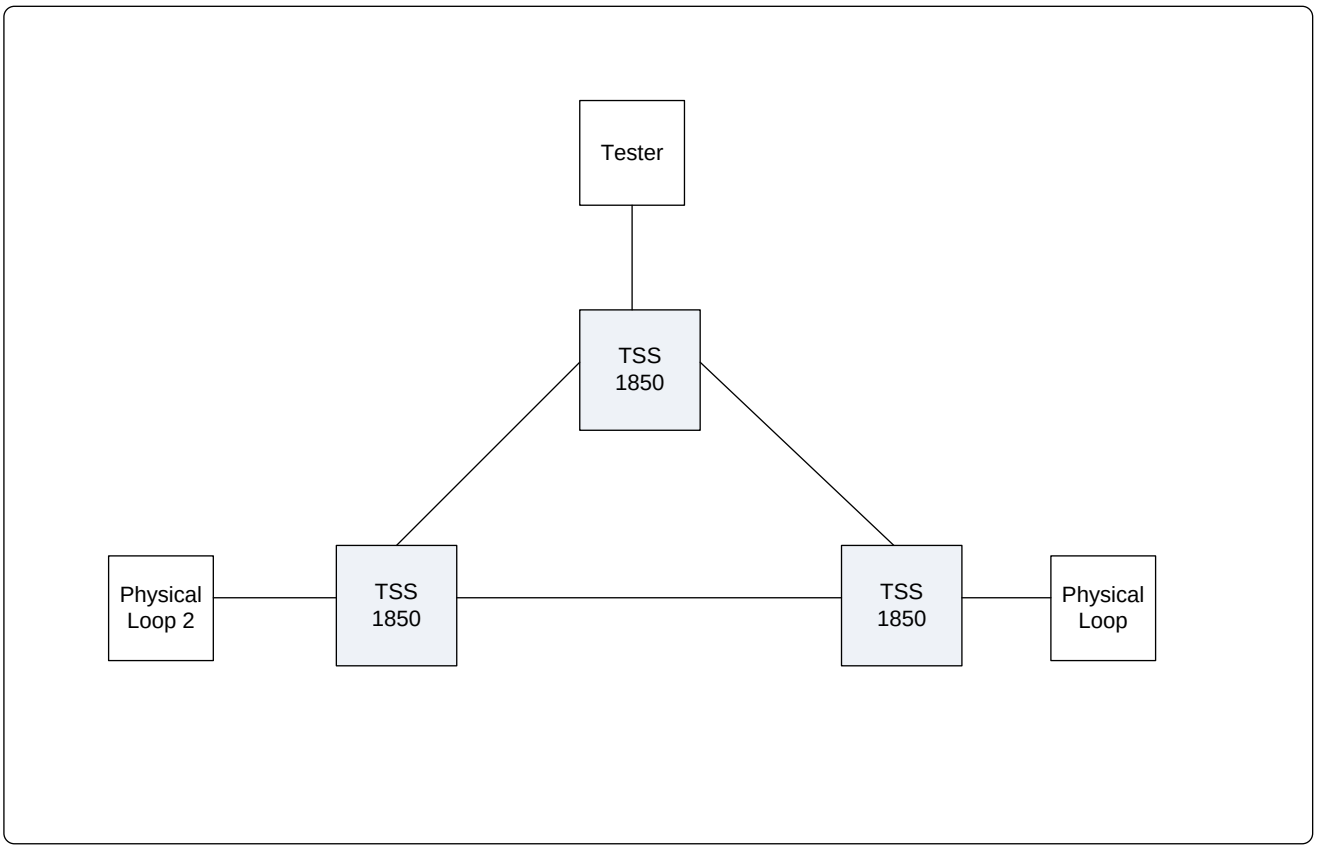


Figure 2.24: EVP-LAN service configuration

Figure 2.25 shows the configuration window of the NMS. In this case an any-to-any topology must be chosen and resource sharing set to “Full” to allow service multiplexing at the UNI. During the configuration process it is also necessary to define the traffic profile and the filtering profile, as for the EVPL service.

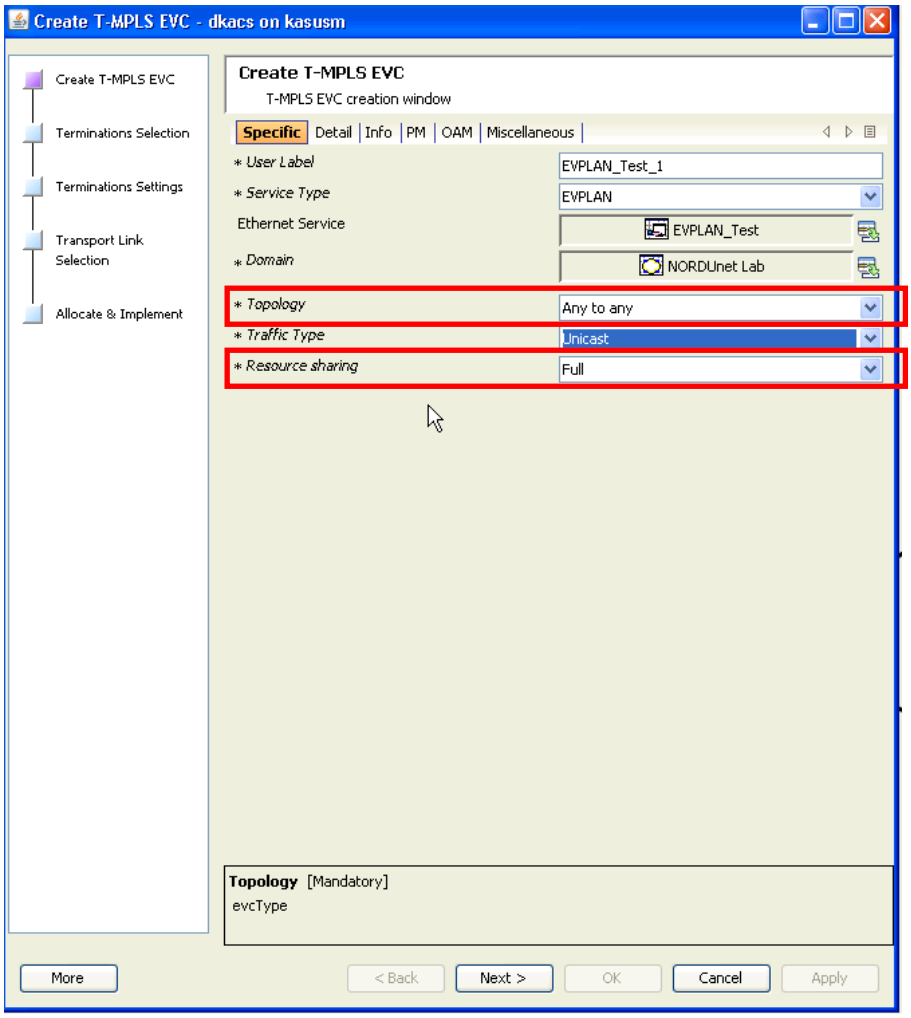
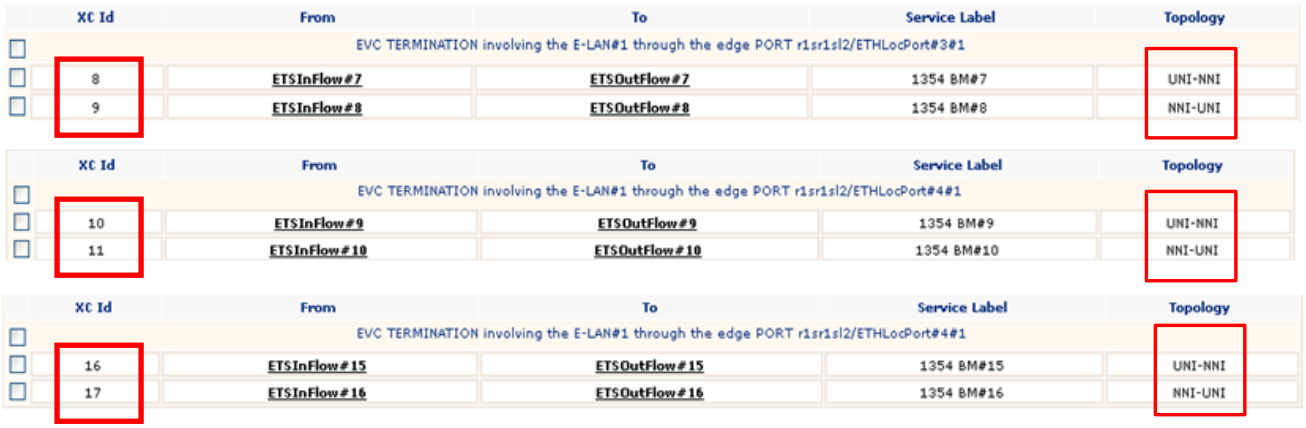


Figure 2.25: EVP-LAN configuration window

Figure 2.26 shows the traffic flows for each end of the service. Again, each traffic flow has a traffic profile and filter profile that defines the bandwidth and the type of frames carried by this service. As shown in the figure, each EVC termination has two flows: the first EVC has flows 8 and 9, the second EVC has flows 10 and 11 and the last EVC has flows 16 and 17. One flow defines the characteristics of the traffic from the UNI to NNI interface while the other flow defines the traffic characteristics in the other direction, meaning from the NNI to the UNI.



XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#3#1				
8	ETSInFlow#7	ETSOutFlow#7	1354 BM#7	UNI-NNI
9	ETSInFlow#8	ETSOutFlow#8	1354 BM#8	NNI-UNI

XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#4#1				
10	ETSInFlow#9	ETSOutFlow#9	1354 BM#9	UNI-NNI
11	ETSInFlow#10	ETSOutFlow#10	1354 BM#10	NNI-UNI

XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#4#1				
16	ETSInFlow#15	ETSOutFlow#15	1354 BM#15	UNI-NNI
17	ETSInFlow#16	ETSOutFlow#16	1354 BM#16	NNI-UNI

Figure 2.26: Traffic flows for EVP-LAN service

Verification of this service was not possible, as the necessary equipment for this purpose was not available. However, according to the NMS and the craft terminal the service was active and operational and no alarms appeared during the process.

2.6.4.4 *Ethernet Virtual Private Tree Service*

An E-Tree service is based upon a rooted-multipoint EVC. The simplest form of this type of service has one termination point acting as a root and several termination points acting as leaves. Leaf termination points can only exchange data with the root termination point.

The MEF defines two types of E-Tree services: an Ethernet Private Tree (EP-Tree) service, where the resources of the termination points of the service are not shared, and an Ethernet Virtual Private Tree (EVP-Tree) service, where the resources are shared. In this case, EVP-Tree was tested.

This type of service is called a Broadcast TV (BTV) service by Alcatel-Lucent, because it is mainly used for multicast transport of TV traffic. The service is similar to the EVP-LAN service, with the only difference that in this case all traffic flows from a common point in the network known as the root. Figure 2.27 shows the configuration window for EVP-Tree services. In this case the topology is set to “Rooted multipoint” and the **Resources sharing** field is again set to “Full” to allow service multiplexing.

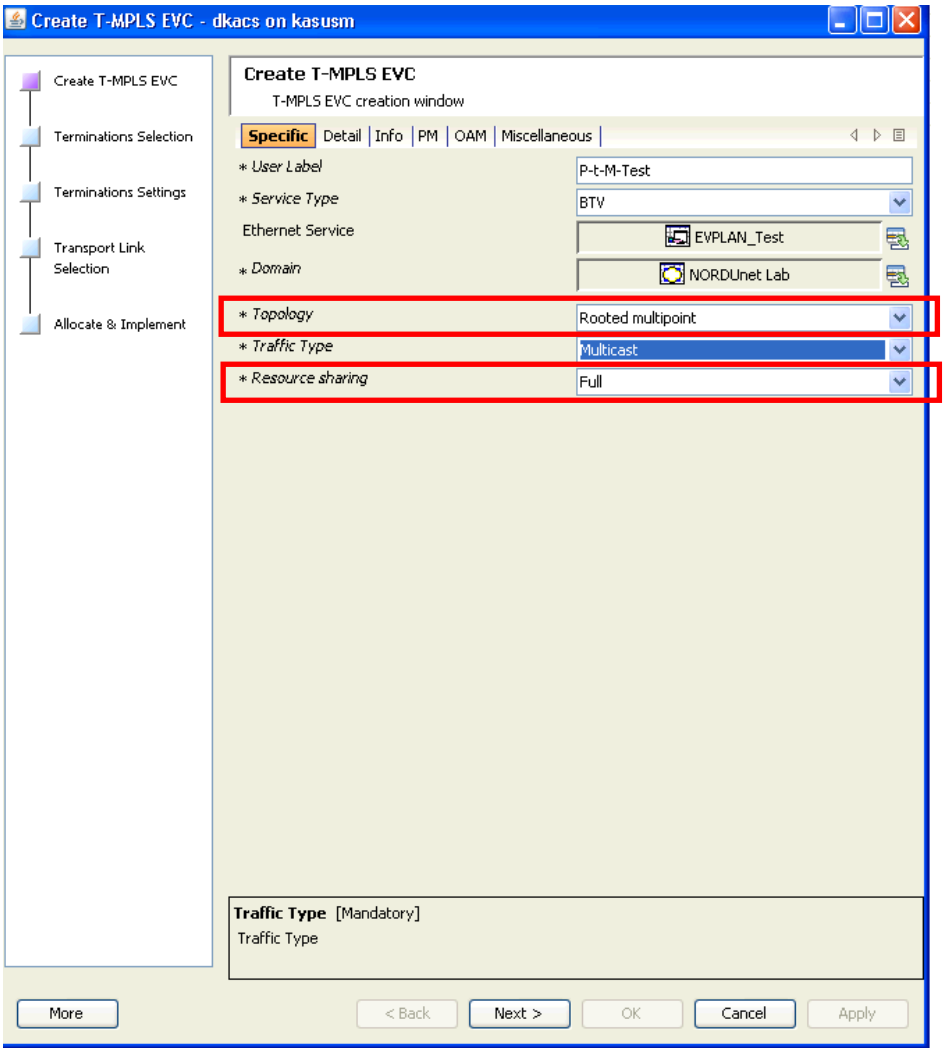


Figure 2.27: EVP-Tree service configuration window

During the configuration it is necessary to choose the roles of each termination point of the service. As shown in Figure 2.28, one of the end points is configured as “Root” while the rest are configured as “Leaf”.

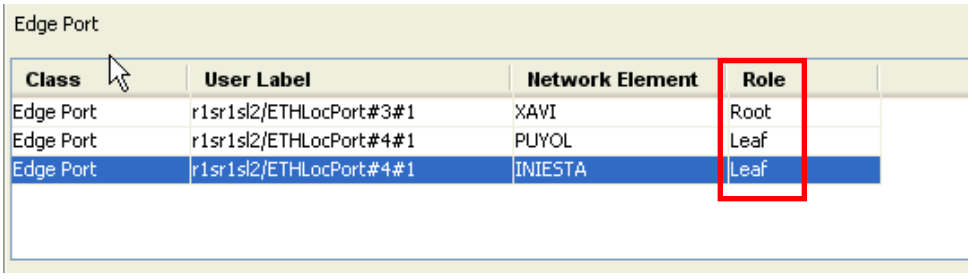
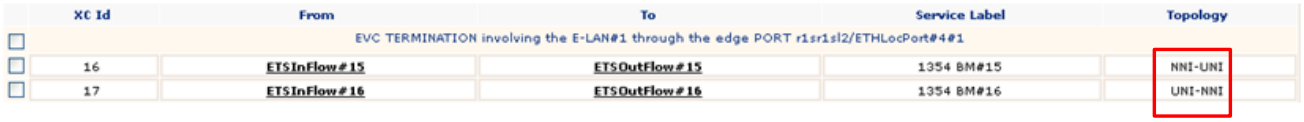


Figure 2.28: Port role configuration



XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#4#1				
16	ETSInFlow#15	ETSOutFlow#15	1354 BM#15	NNI-UNI
17	ETSInFlow#16	ETSOutFlow#16	1354 BM#16	UNI-NNI

Figure 2.29 shows the configured flows for each termination point. Again, each termination point has two flows; one flow defined the characteristics in the UNI to NNI direction, while the second flow defined the traffic in the opposite direction.

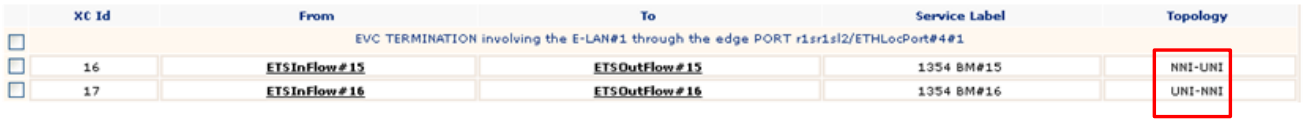
Root:

XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#3#1				
8	ETSInFlow#7	ETSOutFlow#7	1354 BM#7	NNI-UNI
9	ETSInFlow#8	ETSOutFlow#8	1354 BM#8	UNI-NNI

Leaf:

XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#4#1				
10	ETSInFlow#9	ETSOutFlow#9	1354 BM#9	NNI-UNI
11	ETSInFlow#10	ETSOutFlow#10	1354 BM#10	UNI-NNI

Leaf:



XC Id	From	To	Service Label	Topology
EVC TERMINATION involving the E-LAN#1 through the edge PORT r1sr1sl2/ETHLocPort#4#1				
16	ETSInFlow#15	ETSOutFlow#15	1354 BM#15	NNI-UNI
17	ETSInFlow#16	ETSOutFlow#16	1354 BM#16	UNI-NNI

Figure 2.29: Traffic flows configuration

As for the EVP-LAN service, the configuration could not be verified with real traffic, as the necessary equipment for this purpose was not available. However, the configuration could be verified via the NMS and the craft terminal and no alarms were present during the configuration.

2.6.5 Test Results

The testing of the different types of service was satisfactory. It was possible to configure EPL, EVPL, EVP-LAN and EVP-Tree services. It was also possible to send and verify traffic flows for point-to-point services. However, this was not possible for the other types of services as the additional test terminals required for this purpose were not available. MPLS-TP is also capable of carrying ATM or SDH traffic. However, these types of interfaces were not available during the testing.

2.6.6 Test Conclusions

T-MPLS/MPLS-TP allows configuration of a broad range of services. It also allows configuration of different type of filtering and QoS parameters, depending on the needs.

Configuration via NMS can be time-consuming, as there are many steps to go through. However, it allows a better understanding of the process and can be an advantage for less experienced operators. However, operators used to working with the CLI might find this process a bit cumbersome. A combination of both methods, or the possibility of choosing between them, is definitely an advantage. However, the CLI option could not be tested, as it was not fully supported.

2.7 OAM Testing

2.7.1 Test Objective

The purpose of this test was to verify the OAM features available in the Alcatel-Lucent 1850 TSS. Unfortunately, in the release used only Connectivity Verification (CV), the alarms Forward Defect Indication (FDI) and Remote Defect Indication (RDI) and single-ended Delay Measurements (DM) are available.

2.7.2 Technology Briefing

OAM traffic is carried at different levels of the MPLS-TP protocol. If the frame is carrying OAM information, an OAM label describes the type of OAM traffic that it is being carried, as shown in Figure 2.30. In this test only OAM at the LSP level was supported. The OAM implementation in this software release is based on the ITU-T Y.1731 standard where CV is used for 1+1 LSP protection.

For more information about OAM please refer to JRA1 Task 1's first GN3 deliverable, DJ1.1.1 [DJ1.1.1].

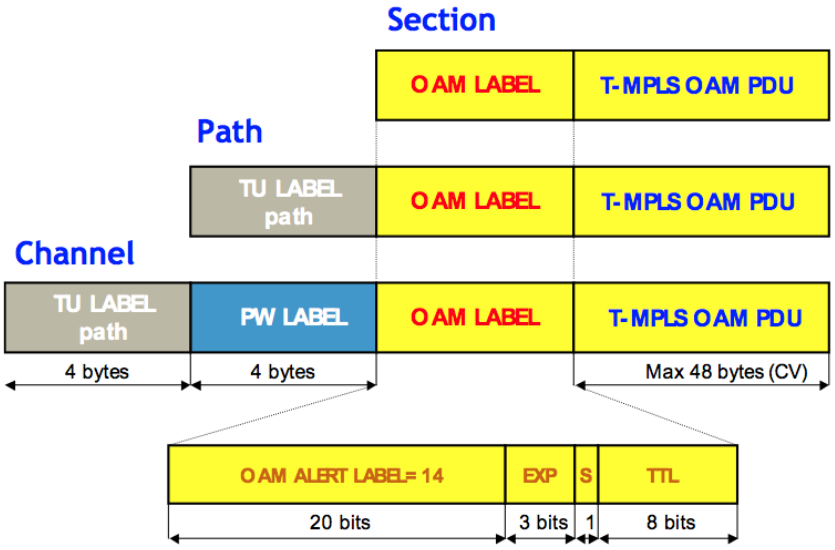


Figure 2.30: OAM MPLS-TP label

2.7.3 Test Setup

The test infrastructure used for this test is the same as described in Section 2.4 above.

2.7.4 Test Description

2.7.4.1 *Connectivity Verification*

The source Maintenance End Point (MEP) sends Connectivity Verification (CV) OAM packets periodically at the configurable rate. The sink MEP monitors the arrival of these CV OAM packets at the configured rate and detects the defect of Loss of Continuity (LOC).

CV is configured during the LSP setup. As shown in Figure 2.31, it can be configured in one or both directions. It is also possible to configure the CV period. In order to achieve 50 ms switching during protection, 3.33 ms is necessary. This means that a CV packet is sent every 3.33 milliseconds.

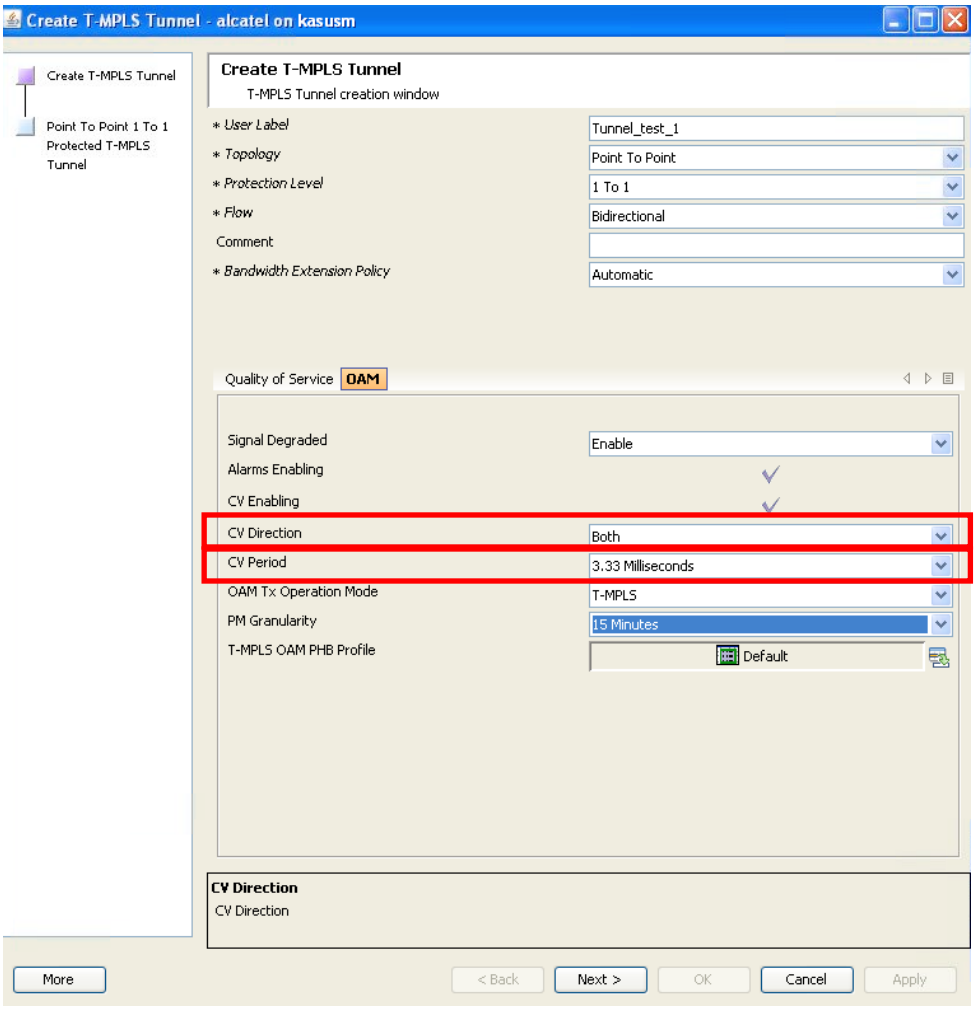


Figure 2.31: CV configuration during LSP setup

During the testing it was not possible to monitor OAM traffic. Therefore the only way to verify CV was to use it in combination with 1+1 protection, as described in Section 2.8.

2.7.4.2 *Forward Defect Indication and Remote Defect Indication*

Forward Defect Indication (FDI) is mainly used to suppress alarms after detection of defect conditions at the server sub-layer. When a server MEP detects Loss of Continuity (LOC) or signal fail, it sets a flag that results in the generation of OAM packets with Alarm Indication Signal (AIS) information that are forwarded in the downstream direction to the sink MEP in the client sub-layer. This allows the suppression of secondary alarms in other layers.

Remote Defect Indication (RDI) is an indicator that is transmitted by a MEP to communicate to its peer MEPs that a signal fail condition exists. When a MEP detects a signal fail condition, it sends RDI to its peer MEP.

Because it was not possible to monitor OAM packets at the MPLS layer, it was not possible to verify FDI and RDI operations. However, it was possible to verify that when a link failed an RDI alarm was raised in the alarm monitoring system.

2.7.4.3 *On-Demand Delay Measurement*

Dual-ended Delay Measurement (DM) can be used proactively and on demand. The originating MEP periodically sends DM packets indicating the timestamp of the transmitting time. The receiver MEP copies the timestamp and sends it back. The originating MEP can then compare and calculate the time delay. The difference between two consecutive measurements gives the delay variation. This tool is used for performance monitoring.

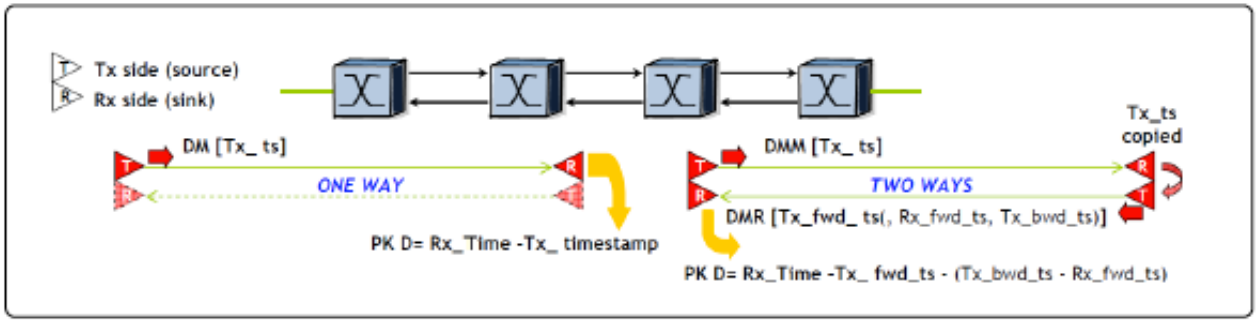


Figure 2.32: Single- and dual-ended Delay Measurement concept

DM was only supported via the craft terminal, as shown in Figure 2.33. It was possible to configure the exact starting point and the duration of the measurement. The tool also allowed the configuration of different frame sizes.

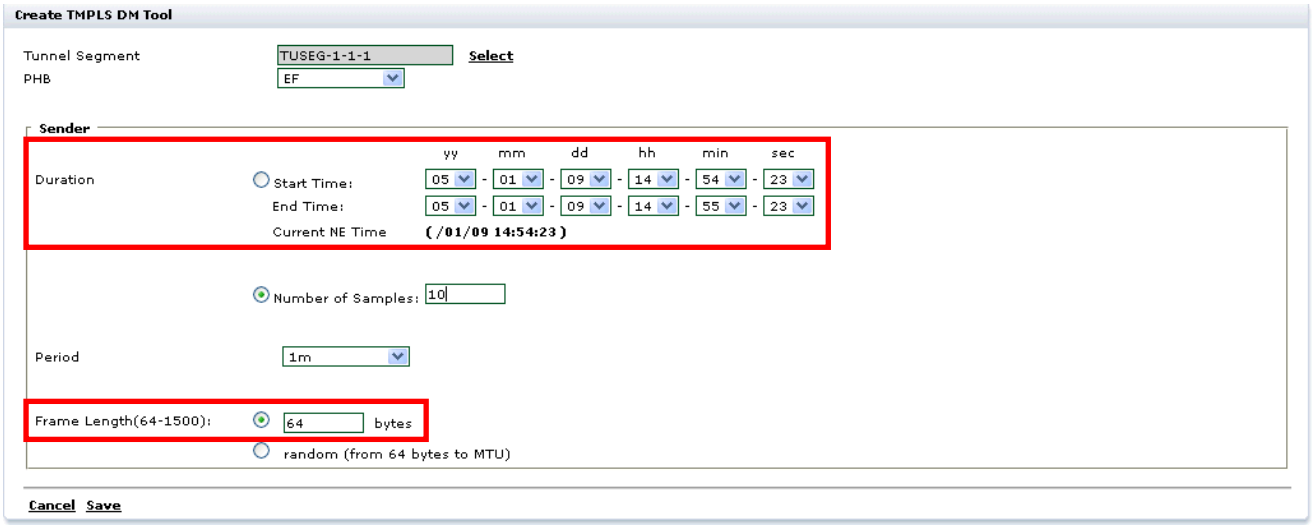


Figure 2.33: DM configuration window via craft terminal

The measurement results were then archived in a file that could be exported to an Excel file, as shown in Figure 2.34.

ENABLE: DM	period=1m	sample-num=10	length=64				
Elapsed Time	Successful DM	TxTimeStampf	RxTimeStampf	TxTimeStampb	RxTimeStampb	Unsuccessful DM	RT Delay
000:00:00:50	1	1357743421:39593829	1357743421:208988489	1357743421:208991925	1357743421:39603228	0	00:000:005:963
000:00:01:50	2	1357743481:39593829	1357743481:208980438	1357743481:208985521	1357743481:39604874	0	00:000:005:962
000:00:02:50	3	1357743541:39593829	1357743541:208972336	1357743541:208975979	1357743541:39603433	0	00:000:005:961
000:00:03:50	4	1357743601:39593829	1357743601:208964234	1357743601:208966334	1357743601:39601865	0	00:000:005:936
000:00:04:50	5	1357743661:39593829	1357743661:208956158	1357743661:208959930	1357743661:39603613	0	00:000:006:012
000:00:05:50	6	1357743721:39593829	1357743721:208948005	1357743721:208950362	1357743721:39602225	0	00:000:006:039
000:00:06:50	7	1357743781:39593829	1357743781:208939903	1357743781:208943983	1357743781:39603871	0	00:000:005:962
000:00:07:50	8	1357743841:39593829	1357743841:208931801	1357743841:208934312	1357743841:39602327	0	00:000:005:987
000:00:08:50	9	1357743901:39593829	1357743901:208923700	1357743901:208927934	1357743901:39604025	0	00:000:005:962
000:00:09:50	10	1357743961:39593829	1357743961:208915598	1357743961:208918366	1357743961:39602636	0	00:000:006:039

Figure 2.34: DM results

2.7.5 Test Results

The main purpose of this test was to prove that it is possible to configure OAM features. Actual verification was not possible, as the setup did not support monitoring T-MPLS/MPLS-TP OAM packets. CV verification was done in relation to 1+1 LSP protection as described in Section 2.8.

Dual-ended DM measurements were also possible via the craft terminal.

2.7.6 Test Conclusions

As already mentioned, the set of T-MPLS/MPLS-TP OAM features available in this release of the Alcatel-Lucent 1850 TSS was very limited. The current OAM implementation is at a very early stage and at this point only Connectivity Verification (CV) is useful in relation to 1+1 protection. Other tools like Loopback (LB) measurements and Linktrace (LT) will be available in newer software releases. Configuration of CV features was very easy and straightforward.

The OAM tools available in this release were based on the ITU-T Y.1731 standard and therefore will not interoperate with other implementations based on IETF tools such as LSP ping and BFD.

The TSS has support for a more extensive set of OAM features at the Ethernet level but the testing of these features was outside the scope of this document.

2.8 LSP 1+1 Linear Protection Test

2.8.1 Test Objective

The purpose of this test was to verify the LSP 1+1 linear protection scheme based on Connectivity Verification (CV) and Automatic Protection Switching (APS).

2.8.2 Technology Briefing

MPLS-TP 1+1 protection uses OAM CV to detect a condition or defect in the LSP. CV can be configured with different periods; 3.33 ms is usually the recommended period for 1+1 protection. In this case, CV OAM packets are sent every 3.33 ms. If the end MEP does not detect an OAM packet for 3.5 times the CV period, it declares a Loss of Connectivity alarm.

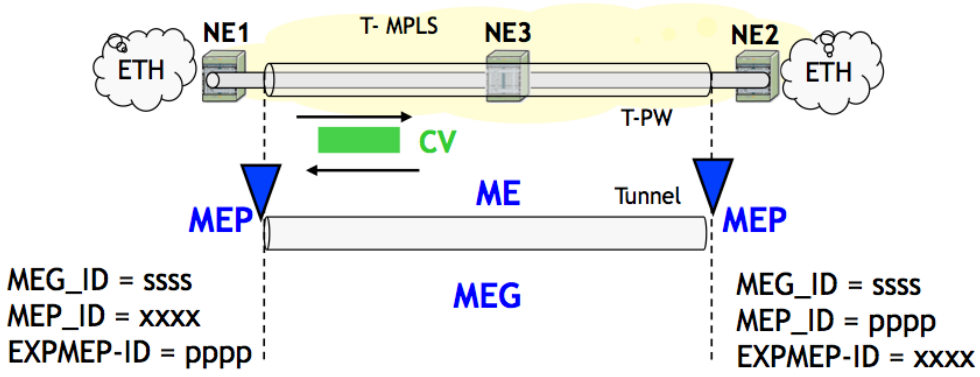


Figure 2.35: 1+1 protection with OAM CV

The Automatic Protection Switching (APS) protocol is used to signal the switching between the two end points of the LSP. APS signalling is used to synchronise between head end and tail end. It is critical that head end and tail end are always in the same state. The MPLS-TP protocol defines another protocol, called Protection State Coordination (PSC), which is very similar to APS and used for the same purpose. This implementation reuses the ITU-T-defined protocol APS.

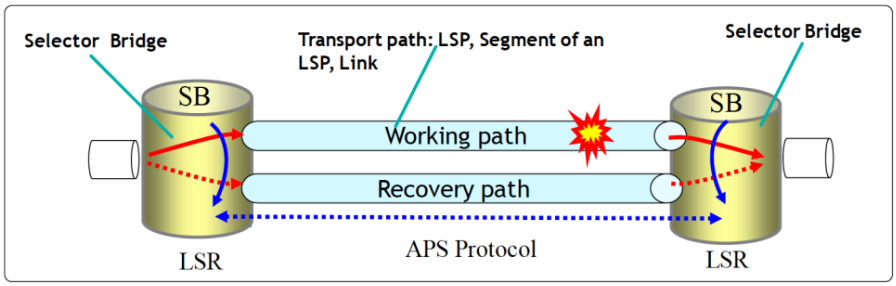


Figure 2.36: APS protocol

2.8.3 Test Setup

The test infrastructure used for this test is the same as described in Section 2.4 above.

2.8.4 Test Description

For the purpose of this test three LSPs are configured: one LSP or tunnel for each section of the network. Tunnel 1 is configured with protection and the Connectivity Verification period is set to 3.33 ms.

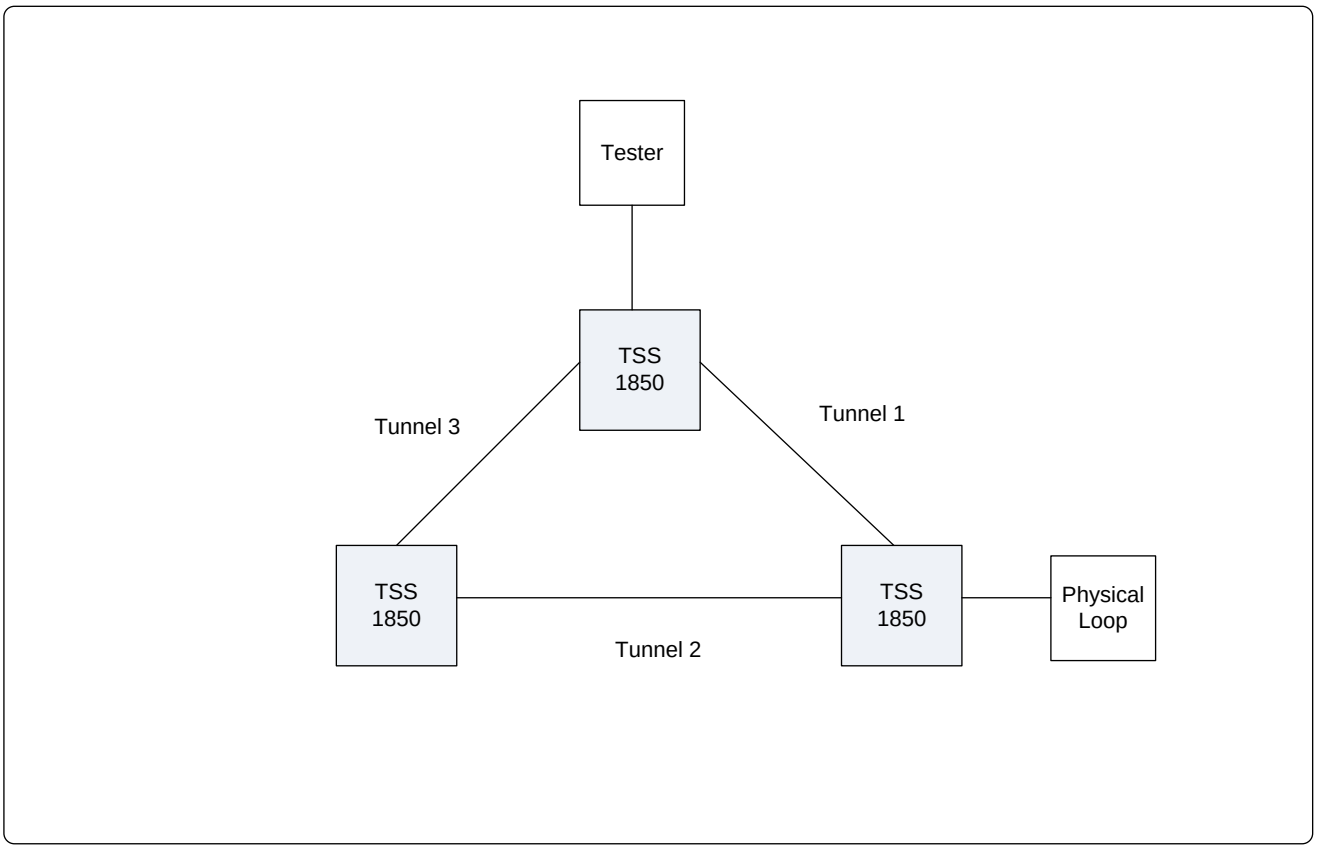


Figure 2.37: Test setup

The PWs carried by Tunnel 1 are shown in Figure 2.38.

User Label	Working S...	Last Action	Configurat...
PW_Test_1	Normal	Implementation	Implemen...
PW_test_A	Normal	Implementation	Implemen...
EVPL_Service_test_VID_...	Normal	Implementation	Implemen...
EVPL_Service_test_VID_...	Normal	Implementation	Implemen...
EVPLAN_PW_1	Normal	Implementation	Implemen...

Figure 2.38: Current PWs in Tunnel 1

The tester is connected as described in Figure 2.37 and a loop is placed in the node called “Puyol”. The tester is configured to send traffic in a specific VLAN. In this case, VLAN 10 was used. As shown in Figure 2.39, traffic is running: the alarms are green, indicating that they have all been cleared.

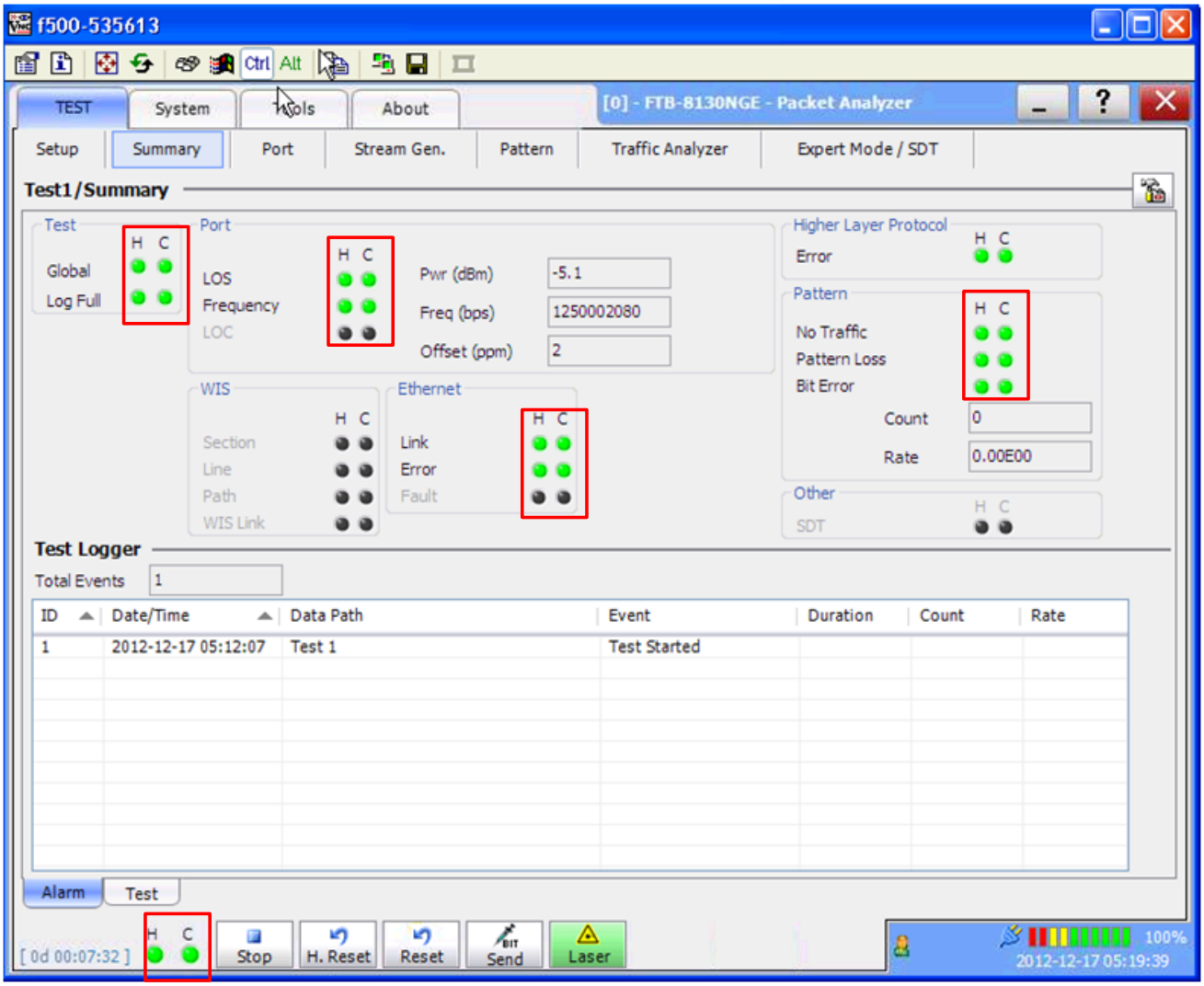


Figure 2.39: Traffic running in VLAN 10

At this point a fibre cut is simulated between the two nodes carrying Tunnel 1. As shown in Figure 2.40, the tester detects a very short Pattern Loss event. However, as the **Duration** field shows, the traffic stars running after a very short period of time.

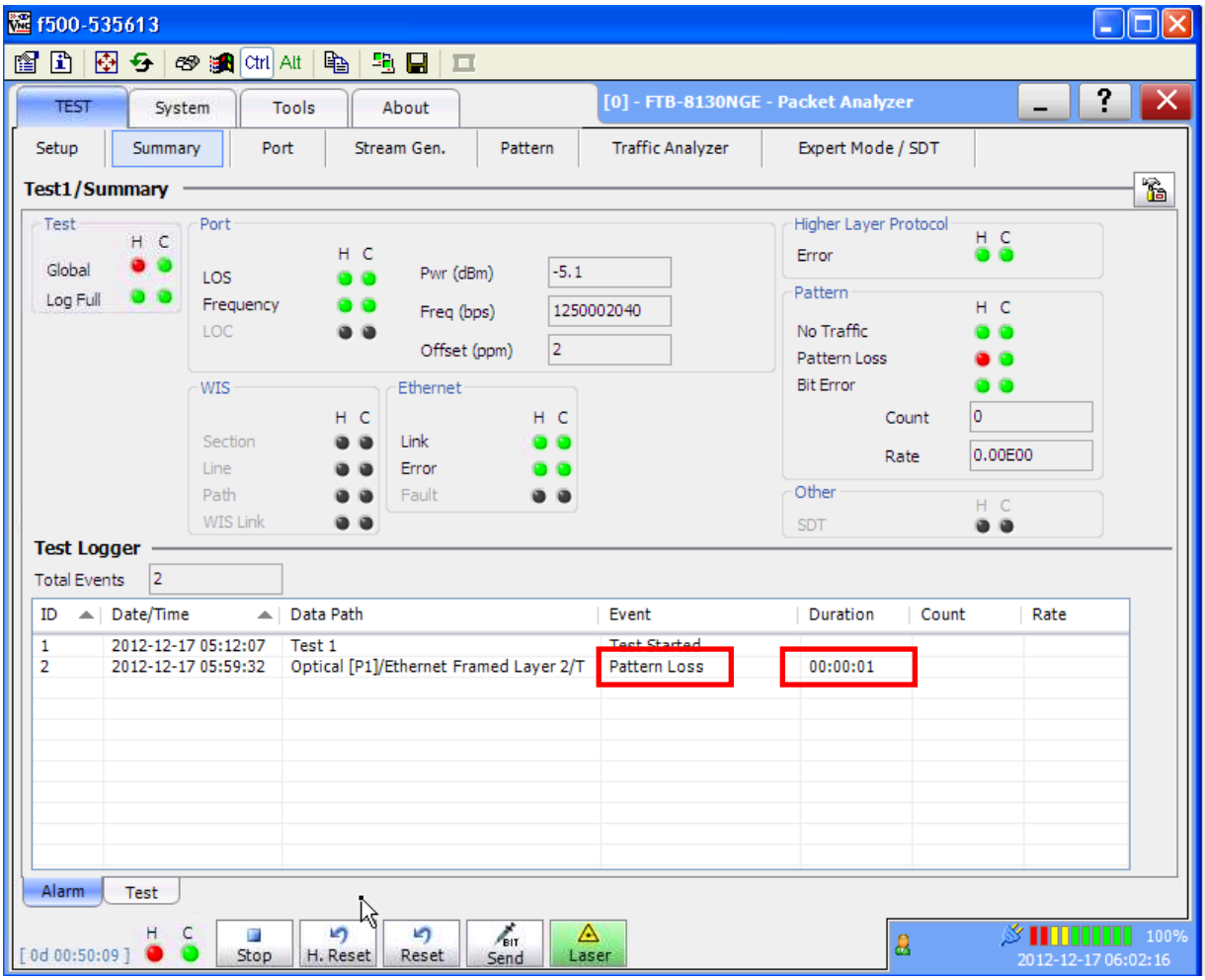


Figure 2.40: Traffic running after fibre cut

Figure 2.41 shows how the NMS also notifies the failure. The main LSP (Tunnel_test_1) is “In Failure” and is now in “Stand By” mode, while the protection LSP Tunnel_test_1_spare) is now in “Active” mode.

Figure 2.41: NMS notification of protection status

Figure 2.42, also taken from the NMS, shows how the tunnel between the nodes called “Xavi” and “Puyol” is down and the traffic is going through the alternative path.

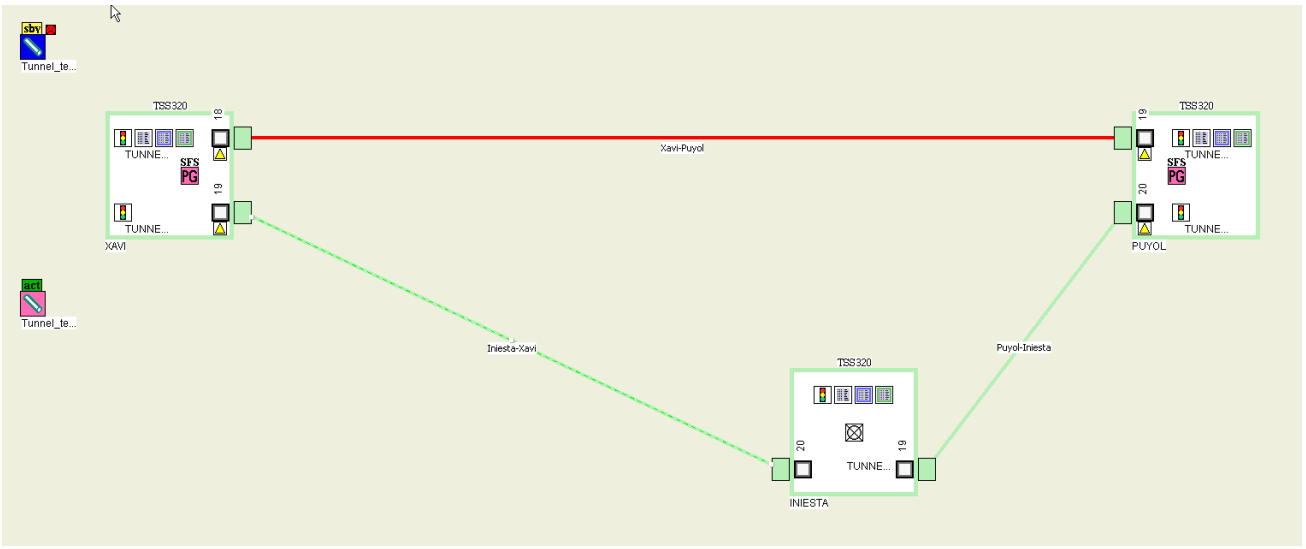


Figure 2.42: 1+1 LSP protection

It was possible to verify via the NMS that all PWs being carried by Tunnel 1 are now being carried by the protection LSP.

2.8.4.1 *Revert Time Test*

The **Revert Time** parameter is a timer used to switch back to the initial path once the problem causing the outage has been solved. This is also known as the Wait To Restore (WTR) timer. As shown in Figure 2.43, the protection is configured with a revert time of one minute. This means that the traffic will switch back to the main LSP one minute after the fibre break has been solved. Notice that a **Hold Off Time** parameter can also be configured. Hold Off Time is the time from the moment a condition is detected to the moment the APS protocol starts acting. This is used to avoid switching to the protection path in the event of flapping alarms. The APS protocol will wait for the configured Hold Off Time to ensure that the condition is permanent.

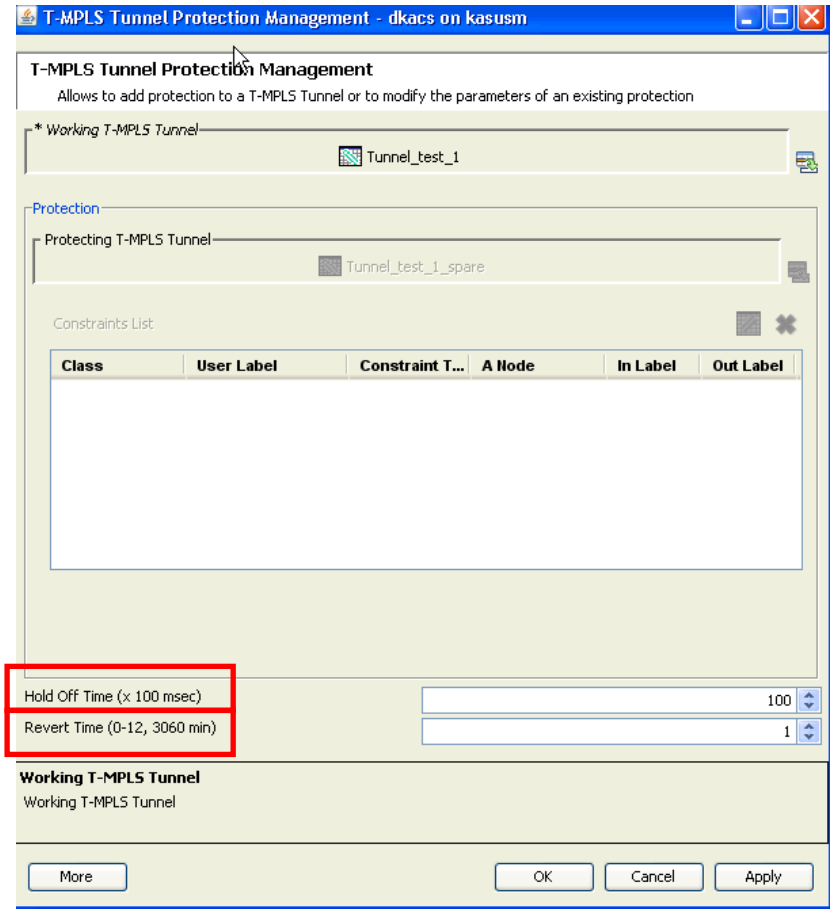


Figure 2.43: Protection Management window

The fibre between the two nodes is reconnected and, according to the configured Revert Time, the traffic should switch back to the main LSP after one minute. This is shown in Figure 2.44.

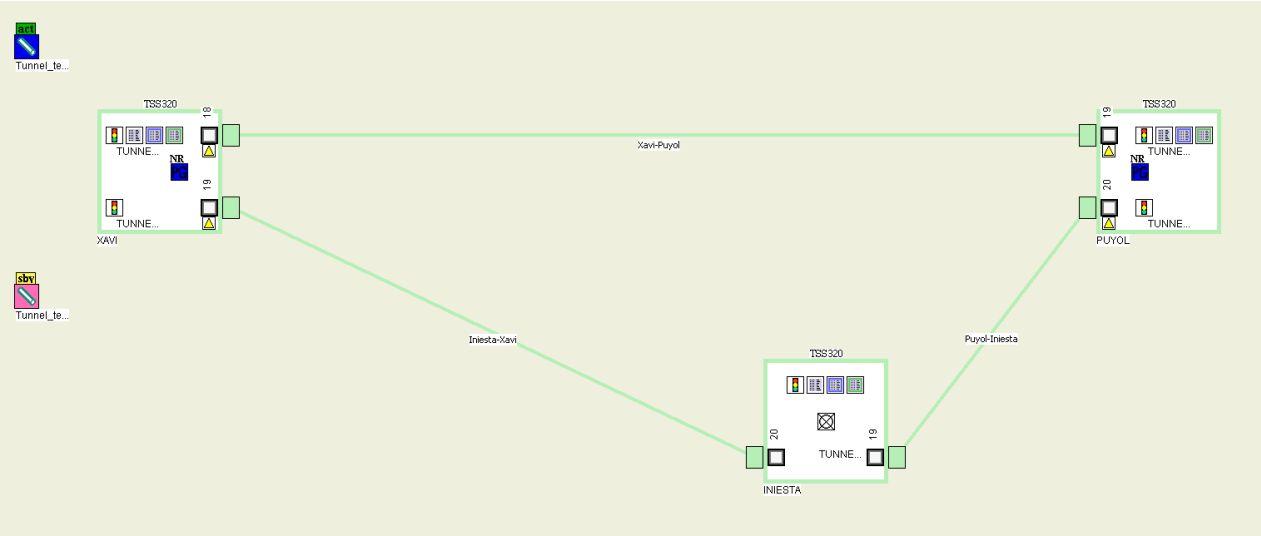


Figure 2.44: Traffic switched back to main LSP

As shown in Figure 2.45, taken from the tester, there is a minor Pattern Loss when the switch to the main LSP occurs.

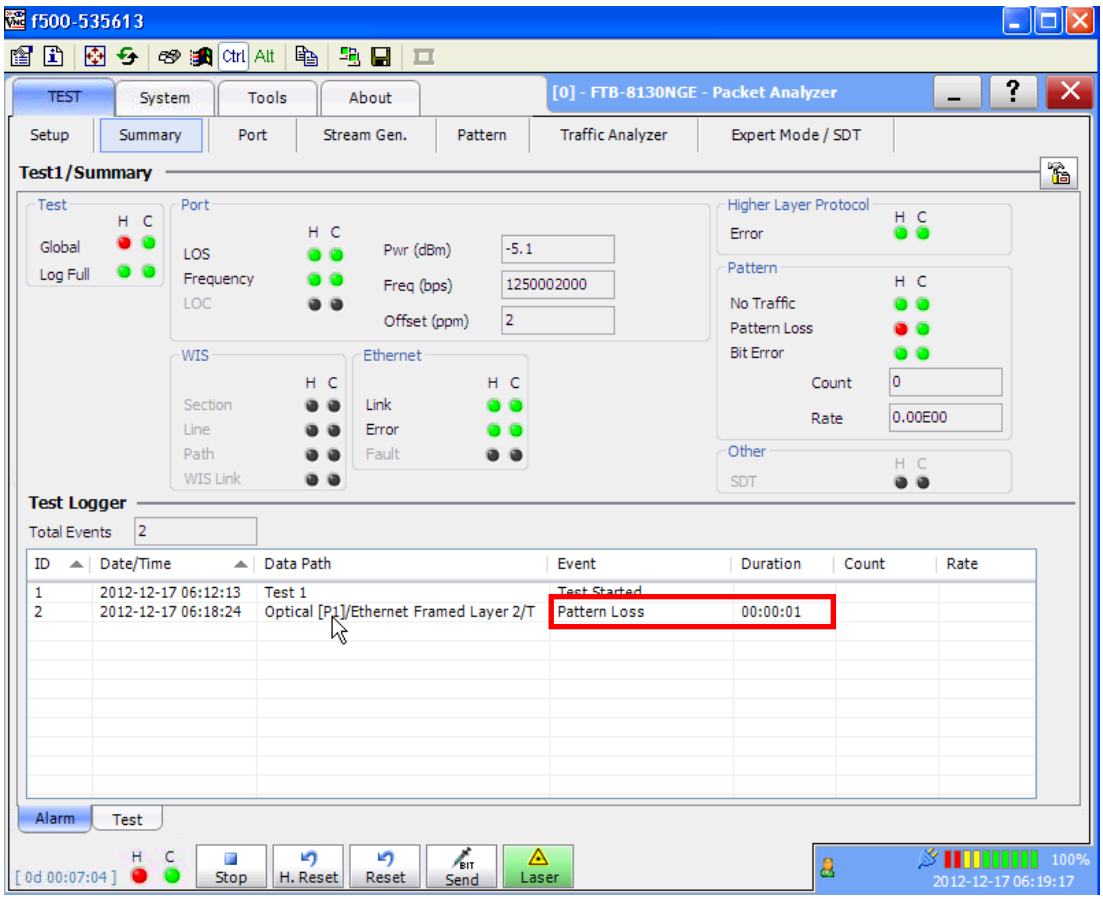


Figure 2.45: Pattern Loss when traffic switches to main LSP

2.8.4.2 *Force Switch Testing*

Force switch is a feature that is used during troubleshooting or for maintenance purposes. The operator is able to move traffic to the protection LSP while maintenance is carried out in the main link.

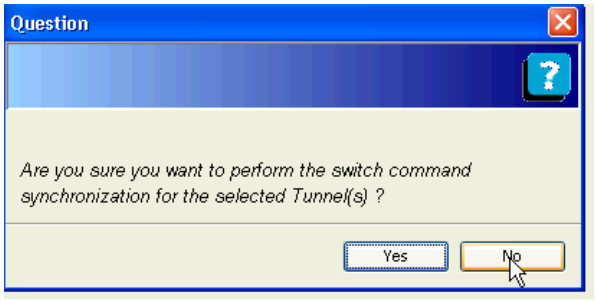


Figure 2.46: Force switch activation window

Figure 2.47 shows how the operator is able to switch to the protection path and back to the main LSP with the use of the force switch command.

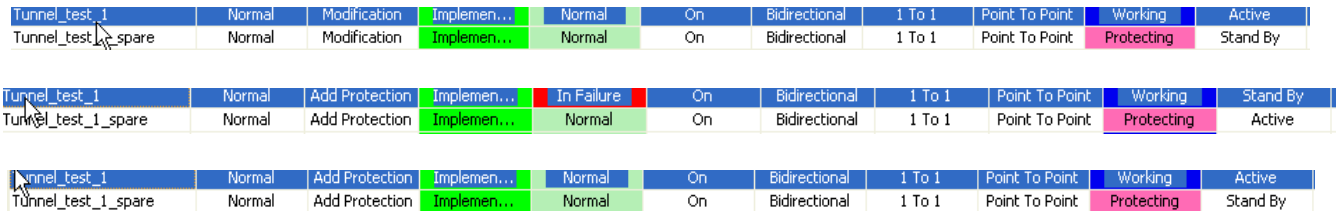


Figure 2.47: Force switch process

As shown in Figure 2.48, the traffic is still running after switching several times: the alarms are green, which indicates that they have been cleared. However, the tester shows that there has been an alarm that lasted 00:01, which is now cleared. This is the reason for the red dot under the historical alarms (“H” in the figure). This alarm shows that a few frames have been lost during the switching. It also proves that the switching actually happened.

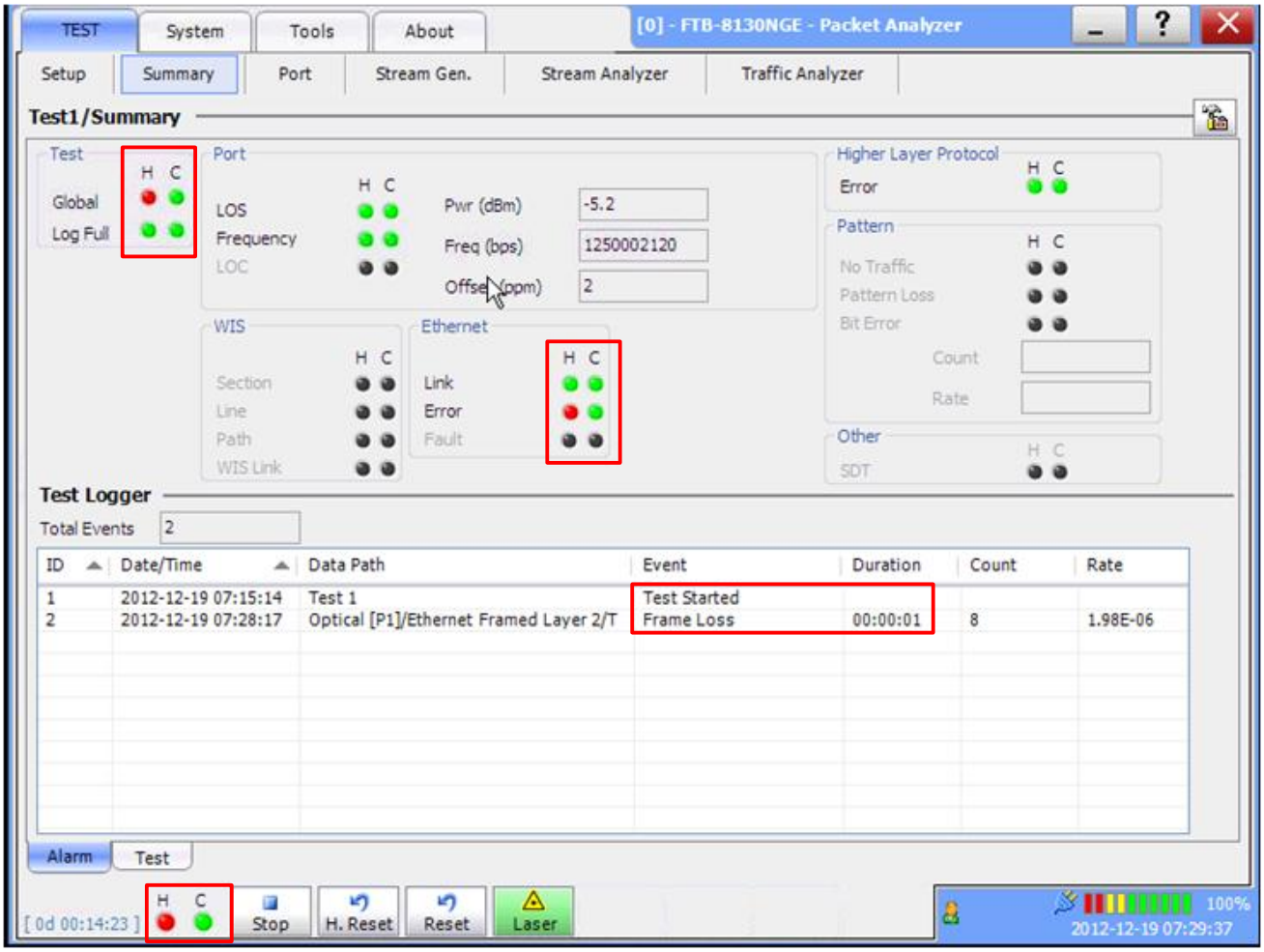


Figure 2.48: Traffic running after force switch command

2.8.5 Test Results

1+1 LSP linear protection was demonstrated and worked as expected and without error in conjunction with CV and the APS protocol. The protection LSP changed its status to active and the PWs were moved from the protected LSP to the protection LSP.

2.8.6 Test Conclusions

1+1 LSP linear protection is a very useful feature both for providing transport functionality and for ensuring service availability. 1+1 protection was very easy to configure as it was done during the LSP configuration process. The APS protocol was also verified and Wait To Restore (WTR) and Hold Off timers worked as expected, giving operators similar functionality to legacy services such as SDH transport networks. On the negative side, 1+1 LSP protection requires traffic reservation in the alternative path, which decreases the available bandwidth in the network.

2.9 Conclusions

The testing carried out was very useful to understand the current status of T-MPLS in the Alcatel-Lucent 1850 TSS-320 and the usability of the NMS to configure services, LSPs and pseudowires. The software installed in the Alcatel equipment (release 3.4) supported T-MPLS, an earlier version of MPLS-TP developed by the ITU-T. Full support for MPLS-TP would have required release 4.0. In the installed version the complete set of features was not available. However, the T-MPLS architecture, services and some basic OAM functionality were available.

The MPLS-TP protocol specifies the possibility of operating with and without the control plane (Generalised Multi-Protocol Label Switching (GMPLS)). However the equipment available for the testing did not support control plane. Instead, a Network Management System (NMS) was used to configure and monitor the network. The use of the NMS gives a better understanding of the configuration process and a good visibility of the overall network. However, the use of NMS for configuration purposes is a little cumbersome due to the number of steps that need to be taken. Comparison with the use of CLI would have been beneficial for a proper evaluation of the solution. However, this was not possible as the CLI option was not fully supported.

The use of the NMS can be an advantage for monitoring the network, while configuration of services would require a more agile method.

The testing of T-MPLS/MPLS-TP architecture, and Ethernet services gave positive results. It was possible to configure 1+1 LSP protection and verify Connectivity Verification (CV) functionality. OAM functionality was very limited in the release available and therefore it was not possible to evaluate tools like Loopback and Linktrace. These tools are already available in release 4.0, according to roadmaps of the vendor of the equipment used in the test. The OAM features available in the release used were based on the ITU-T Y.1731 standard, while other implementations in the market are based on IETF tools such as LSP ping and BFD. This duality of protocols makes interoperability impossible unless operators are given the option of choosing between both protocols.

Overall the testing was very useful for gaining experience with T-MPLS/MPLS-TP technology and for becoming familiar with protection and OAM functionality in packet-based networks. However, the implementations need to become more mature, with the introduction of a control plane and the full set of OAM features. A more agile NMS in combination with a CLI would be the best solution for allowing operators with different skill sets to choose the best option for their profile and their needs.

Finally, MPLS-TP protocol needs to be fully interoperable with IP/MPLS in order to achieve seamless connectivity of services across domains; such standardised interoperability does not yet exist.

3 EoMPLS Interoperability Tests

3.1 Overview

This chapter presents the results of testing the interoperability between the Ethernet over MPLS (EoMPLS) implementations of two leading manufacturers of the equipment for carrier class IP/MPLS networks: Juniper Networks and Cisco Systems.

3.2 Introduction

Transport of Layer 2 (Ethernet) traffic over an IP/MPLS network was identified in "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1] as one of the key Carrier Class Transport Network Technologies. The technology is gaining more and more popularity in carriers' networks as well as between research networks, replacing the legacy Time Division Multiplexing (TDM) technologies used for transporting Ethernet frames. One of the networks that has recently taken a huge step towards the use of Ethernet transport over IP/MPLS is the GÉANT network, which has replaced the Alcatel devices used for transporting Ethernet traffic over TDM technology with Juniper Networks MX-series routers, designed for transporting Ethernet over an IP/MPLS backbone.

For more information about the Ethernet over MPLS technology, please refer to "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1] and "Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing" [DJ1.1.2].

For the purpose of this chapter the term "Ethernet over MPLS" (EoMPLS) means Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS). Both VPWS and VPLS are services used for transporting Ethernet frames over an MPLS network. VPWS is an emulated point-to-point connection between two customer-facing interfaces of providers' access devices. VPLS is a multipoint service that allows multiple customer-facing interfaces to be connected and Ethernet frames to be transported over an MPLS network between any pair of the access interfaces (in a full-mesh topology).

The point-to-point VPWS is referred to in this chapter as Layer 2 Virtual Private Network (L2VPN) and sometimes the term "pseudowire" (PW) is used in order to distinguish between point-to-point and multipoint services (the term "L2VPN" used in a more general sense may refer to both point-to-point and multipoint services). The multipoint service is referred to as "VPLS" in this chapter.

The tests focus on three aspects of Ethernet over MPLS services:

- The basic interoperability of the service implementation, allowing the provisioning of Ethernet over MPLS services in a multi-vendor network.
- Use of Ethernet Operation, Administration and Maintenance (OAM) mechanisms (such as Continuity Check Protocol, loopback, linktrace) to monitor and troubleshoot an Ethernet over MPLS service in a multi-vendor network.
- Fast recovery of an Ethernet over MPLS service by using the Fast Reroute mechanisms provided by the MPLS transport plane in a multi-vendor network.

Six tests are described in this chapter, three dealing with the point-to-point service and three dealing with multipoint VPLS services. Each of the tests verifies one of the three aspects of Ethernet over MPLS services listed above.

1. Test 1 verifies the possibility of provisioning a point-to-point L2VPN service in a multi-vendor network.
2. Test 2 investigates the use of Ethernet OAM mechanisms to monitor and troubleshoot a point-to-point L2VPN service in a multi-vendor network.
3. Test 3 examines using the Fast Reroute mechanisms for faster recovery of a point-to-point L2VPN service in a multi-vendor network.
4. Test 4 verifies the possibility of provisioning a multipoint VPLS service in a multi-vendor network.
5. Test 5 explores the use of Ethernet OAM mechanisms to monitor and troubleshoot a multipoint VPLS service in a multi-vendor network.
6. Test 6 examines using the Fast Reroute mechanisms for faster recovery of a multipoint VPLS service in a multi-vendor network.

The tests were executed in the first quarter of 2013 by PSNC and CESNET using the Juniper and Cisco equipment provided by the test participants. The Juniper routers used for the test represent the MX family and the Cisco routers represent the Catalyst 6500 and ME3600 families. The Juniper equipment was installed at PSNC in Poznan and the Cisco equipment at CESNET in Prague. The two parts of the testbed were connected using the PIONIER and CESNET National Research and Education Networks.

3.3 Test Objective

There are two leading vendors of equipment for IP/MPLS networks: Juniper Networks and Cisco Systems. Both vendors' equipment is used in research networks. Each of the two vendors produces equipment that supports Ethernet transport over IP/MPLS networks (Ethernet over MPLS). The tests described in this chapter aim at verifying whether the two vendors' implementations of Ethernet over MPLS services are interoperable and can be used together in a single network which provides Ethernet over MPLS services.

3.4 Test Infrastructure

3.4.1 Description

Four Juniper Networks routers (MX-series) and four Cisco Systems routers were used to emulate a carrier's network. Four routers served as backbone routers (P routers in the MPLS terminology) and four as access routers (PE routers). Four other devices were connected to PE routers to emulate customer devices (CE devices). The CE devices supported IP ping and Ethernet OAM (CFM). For service recovery tests two of the CE devices (CE1 and CE3) were replaced with traffic generators/analysers.

The PSNC part of the testbed was built from Juniper MX80 routers, while the CESNET part was built from Cisco routers. The two parts of the testbed were connected using two links provisioned over the PSNC and CESNET national networks and their interconnection in Bielsko-Biala.

The routers inside the PSNC domain were interconnected using 10GE links, while the connections between the two domains and inside the CESNET domain were provisioned using Gigabit Ethernet technology. The CE devices were connected to the PE routers using Gigabit Ethernet links.

The topology of the testbed is shown in Figure 3.1.

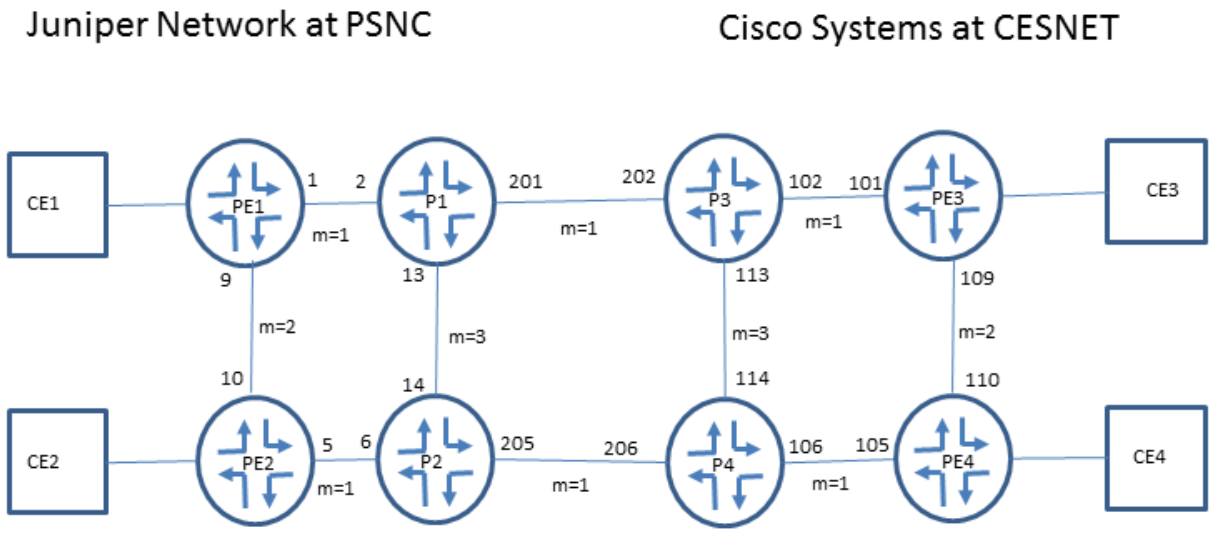


Figure 3.1: Topology of the testbed

3.4.2 Configurations

IP addresses for the links were in the form of 10.0.0.*X*/30 where *X* is given in Figure 3.1.

Loopback IP addresses:

PE1 10.1.1.1/32

PE2 10.1.1.2/32

PE3 10.1.1.3/32

PE4 10.1.1.4/32

P1 10.1.2.1/32

P2 10.1.2.2/32

P3 10.1.2.3/32

P4 10.1.2.4/32

IP addresses on CE devices:

CE1 interface IP 10.2.1.1/24

CE2 interface IP 10.2.1.2/24

CE3 interface IP 10.2.1.3/24

CE4 interface IP 10.2.1.4/24

Initial configuration:

- Intermediate System to Intermediate System (IS-IS) between all routers (Level 2 only, wide metrics, metrics for links given in Figure 3.1).
- Resource Reservation Protocol (RSVP)-provisioned Label Switched Paths (LSPs) between all PE routers.

Targeted Label Distribution Protocol (LDP) sessions were used for the L2VPN (pseudowire) tests (Tests 1–3). Border Gateway Protocol (BGP) was used for the VPLS tests (Tests 4–6). Further details are given in the descriptions of Tests 1 and 4.

3.4.3 Hardware and Software Versions

All routers in the PSNC part of the testbed (P1, P2, PE1, PE2), as well as customer devices in this part of the testbed (CE1, CE2), are Juniper Networks MX80 with JUNOS software version 11.2R4.3.

All routers and customer devices in the CESNET part of the testbed are different types of Cisco Systems products:

- P3 and P4 – Cisco c6504 (chassis WS-C6504-E, supervisor VS-SUP2T-10G, software IOS version 15.1(1)SY).
- PE3 – Cisco ME3600X (software IOS version 15.3(1)S).

- PE4 – Cisco ME-3600X-24CX-M, (software IOS version 15.3(1)S).
- CE3 and CE4 – Cisco ME-3400 (software IOS version 12.2(58)SE2).

For the service restoration (Fast Reroute) tests, traffic generators/analysers were used as customer devices (CE1 and CE3) in order to measure the number of frames lost during restoration. Agilent N2X was used as CE1 and Sunrise Telecom STT Platform with Ethernet+ module was used as CE3.

3.5 Test 1 – L2VPN (Pseudowire) Introductory Test

. It must be noted that Cisco Systems uses the term “L2VPN” to refer to both point-to-point pseudowire services and multipoint VPLS services. Juniper, in the routers’ configurations, uses the same term in the narrower meaning of a point-to-point pseudowire service with BGP signalling. The tested service (pseudowire with LDP signalling) is referred to as “l2circuit” in Juniper configurations and as “l2vpn xconnect” in Cisco configurations. The term “L2VPN” is used in this chapter for a point-to-point (pseudowire) service.

3.5.1 Test Setup

The topology of the testbed is shown in Figure 3.1.

3.5.2 Configuration

The initial configuration is shown in Section 3.4.2. Additionally, an LDP-signalled L2VPN (pseudowire) was provisioned between PE1 and PE3. The access circuits of the L2VPN were the links to CE1 and CE3.

LDP signalling was selected for this test because the Cisco routers that were used as PE devices did not support BGP signalling and auto-discovery for point-to-point L2VPN services. (BGP signalling and auto-discovery was, however, used for the multipoint VPLS tests (Tests 4–6).)

RSVP LSPs were used as transport for the service and targeted LDP sessions were used to signal VPN labels.

VLAN tagging was used on the PE–CE links. The encapsulation of the L2VPN service was “Ethernet virtual LAN” (Encapsulation type = 4).

Configuration:

- Targeted LDP sessions between all PE routers.
- VPN between PE1 and PE3 (virtual-circuit-id 3101, no control word).
- PE1–PE3 and PE3–PE1 LSPs used as transport for the L2VPN.

3.5.3 Test Description

1. Verify the state of the L2VPN using the monitoring commands available on the PE routers (and store the results for a report).
2. When the L2VPN is operational, verify the connectivity between CE1 and CE3 using ping.
3. Store the configurations for future use.

3.5.4 Expected Results

It is expected that the L2VPN will be successfully provisioned. The monitoring commands on all routers should indicate that the L2VPN is operational and ping between CE devices should confirm the state of the service.

3.5.5 Test Results

- The L2VPN instance was successfully provisioned. Monitoring commands on the PE routers indicated that the state of the service was operational and all other PE routers were connected to the service.

  Sample output from PE1 and PE3 showing the status of the L2VPN instance:

```
marcinga@PE1> show l2circuit connections
Layer-2 Circuit Connections:
Neighbor: 10.1.1.3
    Interface                 Type  St     Time last up          # Up trans
    ge-1/3/3.0(vc 3101)       rmt   Up     Feb 14 09:41:15 2013           1
      Remote PE: 10.1.1.3, Negotiated control-word: No
      Incoming label: 308880, Outgoing label: 24
      Negotiated PW status TLV: No
      Local interface: ge-1/3/3.0, Status: Up, Encapsulation: VLAN
```

- It was also verified that the L2VPN service used LSP PE1–PE3 (provisioned by RSVP).

  Output of `show route` command on PE1:

```
l2circuit.0: 3 destinations, 3 routes (3 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

10.1.1.2:NoCtrlWord:5:3101:Remote/96
                   *[LDP/9] 1w0d 02:33:43
                      Discard
10.1.1.3:NoCtrlWord:4:3101:Local/96
                   *[L2CKT/7] 1w0d 00:57:40, metric2 21
                    > to 10.0.0.2 via xe-0/0/3.0, label-switched-path PE1-PE3
                      to 10.0.0.10 via xe-0/0/1.0, label-switched-path Bypass-
      >10.0.0.2->10.0.0.102
10.1.1.3:NoCtrlWord:4:3101:Remote/96
```

```
                    *[LDP/9] 5d 04:39:48
                       Discard
```

- IP ping between CE devices proved that each CE device could communicate with all other CE devices over the provisioned L2VPN service.

  Output from CE1 confirming connectivity to CE3:

```
marcinga@CE1> ping 10.2.1.3
PING 10.2.1.3 (10.2.1.3): 56 data bytes
64 bytes from 10.2.1.3: icmp_seq=0 ttl=255 time=15.695 ms
64 bytes from 10.2.1.3: icmp_seq=1 ttl=255 time=15.777 ms
64 bytes from 10.2.1.3: icmp_seq=2 ttl=255 time=15.692 ms
64 bytes from 10.2.1.3: icmp_seq=3 ttl=255 time=46.289 ms
^C
--- 10.2.1.3 ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
round-trip min/avg/max/stddev = 15.692/23.363/46.289/13.236 ms
```

3.5.6 Test Conclusions

- The test proved interoperability between Juniper and Cisco L2VPN (pseudowire) implementations. Also, the protocols that are necessary to establish a pseudowire service and RSVP LSPs proved to be interoperable.
- The test shows that L2VPN (pseudowire) services with LSP signalling and RSVP transport can be successfully deployed in a network built from Cisco and Juniper routers. Both Cisco and Juniper routers can serve as core (P) as well as access (PE) routers in a heterogeneous network.
- The Cisco devices that were used as CE routers did not support BGP for signalling and auto-discovery of L2VPN (pseudowire) services during the testing. It is expected to be supported in the future. In the meantime, LDP can be used for signalling L2VPN labels. As targeted LDP sessions and RSVP transport of L2VPN services were successfully tested, it will be enough to activate LDP on the PE devices on which the L2VPN is terminated without any modification of the network core (P routers).
- A terminology issue was found between Cisco and Juniper configurations and documentation. The term “L2VPN” is used by Cisco for both a point-to-point pseudowire service and a multipoint VPLS service while Juniper uses the same term for point-to-point pseudowire services with BGP signalling. The exact name of the tested service in Juniper configurations and documents is “l2circuit”; Cisco Systems uses the “l2vpn xconnect ….” statement in the router configuration for the same type of service. In its documents, Cisco Systems refers to this type of service as “VPWS”.
- In addition, the terms used by the two manufacturers for Label Switched Paths are different. Juniper refers to them as “LSPs”, while Cisco uses the term “TE tunnel”.

3.6 Test 2 – Pseudowire OAM Test

3.6.1 Test Setup

The topology of the testbed is shown in Figure 3.1.

The Ethernet OAM mechanisms examined in this test are part of the Continuity Fault Management (CFM) protocols. Continuity Fault Management contains three protocols:

- Continuity Check Protocol – messages transmitted periodically by Maintenance End Points (MEPs) to other MEPs in order to verify connectivity with the other MEPs.
- Linktrace – messages sent by one MEP to another MEP in order to track the path between the two MEPs in a hop-by-hop manner and report the Maintenance Intermediate Points (MIPs) along the path.
- Loopback – messages sent by one MEP to another MEP or MIP and responses sent to the initiating MEP in order to verify connectivity. Unlike the Continuity Check Protocol, Loopback messages are initiated administratively.

3.6.2 Configuration

The configuration stored in Test 1 was used as the initial configuration for this test. Additionally, CFM for the provider domain was configured with MEPs on PE1 and PE3. CFM was associated with the pseudowire service and monitored the state of this service. CFM for the customer domain was configured between CE1 and CE3 (MEPs) with MIPs on PE1 and PE3.

Provider domain:

MD name: MD-Provider

MA name: VPLS1

Level: 5

CCM interval: 1 s

MEP IDs: 11 and 13 (respectively on PE1 and, PE3)

Customer domain:

MD name: MD-Customer

MA name: VPLS1

Level: 6

CCM interval: 1 s

MEP IDs: 101 and 103 (respectively on CE1 and CE3)

The MEPs in the provider domain (MEPs 11 and 13) were located on the customer-facing interfaces of the PE routers. The direction of the MEPs was up (in the direction of the network core). The customer domain MEPs were located on the CEs' interfaces towards the network. The direction of the MEPs was down (in the direction of the link to the network).

The Ethernet over MPLS technology does not allow MIPs to be located on the core routers (P routers) as the customer traffic is transmitted over MPLS Label Switched Paths and there are no Ethernet interfaces belonging to the customer service on core routers, on which MIPs can be located. Due to this limitation, the Ethernet OAM mechanisms cannot be used for monitoring the core routers of a pseudowire or VPLS instance.

3.6.3 Test Description

1. Use the monitoring commands available on PE1 and PE3 to verify that CFM is running between PE1 and PE3 and shows the correct status of the pseudowire service. Check the status of CFM in the customer domain (between CE1 and CE3) using the monitoring commands available on the CE devices.
2. Verify whether linktrace and loopback protocols work between the MEPs on PE1 and PE3 in the provider domain as well as between MEPs on CE1 and CE3 in the customer domain.
3. Disconnect the links between PE1 and P1 as well as between PE1 and PE2 in order to isolate PE1.
4. Check the status of the CFM session between PE1 and PE3.
5. Check the status of the CFM session between CE1 and CE3. Verify whether loopback and linktrace protocols are able to find the location of the outage.
6. Restore the connectivity to PE1.
7. Check the status of the CFM session between PE1 and PE3.
8. Check the status of the CFM session between CE1 and CE3

3.6.4 Expected Results

It is expected that CFM will be successfully deployed between PE1 and PE3 and will monitor the status of the pseudowire. In the customer domain, CFM deployed on CE1 and CE3 should be able to monitor the status of the whole customer service.

Linktrace and loopback should work in both domains. Linktrace in the customer domain should show two MIPs on the PE routers and the final MEP on the other CE device. As there were no MIPs in the provider domain, linktrace in this domain should show only the final MEP.

When the connectivity in the pseudowire is lost, CFM should be able to detect this outage. CFM in the customer domain should be able to find the location of the outage as being between PE1 and PE3. Also, when connectivity is restored, CFM should detect the restoration of connectivity.

3.6.5 Test Results

- CFM sessions between MEPs in the provider domain were successfully established and showed the status of connectivity between PE1 and PE3. Also, in the customer domain, a CFM session between CE1 and CE3 was established and showed the status of the customer service.

  Sample output from PE1 showing the status of the CFM session to PE3:

```
  Remote MEP count: 1
    Identifier    MAC address        State    Interface
       13     d4:a0:2a:55:fd:80       ok    ge-1/3/3.0
```

  Sample output from CE1 showing the status of the CFM session to CE3:

```
  Remote MEP count: 1
    Identifier    MAC address        State    Interface
      103     00:26:98:71:27:06       ok    ge-1/3/4.0
```

  Sample output from CE3 showing the status of the CFM session to CE1:

```
ME3400-1#sh ethernet cfm maintenance-points re
-------------------------------------------------------------------------------
MPID  Domain Name                                MacAddress          IfSt  PtSt
 Lvl  Domain ID                                  Ingress
 RDI  MA Name                                    Type Id             SrvcInst
      EVC Name                                                       Age
-------------------------------------------------------------------------------
101   MD-Customer                                6487.8858.c7ec      Unkn  Unkn
 6    MD-Customer                                Fa0/4
 -    VPLS1                                      Vlan 101            N/A
      N/A                                                            0s
Total Remote MEPs: 1
```

  Sample output from PE3 showing the status of the CFM session to PE1:

```
ME3600X-1#sh etherne cfm ma re
-------------------------------------------------------------------------------
MPID  Domain Name                                MacAddress          IfSt  PtSt
 Lvl  Domain ID                                  Ingress
 RDI  MA Name                                    Type Id             SrvcInst
      EVC Name                                                       Age
      Local MEP Info
-------------------------------------------------------------------------------
11    MD-Provider                                6487.8858.c7eb      Unkn  Unkn
 5    MD-Provider                                Vl101
 -    PE3-PE1                                      BD-V 101           N/A
      VPLS1                                                          0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
Total Remote MEPs: 1
```

- Loopback from PE3 to PE1 in the provider domain worked correctly. However, it did not work in the opposite direction (PE1 did not receive responses from PE3).

  Sample output from PE3 and PE1:

```
ME3600X-1#ping ethernet mpid 11 domain MD-Provider service VPLS1
Type escape sequence to abort.
Sending 5 Ethernet CFM loopback messages to 6487.8858.c7eb, timeout is 5 seconds:!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 16/16/16 ms

marcinga@PE1>ping Ethernet d4:a0:2a:55:fd:80 maintenance-domain MD-Provider
       maintenance-association VPLS1
PING to d4:a0:2a:55:fd:80, Interface ge-1/3/3.0
^C--- ping statistics ---
4 packets transmitted, 0 packets received, 100% packet loss
```

- In the customer domain, loopback worked correctly between CE devices. Each CE device received responses from the other CE without any loss of packets.

  Sample output from CE3 and CE1:

```
ME3400-1#ping ethernet mpid 101 domain MD-Customer service VPLS1
Type escape sequence to abort.
Sending 5 Ethernet CFM loopback messages to 6487.8858.c7ec, timeout is 5 seconds:!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 16/16/17 ms
marcinga@CE1>ping ethernet 00:26:98:71:27:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
PING to 00:26:98:71:27:06, Interface ge-1/3/4.0
64 bytes from 00:26:98:71:27:06: lbm_seq=9
64 bytes from 00:26:98:71:27:06: lbm_seq=10
64 bytes from 00:26:98:71:27:06: lbm_seq=11
64 bytes from 00:26:98:71:27:06: lbm_seq=12
--- ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
```

- Linktrace did not work correctly in the provider domain. None of the PE routers received responses from the other PE. As there were no MIPs in the provider domain, no intermediate hops were expected in the linktrace output.
- In the customer domain, linktrace from CE1 showed the MIP on PE1 and the destination MEP on PE3. It did not show the MIP on PE3. However, the MIP on PE3 decremented the Time to Live (TTL) as expected. (Please note there are responses with TTL 63 and 61 in the output below and the response with TTL 62 is missing.) In the opposite direction, linktrace from CE3 towards CE1 showed only the MIP on PE3. It did not show the MIP on PE1 or the destination MEP on CE1.

  Sample output from CE1 and CE3:

```
marcinga@CE1> traceroute ethernet 00:26:98:71:27:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
  Linktrace to 00:26:98:71:27:06, Interface : ge-1/3/4.0
      Maintenance Domain: MD-Customer, Level: 6
```

```
      Maintenance Association: VPLS1, Local Mep: 101
      Transaction Identifier: 7
    Hop   TTL    Source MAC address        Next-hop MAC address
    1         61     00:26:98:71:27:06         00:26:98:71:27:06
    2         63     64:87:88:58:c7:eb         00:00:00:00:00:00
marcinga@CE1>
ME3600X-1#traceroute ethernet mpid 11 domain MD-Provider service VPLS1
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 6487.8858.c7eb on Domain MD-Provider, Level 5, service VPLS1,
       evc PE1-PE3 vlan 101
Traceroute sent via Vl101, path found via MPDB
B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
-------------------------------------------------------------------------------
                             MAC      Ingress        Ingr Action  Relay Action
  Hops   Host             Forwarded   Egress         Egr Action   Previous Hop
-------------------------------------------------------------------------------
B 1                   d4a0.2a55.fd80 Gi0/4          IngOk        RlyMPDB
                      Forwarded
*
*
*
```

- To check whether the MIP on PE3 was blocking linktrace in from CE3 towards CE1, the MIP was disabled. When the MIP was disabled, linktrace in the customer domain worked correctly in both directions. It did not show the MIP on PE3 as the MIP had been disabled but it showed the MIP on PE1 and the destination MEP.

  Sample output from CE1 and CE3:

```
marcinga@CE1>traceroute ethernet 00:26:98:71:27:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
  Linktrace to 00:26:98:71:27:06, Interface : ge-1/3/4.0
      Maintenance Domain: MD-Customer, Level: 6
      Maintenance Association: VPLS1, Local Mep: 101
      Transaction Identifier: 8
    Hop   TTL    Source MAC address        Next-hop MAC address
    1     62     00:26:98:71:27:06         00:26:98:71:27:06
    2     63     64:87:88:58:c7:eb         00:00:00:00:00:00
ME3400-1#traceroute ethernet mpid 101 domain MD-Customer service VPLS1
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 6487.8858.c7ec on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4
B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
```

```
--------------------------------------------------------------------------------
                              MAC      Ingress        Ingr Action  Relay Action
  Hops    Host              Forwarded  Egress         Egr Action   Previous Hop
--------------------------------------------------------------------------------
B 1                     6487.8858.c7eb                            RlyFDB
                        Not Forwarded                EgrOK        0026.9871.2706
! 2                     6487.8858.c7ec               IngOk        RlyHit:MEP
                        Not Forwarded
```

- When PE1 was isolated and the communication in both provider and customer domains was lost, the CFM sessions in both domains correctly showed loss of connectivity.

  Sample output from PE1 and CE1 showing loss of connectivity to the remote MEP (failed status):

```
Remote MEP count: 1
    Identifier    MAC address        State    Interface
       13     d4:a0:2a:55:fd:80   failed    ge-1/3/3.0
  Remote MEP count: 1
    Identifier    MAC address        State    Interface
      103     00:26:98:71:27:06   failed    ge-1/3/4.0
```

- Loopback also confirmed loss of connectivity in both domains (it did not receive responses from the remote MEPs).
- Linktrace from CE1 in the customer domain showed only the MIP on PE1. The destination MEP on CE3 was not available due to loss of connectivity between PE1 and PE3. The MIP on PE3 was disabled.

  Sample output from CE1:

```
marcinga@CE1>traceroute ethernet 00:26:98:71:27:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
  Linktrace to 00:26:98:71:27:06, Interface : ge-1/3/4.0
      Maintenance Domain: MD-Customer, Level: 6
      Maintenance Association: VPLS1, Local Mep: 101
      Transaction Identifier: 9
    Hop    TTL    Source MAC address        Next-hop MAC address
    1      63     64:87:88:58:c7:eb         00:00:00:00:00:00
    .
    .
    .
```

- Linktrace from CE3 in the customer domain did not show any hops as the MIP on PE1 and the remote MEP on CE1 were not available and the MIP on PE3 was disabled.
- When PE1 was isolated, linktrace from the Cisco PE3 device towards the unreachable PE1 did not start due to lack of the correct entry in the Continuity Check database (CCDB). This means that once the Continuity Check Messages (CCMs) from PE1 stop being received, the switch is not aware of any MEP with MPID 13 and does not know the Media Access Control (MAC) address to which a linktrace message should be sent. The MAC address can be used (if known) instead of the MPID in the linktrace command and in this case the linktrace would start.

- When connectivity between PE1 and PE3 was restored, both CFM sessions reacted by changing the state of the remote MEPs to “ok”.

3.6.6 Test Conclusions

- Monitoring the state of CFM sessions gave the correct information about the status of connectivity in both the provider domain (between provider edge routers) and the customer domain (between customer devices).
- CFM worked correctly between Juniper and Cisco devices and can be used to monitor Ethernet over MPLS L2VPNs.
- Loopback worked correctly in the customer domain; in the provider domain it worked in the direction from Cisco to Juniper and did not work in the opposite direction (the MEP on a Juniper PE router did not receive responses from a Cisco PE router). The reason for this was not found. It is worth noting that the customer domain MEPs were established on native Ethernet devices and not on Ethernet over MPLS devices. Ethernet over MPLS was only used as transparent transport for the customer service. The provider domain MEPs were established on Ethernet over MPLS devices. It is possible that loopback is interoperable between Juniper and Cisco when MEPs are located in native Ethernet but is not interoperable when MEPs are located on edge routers of an Ethernet over MPLS L2VPN.
- Linktrace did not work in the provider domain. It worked correctly in the customer domain between CE1 and CE3 when the MIP on PE3 was disabled. When the MIP on PE3 was enabled, it seemed to be blocking linktrace in the customer domain. This may suggest that linktrace works correctly and is interoperable between Juniper and Cisco on native Ethernet devices while it does not work correctly on Ethernet over MPLS L2VPN edge routers.
- Loopback and linktrace were also tested in a pure Cisco environment between PE3 and PE4 in a similar configuration (not part of the test scenario). The result was similar: loopback and linktrace did not work between two Cisco L2VPN edge routers. This may suggest that the loopback and linktrace issue is not an interoperability problem but rather a Cisco problem, a configuration problem or a problem with a particular Cisco device used for the test. Loopback and linktrace between two Juniper routers (PE1 and PE2) in a similar configuration worked correctly.
- Linktrace worked in the customer domain when the MIP on PE3 was disabled, but due to the lack of a MIP on the border between the customer network (represented by CE3) and the provider network (on PE3) linktrace was not able to locate the outage as being in the provider network (between PE1 and PE3) or on the customer–provider connection (between CE3 and PE3).
- The test shows that linktrace and loopback do not work correctly in a L2VPN instance provisioned on PE devices from different vendors. The problem may be in the operability between the two vendors’ devices or just in one vendor’s products.

3.7 Test 3 – L2VPN Fast Reroute (Node Protection, Facility Backup) Test

3.7.1 Test Setup

The topology of the testbed is shown in Figure 3.1. Traffic generators were employed as CE1 and CE3. Each of the generators generated a one-directional packet stream towards the other generator. The packet streams did not require any reverse communication. Each of the generators also served as a receiver/analyser for the stream generated by the other generator. This feature was used to determine the amount of time needed for service restoration on the basis of the number of lost packets (difference between the number of packets sent by CE1 and received by CE3 and vice versa).

Each generator generated 100,000 packets per second.

The Fast Reroute mode selected for testing was Node Protection and Facility Backup. In the Facility Backup approach each router maintains a single backup path for a set of LSPs traversing the same interface. This makes Fast Reroute scale much better than in the One-to-One approach in which a router maintains a separate backup path for each protected LSP. The Node Protection mode makes each router establish a backup path to the next-next-hop router on the protected LSP, protecting the service not only against link failure (as in Link Protection mode, in which the backup path is established to the next-hop router) but also against a failure of the immediate next-hop router.

3.7.2 Configuration

The configuration stored in Test 1 was used as the initial configuration for this test. The only change was that the LSP between PE1 and PE3 (in both directions) was signalled via RSVP with Fast Reroute (Node Protection) capability.

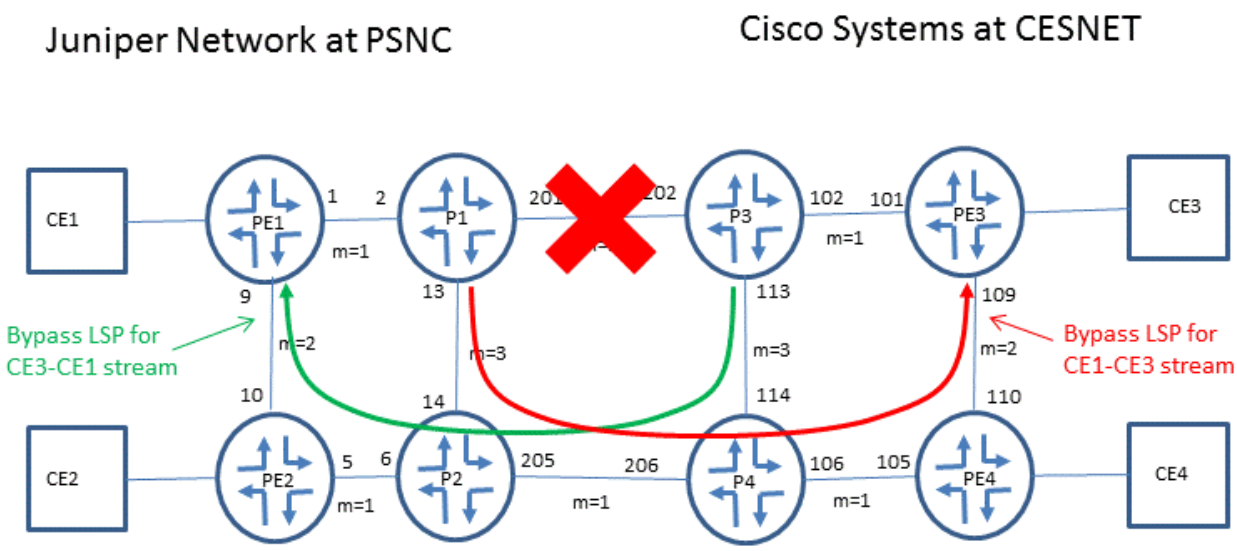


Figure 3.2: The bypass LSPs for the Fast Reroute tests

3.7.3 Test Description

1. Make sure the LSP has been established on the route PE1–P1–P3–PE3 (in both directions).
2. Use the monitoring commands available on the routers to verify the existence of bypass LSPs between P1 and PE3 as well as between P3 and PE1.
3. Disconnect the P1–P3 link. (As this link uses the PIONIER and CESNET transmission services, only one end of the link will be able to discover Loss of Signal on its interface towards the other end (the device on the other end will still see the signal from the PIONIER or CESNET access device). Make sure the link is disconnected on the P1 side to allow P1 to discover Loss of Signal.)
4. Use the monitoring commands available on the routers to verify that traffic has been rerouted to the bypass LSPs and then to a new primary path.
5. If possible, calculate the amount of time needed for service restoration based on the number of lost packets (difference between the number of packets sent by CE1 and received by CE3 and vice versa).
6. Restore the connectivity between P1 and P3 and make sure the primary path has been rerouted to the optimal route (PE1–P1–P3–PE3) in both directions.
7. Repeat steps 3–5, disconnecting the P1–P3 link on the P3 side.

3.7.4 Expected Results

It is expected that traffic will be rerouted to the bypass LSPs when connectivity between P1 and P3 is lost. Traffic should be rerouted to a new primary path when this path is established in the modified topology.

3.7.5 Test Results

- The protection was successfully established on both Juniper and Cisco routers. The existence of bypass LSPs was verified with the available `show` commands.

  Sample output:

  LSP PE1–PE3 on P1:

```
10.1.1.3
  From: 10.1.1.1, LSPstate: Up, ActiveRoute: 0
  LSPname: PE1-PE3, LSPpath: Primary
/…/
  Node/Link protection desired
  Type: Node/Link protected LSP, using Bypass->10.0.0.102->10.0.0.101
/…/
  Record route: 10.0.0.1 <self> 10.1.2.3 (node-id) 10.0.0.102 10.1.1.3 (node-id)
       10.0.0.101
```

  Bypass LSP on P1:

```
10.1.1.3
  From: 10.1.2.1, LSPstate: Up, ActiveRoute: 0
  LSPname: Bypass->10.0.0.102->10.0.0.101
/…/
  Type: Bypass LSP
    Number of data route tunnel through: 1
  Record route: <self> 10.0.0.14 10.0.0.106 10.0.0.110 10.0.0.109
```

  FRR status on P3:

```
lab-6504a#show ip rsvp fast-reroute
Primary                   Protect BW          Backup
Tunnel                    I/F     BPS:Type    Tunnel:Label  State  Level    Type
------                    ------- --------    ------------- ------ -----    ------
PE3-PE1                   Gi3/1   0:G         Tu10002:3     Active any-unl N-Nhop
PE1-PE3                   Gi3/3   0:G         Tu10000:3     Ready  any-unl Nhop
```

- When the P1–P3 link was disconnected on the PSNC side, the stream generated at CE1 towards CE3 was successfully rerouted by P1 to a new path via P2, P4 and PE2 (LSP Bypass->10.0.0.102->10.0.0.101).
- The number of frames lost during service recovery was 3,670. The calculated service recovery time was 37 ms.
- The PE1–PE3 LSP was established in the new topology (using the P2–P4 link) and traffic was moved back from the bypass LSP to the PE1–PE3 LSP. The bypass LSP was used for 1 second.
- The PE1–PE3 LSP was rerouted to the optimal path (PE1–P1–P3–PE3) when the P1–P3 link was reconnected.

- When the P1–P3 link was disconnected on the CESNET side, the stream generated at CE3 towards CE1 was successfully rerouted by P1 to a new path via P2, P4 and PE2 (LSP lab-6504a_t10002).
- The number of frames lost during service recovery was 4,258. The calculated service recovery time was 42 ms.
- The PE3–PE1 LSP was established in the new topology (using the P2–P4 link) As the duration of the experiment was very short, the traffic stream was terminated before re-establishing the LSP, so the time for which the bypass LSP was used was not calculated.
- The PE3–PE1 LSP was rerouted to the optimal path (PE3–P3–P1–PE1) when the P1–P3 link was reconnected.

3.7.6 Test Conclusions

- Fast Reroute (Node Protection, Facility Backup) was successfully tested in a multi-vendor network allowing recovery of a L2VPN service in sub-50 ms times.
- Using Fast Reroute, an MPLS network with L2VPN services can offer the same level of service recovery as TDM networks, with recovery times below 50 ms, which is required by some voice services.
- It was necessary to tune an RSVP timer on the Cisco routers to 10 ms from its default 200 ms to allow sub-50 ms recovery time. The minimal allowed time is 10 ms, but it should be taken into account that having too short an RSVP hallo interval could trigger unnecessary reroutes.

3.8 Test 4 – VPLS Introductory Test

3.8.1 Test Setup

The topology of the testbed is shown in Figure 3.1.

3.8.2 Configuration

The initial configuration is shown in Section 3.4.2 Additionally, a BGP-signalled VPLS instance was provisioned between PE1, PE2, PE3 and PE4 connecting all CE devices.

BGP configuration:

- BGP for label advertisement and auto-discovery between all PE routers.
- All PE routers in AS 10.
- Route reflectors on PE1 serving PE2 and on PE3 serving PE4.
- Cluster-id 1.1.1.1 on PE1 and 3.3.3.3 on PE3.
- iBGP session between PE1 and PE3 on loopback addresses.
- BGP for L2VPN and VPLS.

VPLS configuration

- VPLS name VPLS1
- vrf-target target:10:1
- Route-distinguisher 10.1.1.1:1
- Site name PE1, site ID 1
- Site name PE2, site ID 2
- Site name PE3, site ID 3
- Site name PE4, site ID 4
- Site-range 4

3.8.3 Test Description

1. Verify the state of the VPLS instance using the monitoring commands available on the PE routers (and store the results for a report).
2. When the VPLS is up, verify connectivity between CE2 and CE4 using ping.
3. Store the configurations for future use.

3.8.4 Expected Results

It is expected that the VPLS instance will be successfully provisioned. The monitoring commands on all routers should indicate that the VPLS is operational and ping between CE devices should confirm the state of the service.

3.8.5 Test Results

- The VPLS instance was successfully provisioned. Monitoring commands on the PE routers indicated that the state of the service was operational and all other PE routers were connected to the service. IP ping between CE devices proved that each CE device could communicate with all other CE devices over the provisioned VPLS service.

  Output from PE1 and PE3 showing the status of connections to other PE routers in the VPLS instance:

```
marcinga@PE1> show vpls connections
Layer-2 VPN connections:

Instance: VPLS1
  Local site: PE1 (1)
    connection-site           Type  St     Time last up          # Up trans
    2                         rmt   Up     Jan 17 11:15:34 2013           1
      Remote PE: 10.1.1.2, Negotiated control-word: No
      Incoming label: 262162, Outgoing label: 262161
```

```
      Local interface: lsi.17826321, Status: Up, Encapsulation: VPLS
        Description: Intf - vpls VPLS1 local site 1 remote site 2
    3                         rmt   Up     Jan 17 12:07:18 2013           1
      Remote PE: 10.1.1.3, Negotiated control-word: No
      Incoming label: 262163, Outgoing label: 35
      Local interface: lsi.17826326, Status: Up, Encapsulation: VPLS
        Description: Intf - vpls VPLS1 local site 1 remote site 3
    4                         rmt   Up     Jan 17 23:00:47 2013           1
      Remote PE: 10.1.1.4, Negotiated control-word: No
      Incoming label: 262164, Outgoing label: 47
      Local interface: lsi.17826330, Status: Up, Encapsulation: VPLS
        Description: Intf - vpls VPLS1 local site 1 remote site 4

ME3600X-1#sh bgp l2vpn vpls all
BGP table version is 58, local router ID is 10.1.1.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 10:4002
 *>  10:4002:VEID-3:Blk-1/136
                      0.0.0.0                              32768 ?
 *>i 10:4002:VEID-4:Blk-1/136
                      10.1.1.4                 0    100      0 ?
Route Distinguisher: 10.1.1.1:1
 *>i 10.1.1.1:1:VEID-1:Blk-1/136
                      10.1.1.1                      100      0 i
 *>i 10.1.1.1:1:VEID-2:Blk-1/136
                      10.1.1.2                      100      0 i
 *>  10.1.1.1:1:VEID-3:Blk-1/136
                      0.0.0.0                              32768 ?
 *>i 10.1.1.1:1:VEID-4:Blk-1/136
                      10.1.1.4                 0    100      0 ?
```

Output from CE1 confirming connectivity to other CE devices:

```
marcinga@CE1> ping 10.2.1.2
PING 10.2.1.2 (10.2.1.2): 56 data bytes
64 bytes from 10.2.1.2: icmp_seq=0 ttl=64 time=0.512 ms
64 bytes from 10.2.1.2: icmp_seq=1 ttl=64 time=1.855 ms
64 bytes from 10.2.1.2: icmp_seq=2 ttl=64 time=1.138 ms
64 bytes from 10.2.1.2: icmp_seq=3 ttl=64 time=15.222 ms
64 bytes from 10.2.1.2: icmp_seq=4 ttl=64 time=0.445 ms
^C
```

```
--- 10.2.1.2 ping statistics ---
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max/stddev = 0.445/3.834/15.222/5.717 ms

marcinga@CE1> ping 10.2.1.3
PING 10.2.1.3 (10.2.1.3): 56 data bytes
64 bytes from 10.2.1.3: icmp_seq=0 ttl=255 time=17.088 ms
64 bytes from 10.2.1.3: icmp_seq=1 ttl=255 time=17.101 ms
64 bytes from 10.2.1.3: icmp_seq=2 ttl=255 time=16.749 ms
64 bytes from 10.2.1.3: icmp_seq=3 ttl=255 time=15.684 ms
64 bytes from 10.2.1.3: icmp_seq=4 ttl=255 time=15.861 ms
^C
--- 10.2.1.3 ping statistics ---
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max/stddev = 15.684/16.497/17.101/0.607 ms

marcinga@CE1> ping 10.2.1.4
PING 10.2.1.4 (10.2.1.4): 56 data bytes
64 bytes from 10.2.1.4: icmp_seq=0 ttl=255 time=53.803 ms
64 bytes from 10.2.1.4: icmp_seq=1 ttl=255 time=16.218 ms
64 bytes from 10.2.1.4: icmp_seq=2 ttl=255 time=52.158 ms
64 bytes from 10.2.1.4: icmp_seq=3 ttl=255 time=17.256 ms
64 bytes from 10.2.1.4: icmp_seq=4 ttl=255 time=49.481 ms
^C
--- 10.2.1.4 ping statistics ---
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max/stddev = 16.218/37.783/53.803/17.243 ms

marcinga@CE1>
```

3.8.6 Test Conclusions

- The test proved interoperability between Juniper and Cisco VPLS implementations. Also, the protocols that are necessary to establish a VPLS (including RSVP and BGP) proved to be interoperable.
- The test shows that VPLS services with BGP signalling and RSVP transport can be successfully deployed in a network built from Cisco and Juniper routers. Both Cisco and Juniper routers can serve as core (P) as well as access (PE) routers in a heterogeneous network.
- One interoperability issue was encountered during the test. The Cisco VPLS implementation requires that Route Distinguishers on all PE routers participating in a VPLS instance are the same, while the Juniper implementation does not have such a requirement and allows each PE router to use a different Route Distinguisher. This issue will not cause a serious problem when manually configured Route Distinguishers are used (i.e. configured by the network administrator on each router, for each VPLS instance). In this case, the network administrator must configure the same Route Distinguisher value on all PE routers. The issue will cause a broader, more serious problem when automatically generated

Route Distinguishers are used. All Cisco routers will generate the same Route Distinguisher value based on the common AS number and VPN identifier, while each Juniper router will generate a different Route Distinguisher. This means that automatic generation of Route Distinguishers is not possible on Juniper routers when Cisco PE routers participate in the same VPLS instance and that Route Distinguishers have to be manually configured on all Juniper PE routers.

3.9 Test 5 – VPLS OAM Test

3.9.1 Test Setup

The topology of the testbed is shown in Figure 3.1.

3.9.2 Configuration

The configuration stored in Test 4 was used as the initial configuration for this test. Additionally, Continuity Fault Management (CFM) for the provider domain was configured with MEPs on all PE routers. CFM was associated with the VPLS service, to monitor the state of this service. CFM for the customer domain was configured with MEPs on all CE devices and MIPs on all PE routers.

Provider domain:

- MD name: MD-Provider
- MA name: VPLS1
- Level: 5
- CCM interval: 1 s
- MEP IDs: 11, 12, 13, 14 (respectively on PE1, PE2, PE3, PE4)

Provider domain:

- MD name: MD-Customer
- MA name: VPLS1
- Level: 6
- CCM interval: 1 s
- MEP IDs: 101, 102, 103, 104 (respectively on CE1, CE2, CE3, CE4)

The MEPs in the provider domain (MEPs 11, 12, 13 and 14) were located on the customer-facing interfaces of the PE routers. The direction of the MEPs was up (in the direction of the network core). The customer domain MEPs were located on the CEs' interfaces towards the network. The direction of the MEPs was down (in the direction of the link to the network).

The Ethernet over MPLS technology does not allow MIPs to be located on the core routers (P routers) as the customer traffic is transmitted over MPLS Label Switched Paths and there are no Ethernet interfaces belonging to the customer service on core routers, on which MIPs can be located. Due to this limitation, the Ethernet OAM mechanisms cannot be used for monitoring the core routers of a VPLS instance.

3.9.3 Test Description

1. Use the monitoring commands available on the PE routers to verify that CFM is running between the PEs and shows the correct status of connectivity in the provider domain. Check the status of CFM in the customer domain (between all CE devices) using the monitoring commands available on the CE devices.
2. Verify whether linktrace and loopback protocols work between all MEPs in the provider domain as well as between all MEPs in the customer domain.
3. Disconnect the links between PE2 and P2 as well as between PE1 and PE2 in order to isolate PE2.
4. Check the status of CFM on the PE and CE routers. Verify whether linktrace in the customer domain is able to find the location of the outage.
5. Restore the connectivity.
6. Check the status of CFM on all the PE and CE routers.
7. Disconnect the links between PE4 and P4 as well as between PE3 and PE4 in order to isolate PE4.
8. Check the status of CFM on PE and CE routers. Verify whether linktrace in the customer domain is able to find the location of the outage.
9. Restore the connectivity.
10. Check the status of CFM on all the PE and CE routers.

3.9.4 Expected Results

It is expected that CFM will be successfully deployed on all PE devices and all CE devices and will monitor the connectivity over the VPLS instance. Loopback and linktrace protocols should be able to verify connectivity to all remote MEPs in both domains. When the connectivity to PE2 and then to PE4 is lost, CFM in both domains should be able to detect this outage. When connectivity is restored, CFM should detect the restoration of connectivity.

3.9.5 Test Results

- CFM sessions were successfully established in both domains between all MEPs belonging to a particular domain.

  Sample output from PE1 showing the status of remote MEPs on other PE devices (in the provider domain):

```
Remote MEP count: 3
  Identifier     MAC address        State     Interface
```

```
       12      64:87:88:59:03:03       ok  lsi.17827075
       13      d4:a0:2a:55:fd:80       ok  lsi.17827073
       14      fc:99:47:79:91:00       ok  lsi.17827076
Sample output from PE3 showing the status of remote MEPs on other CE devices (in the
      provider domain):
ME3600X-1#sh etherne cfm ma re
-------------------------------------------------------------------------------
MPID  Domain Name                                 MacAddress          IfSt  PtSt
 Lvl  Domain ID                                   Ingress
 RDI  MA Name                                     Type Id             SrvcInst
      EVC Name                                                        Age
      Local MEP Info
-------------------------------------------------------------------------------
11    MD-Provider                                 6487.8858.c7eb      Unkn  Unkn
 5    MD-Provider                                 Vl3001
 -    VPLS1                                       BD-V 3001           N/A
      VPLS1                                                           0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
12    MD-Provider                                 6487.8859.0303      Unkn  Unkn
 5    MD-Provider                                 Vl3001
 -    VPLS1                                       BD-V 3001           N/A
      VPLS1                                                           0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
14    MD-Provider                                 fc99.4779.9100      Up    Up
 5    MD-Provider                                 Vl3001
 -    VPLS1                                       BD-V 3001           N/A
      VPLS1                                                           0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
Total Remote MEPs: 3
```

Sample output from CE1 showing the status of remote MEPs on other CE devices (in the customer domain):

```
Remote MEP count: 3
    Identifier    MAC address        State    Interface
      102     64:87:88:59:03:04       ok    ge-1/3/4.0
      103     00:26:98:71:27:06       ok    ge-1/3/4.0
      104     00:26:98:71:97:06       ok    ge-1/3/4.0
Sample output from CE3 showing the status of remote MEPs on other CE devices (in the
      customer domain):
ME3400-1#sh ethernet cfm maintenance-points re
-------------------------------------------------------------------------------
MPID  Domain Name                                 MacAddress          IfSt  PtSt
 Lvl  Domain ID                                   Ingress
 RDI  MA Name                                     Type Id             SrvcInst
      EVC Name                                                        Age
-------------------------------------------------------------------------------
```

```
101    MD-Customer                              6487.8858.c7ec      Unkn  Unkn
 6     MD-Customer                              Fa0/4
 -     VPLS1                                    Vlan 101            N/A
       N/A                                                          0s
102    MD-Customer                              6487.8859.0304      Unkn  Unkn
 6     MD-Customer                              Fa0/4
 -     VPLS1                                    Vlan 101            N/A
       N/A                                                          0s
104    MD-Customer                              0026.9871.9706      Up    Up
 6     MD-Customer                              Fa0/4
 RDI   VPLS1                                    Vlan 101            N/A
       N/A                                                          0s
Total Remote MEPs: 3
```

- Linktrace in the provider domain did not work between Juniper PE devices and Cisco PE devices, or between one Cisco PE and another. No response to the linktrace request was received in these cases. (Due to the lack of MIPs in the provider domain, the response from the remote MEPs was the only response expected.) Linktrace between two Juniper PEs worked well.
- Linktrace in the customer domain from Juniper CE devices towards Cisco PE devices showed the MIP on the Juniper PE and the MEP on the destination Cisco CE. It did not show the MIP on the Cisco PE. However, the MIP decremented TTL as expected. Linktrace between Juniper CE devices worked correctly, showing two MIPs and the destination MEP.

  Sample output showing the results of linktrace from CE1 towards CE3 and CE4 (there is no response with TTL 62 – from the MIP on the remote PE device):

```
marcinga@CE1>traceroute ethernet 00:26:98:71:27:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
  Linktrace to 00:26:98:71:27:06, Interface : ge-1/3/4.0
      Maintenance Domain: MD-Customer, Level: 6
      Maintenance Association: VPLS1, Local Mep: 101
      Transaction Identifier: 10
    Hop    TTL    Source MAC address        Next-hop MAC address
    1      61     00:26:98:71:27:06         00:26:98:71:27:06
    2      63     64:87:88:58:c7:eb         64:87:88:58:c8:32
marcinga@CE1>traceroute ethernet 00:26:98:71:97:06 maintenance-domain MD-Customer
       maintenance-association VPLS1
  Linktrace to 00:26:98:71:97:06, Interface : ge-1/3/4.0
      Maintenance Domain: MD-Customer, Level: 6
      Maintenance Association: VPLS1, Local Mep: 101
      Transaction Identifier: 12
    Hop    TTL    Source MAC address        Next-hop MAC address
    1      61     00:26:98:71:97:06         00:26:98:71:97:06
    2      63     64:87:88:58:c7:eb         64:87:88:58:c8:32
```

- Linktrace from CE3 towards Juniper CE devices in the customer domain did not work (it showed only the MIP on PE3). Linktrace from CE3 towards CE4 worked correctly, showing the MIP on PE3 and the

destination MEP on CE4. The MIP on PE4 was not shown in the linktrace output but it decremented the TTL value as expected.

Sample output showing the results of linktrace from Cisco CE3 towards Juniper CE2:

```
ME3400-1#traceroute ethernet mpid 102 domain MD-Customer vlan 101
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 6487.8859.0304 on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4

B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
-------------------------------------------------------------------------------
                            MAC      Ingress        Ingr Action  Relay Action
  Hops   Host             Forwarded   Egress         Egr Action   Previous Hop
-------------------------------------------------------------------------------
B 1                  d4a0.2a55.fd80 Gi0/4          IngOk        RlyMPDB
                     Forwarded                                  0026.9871.2706
*
*
*
*
```

Sample output showing the results of linktrace from Cisco CE3 towards Cisco CE4:

```
ME3400-1#traceroute ethernet mpid 104 domain MD-Customer vlan 101
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 0026.9871.9706 on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4

B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
-------------------------------------------------------------------------------
                            MAC      Ingress        Ingr Action  Relay Action
  Hops   Host             Forwarded   Egress         Egr Action   Previous Hop
-------------------------------------------------------------------------------
B 1                  d4a0.2a55.fd80 Gi0/4          IngOk        RlyMPDB
                     Forwarded                                  0026.9871.2706
! 3                  0026.9871.9706 Fa0/4          IngOk        RlyHit:MEP
                     Not Forwarded                              fc99.4779.9100
*
*
*
*
```

- Linktrace from CE4 towards the Juniper CE devices in the customer domain worked correctly, showing two MIPs on the PE devices and the destination MEP. Linktrace from CE4 towards CE3 showed the MIP on PE3 and the destination MEP on CE4. It did not show the MIP on PE4.

  Sample output showing the results of linktrace from Cisco CE4 towards Juniper CE1 and CE2:

```
ME3400-2#trace ethernet mpid 101 domain MD-Customer vlan 101
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 6487.8858.c7ec on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4
B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
------------------------------------------------------------------------------
                            MAC      Ingress        Ingr Action  Relay Action
  Hops   Host             Forwarded   Egress         Egr Action   Previous Hop
------------------------------------------------------------------------------
B 1                   fc99.4779.9100 Gi0/4          IngOk        RlyMPDB
                      Forwarded                                  0026.9871.9706
B 2                   6487.8858.c832                             RlyFDB
                      Not Forwarded                EgrOK         fc99.4779.9100
! 3                   6487.8858.c7ec               IngOk         RlyHit:MEP
                      Not Forwarded                              fc99.4779.9100
ME3400-2#trace ethernet mpid 102 domain MD-Customer vlan 101
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 6487.8859.0304 on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4
B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
------------------------------------------------------------------------------
                            MAC      Ingress        Ingr Action  Relay Action
  Hops   Host             Forwarded   Egress         Egr Action   Previous Hop
------------------------------------------------------------------------------
B 1                   fc99.4779.9100 Gi0/4          IngOk        RlyMPDB
                      Forwarded                                  0026.9871.9706
B 2                   6487.8859.034a                             RlyFDB
                      Not Forwarded                EgrOK         fc99.4779.9100
! 3                   6487.8859.0304               IngOk         RlyHit:MEP
                      Not Forwarded                              fc99.4779.9100
ME3400-2#trace ethernet mpid 103 domain MD-Customer vlan 101
Type escape sequence to abort. TTL 64. Linktrace Timeout is 5 seconds
Tracing the route to 0026.9871.2706 on Domain MD-Customer, Level 6, vlan 101
Traceroute sent via Fa0/4
B = Intermediary Bridge
! = Target Destination
* = Per hop Timeout
```

```
-------------------------------------------------------------------------------
                              MAC      Ingress        Ingr Action  Relay Action
  Hops   Host              Forwarded   Egress         Egr Action   Previous Hop
-------------------------------------------------------------------------------
B 1                    fc99.4779.9100 Gi0/4          IngOk        RlyMPDB
                       Forwarded                                  0026.9871.9706
! 3   ME3400-1         0026.9871.2706 Fa0/4          IngOk        RlyHit:MEP
                       Not Forwarded                              d4a0.2a55.fd80
*
*
```

- Loopback in the provider domain did not work from Juniper PE devices towards Cisco PE devices, or between Cisco PE devices (PE3–PE4). It worked correctly from Cisco PE devices towards Juniper PE devices as well as between Juniper PE devices (PE1–PE2).

  Sample output from PE3 showing the results of loopback towards other PE devices:

```
ME3600X-1#ping ethernet mpid 11 domain MD-Provider service VPLS1
Type escape sequence to abort.
Sending 5 Ethernet CFM loopback messages to 6487.8858.c7eb, timeout is 5 seconds:!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 16/18/24 ms
ME3600X-1#ping ethernet mpid 12 domain MD-Provider service VPLS1
Type escape sequence to abort.
Sending 5 Ethernet CFM loopback messages to 6487.8859.0303, timeout is 5 seconds:!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 16/22/48 ms
ME3600X-1#ping ethernet mpid 14 domain MD-Provider service VPLS1
Type escape sequence to abort.
Sending 5 Ethernet CFM loopback messages to fc99.4779.9100, timeout is 5 seconds:.....
Success rate is 0 percent (0/5)
```

- Loopback in the customer domain worked correctly between all CE devices.
- When PE2 was isolated, CFM showed loss of connectivity to the MEPs on PE2 and CE2 respectively in the provider and the customer domain.

  Sample output from PE3 showing loss of connectivity to PE2 in the provider domain (the output shows the MEPs on PE1 and PE4; the MEP on PE2 is missing as it was not reachable due to loss of connectivity):

```
ME3600X-1#sh ethernet cfm maintenance-points remote
-------------------------------------------------------------------------------
MPID  Domain Name                                MacAddress          IfSt  PtSt
 Lvl  Domain ID                                  Ingress
 RDI  MA Name                                    Type Id             SrvcInst
      EVC Name                                                       Age
      Local MEP Info
-------------------------------------------------------------------------------
11    MD-Provider                                6487.8858.c7eb      Unkn  Unkn
 5    MD-Provider                                Vl3001
 RDI  VPLS1                                      BD-V 3001           N/A
```

```
      VPLS1                                                    0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
14    MD-Provider                              fc99.4779.9100  Up    Up
 5    MD-Provider                              Vl3001
 -    VPLS1                                    BD-V 3001       N/A
      VPLS1                                                    0s
      MPID: 13 Domain: MD-Provider MA: VPLS1
Total Remote MEPs: 2
```

- When PE2 was isolated, linktrace from the Cisco PE devices towards the unreachable PE2 did not start due to lack of the correct entry in the Continuity Check database (CCDB). This means that once the CCMs from PE2 stop being received, the switch is not aware of any MEP with MPID 12 and does not know the MAC address to which a linktrace message should be sent. The MAC address can be used (if known) instead of the MPID in the linktrace command; provided the status of the remote MEP is "Up" and "only sending CCM by the remote MEP" is turned off, the linktrace will work.

  Sample output from PE3 showing the result of linktrace towards PE2:

```
ME3600X-1#trace ethernet mpid 12 domain MD-Provider service VPLS1
No entry found in CCDB for mpid 12 at domain MD-Provider service VPLS1, evc VPLS1
       vlan 3001. Traceroute terminated.
```

- When PE2 was isolated, linktrace from the Cisco CE devices towards the unreachable CE2 in the customer domain did not start due to lack of the correct entry in the CCDB. Again, the MAC address can be used instead of the MPID (remote MEP ID).

  Sample output from CE3 showing the result of linktrace towards CE2:

```
ME3400-1#trace ethernet mpid 102 domain MD-Customer vlan 101
No entry found in CCDB for mpid 102 at domain MD-Customer, vlan 101. Traceroute
       terminated.
```

- When PE2 was isolated, loopback from the Cisco PE devices towards the unreachable PE2 as well as from the Cisco CE devices towards the unreachable CE2 did not start due to lack of the correct entry in the CCDB.
- When connectivity to PE2 was restored, CFM changed the state of the sessions to PE2 in the provider domain and to CE2 in the customer domain to reflect the change in connectivity.
- When PE4 was isolated, the state of the CFM sessions showed loss of connectivity (the state of the remote MEP changed to "failed") to PE4 in the provider domain and to CE4 in the customer domain.

  Sample output from PE1 showing the state of CFM sessions (the state of the session to PE4 is "failed", which means that the MEP on PE4 is not reachable):

```
Remote MEP count: 3
  Identifier    MAC address        State    Interface
     12      64:87:88:59:03:03       ok  lsi.17827331
     13      d4:a0:2a:55:fd:80       ok  lsi.17827328
     14      fc:99:47:79:91:00   failed
```

  Sample output from CE1 showing the state of CFM sessions (the state of the session to CE4 is "failed"):

```
Remote MEP count: 3
  Identifier    MAC address        State    Interface
    102     64:87:88:59:03:04       ok    ge-1/3/4.0
    103     00:26:98:71:27:06       ok    ge-1/3/4.0
    104     00:26:98:71:97:06   failed    ge-1/3/4.0
```

- Linktrace and loopback in the provider domain were not tested when PE4 was isolated because linktrace and loopback in this domain did not work from Juniper PE devices towards Cisco PE devices.
- Linktrace from Juniper CE devices towards CE4 in the provider domain did not show any hops when CE4 was unavailable due to loss of connectivity to PE4.
- Loopback from Juniper CE devices towards the unreachable CE4 in the customer domain showed loss of connectivity (lack of response from the MEP on CE4).
- When connectivity to PE4 was restored, CFM changed the state of the sessions to PE4 in the provider domain and to CE4 in the customer domain to reflect the change in connectivity.

3.9.6 Test Conclusions

- Monitoring the state of CFM sessions gave the correct information about the status of connectivity in both the provider domain (between provider edge routers) and the customer domain (between customer devices).
- CFM worked correctly between Juniper and Cisco devices and can be used to monitor VPLS instances.
- Loopback worked correctly in the customer domain; in the provider domain it worked in the direction from Cisco to Juniper and did not work in the opposite direction (the MEP on a Juniper PE router did not receive responses from a Cisco PE router). In addition, loopback did not work between MEPs on Cisco PE devices. The reason for this was not found. It is worth noting that the customer domain MEPs were established on native Ethernet devices and not on Ethernet over MPLS devices. Ethernet over MPLS was only used as transparent transport for the customer service. The provider domain MEPs were established on Ethernet over MPLS devices. It is possible that loopback is interoperable between Juniper and Cisco when MEPs are located in native Ethernet but is not interoperable when MEPs are located on edge routers of a VPLS service. The fact that loopback did not work between two Cisco PE devices may also suggest that the loopback issue is not an interoperability problem but rather a Cisco problem, a configuration problem or a problem with a particular Cisco device used for the test. Loopback between two Juniper routers (PE1 and PE2) in a similar configuration worked correctly.
- Linktrace worked in the customer domain from Juniper CE devices towards Cisco PE devices. However, the response from Cisco MIPs (on Cisco PE devices) was not present in the output.
- Linktrace from CE3 towards Juniper CE devices in the customer domain did not work; linktrace from CE4 towards Juniper CE devices worked correctly. The difference may have been caused by the different PE devices to which the CE devices were connected: PE3 and PE4 had the same software and similar configurations, but the hardware of the devices was different. The PE4 hardware was newer than PE3, which may explain the different results.
- The test shows that linktrace and loopback do not work correctly in a VPLS instance provisioned on PE devices from different vendors. The problem may be in the operability between the two vendors' devices or just in one vendor's products.

- Problems with the MIPs and linktrace on the Cisco PEs could be caused by the incorrect interworking of the implementation of CFM functions and EoMPLS (VPLS) over RSVP LSPs (TE tunnels in the Cisco terminology). Full official support of BGP signalling with Ethernet over MPLS services over RSVP-signalled LSPs is expected in the IOS version 15.3S(2) (scheduled for April 2013); the version used in the tests was 15.3.S(1). Until full support is available, the use of OAM for a VPLS instance in a multi-vendor environment is not practical.

3.10 Test 6 – VPLS Fast Reroute (Node Protection, Facility Backup) Test

3.10.1 Test Setup

The topology of the testbed is shown in Figure 3.1. Traffic generators were employed as CE1 and CE3. Each of the generators generated a one-directional packet stream towards the other generator. The packet streams did not require any reverse communication. However, the packets travelling in the opposite direction were used by the PE routers for learning MAC addresses. Each of the generators also served as a receiver/analyser for the stream generated by the other generator. This feature was used to determine the amount of time needed for service restoration on the basis of the number of lost packets (difference between the number of packets sent by CE1 and received by CE3 and vice versa).

Each generator generated 100,000 packets per second.

The Fast Reroute mode selected for testing was Node Protection and Facility Backup. More information about this mode is given in the setup description of Test 3 (Section 3.7.1).

3.10.2 Configuration

The configuration stored in Test 4 was used as the initial configuration for this test. The only change was that the LSP between PE1 and PE3 (in both directions) was signalled via RSVP with Fast Reroute (Node Protection, Facility Backup) capability.

3.10.3 Test Description

1. Make sure the LSP has been established on the route PE1–P1–P3–PE3 (in both directions). Use the available monitoring commands to verify this.
2. Use the monitoring commands available on the routers to verify the existence of bypass LSPs protecting the P1–P3 link (in both directions).
3. Disconnect the P1–P3 link. (As this link uses the PIONIER and CESNET transmission services, only one end of the link will be able to discover Loss of Signal on its interface towards the other end (the

device on the other end will still see the signal from the PIONIER or CESNET access device). Make sure the link is disconnected on the P1 side to allow P1 discover Loss of Signal.)

4. Use the monitoring commands available on the routers to verify that traffic has been rerouted to the bypass LSPs and then to a new primary path.
5. If possible, calculate the amount of time needed for service restoration based on the number of lost packets (difference between the number of packets sent by CE1 and received by CE3 and vice versa).
6. Restore the connectivity between P1 and P3 and make sure the primary path has been rerouted to the optimal route (PE1–P1–P3–PE3) in both directions.
7. Repeat steps 3–5, disconnecting the P1–P3 link on the P3 side.

3.10.4 Expected Results

It is expected that traffic will be rerouted to the bypass LSPs when connectivity between P1 and P3 is lost. Traffic should be rerouted to a new primary path when this path is established in the modified topology.

It is expected that VPLS services will recover in sub-50 ms times.

3.10.5 Test Results

- The protection was successfully established on both Juniper and Cisco routers. The existence of bypass LSPs was verified with the available `show` commands. (As the configuration of the LSPs was the same as in test 3 and the output of the `show` commands was similar to test 3, the output has been omitted from the description of these results.)
- When the P1–P3 link was disconnected on the PSNC side, the stream generated at CE1 towards CE3 was successfully rerouted by P1 to a new path via P2, P4 and PE2 (LSP Bypass->10.0.0.102->10.0.0.101).
- The number of frames lost during service recovery was 4,047. The calculated service recovery time was 40 ms.
- The PE1–PE3 LSP was established in the new topology (using the P2–P4 link) and traffic was moved back from the bypass LSP to the PE1–PE3 LSP. The bypass LSP was used for 56 seconds.
- The PE1–PE3 LSP was rerouted to the optimal path (PE1–P1–P3–PE3) when the P1–P3 link was reconnected.
- When the P1–P3 link was disconnected on the CESNET side, the stream generated at CE3 towards CE1 was successfully rerouted by P1 to a new path via P2, P4 and PE2 (LSP lab-6504a_t10001).
- The number of frames lost during service recovery was 1,413. The calculated service recovery time was 14 ms.
- The PE3–PE1 LSP was established in the new topology (using the P2-P4 link) As the duration of the experiment was very short, the traffic stream was terminated before re-establishing the LSP, so the time for which the bypass LSP was used was not calculated.
- The PE3–PE1 LSP was rerouted to the optimal path (PE3–P3–P1–PE1) when the P1–P3 link was reconnected.

3.10.6 Test Conclusions

- Fast Reroute (Node Protection, Facility Backup) was successfully tested in a multi-vendor network allowing recovery of a VPLS service in sub-50 ms times.
- Using Fast Reroute, an MPLS network with VPLS services can offer the same level of service recovery as TDM networks, with recovery times below 50 ms, which is required by some voice services.
- It was necessary to tune an RSVP timer on the Cisco routers to 50 ms from its default 200 ms to allow sub-50 ms recovery time. Minimal allowed time is 10 ms, but should be taken into account that having too short an RSVP hallo interval could trigger unnecessary reroutes.
- Bypass LSP was used for a much longer time than in Test 3 (L2VPN Fast Reroute Test, described in Section 3.7). The reason for this is that a router periodically tries to re-establish LSPs that are not operational. There is one timer used by the router to trigger re-establishing of all LSPs, and it does not depend on when a particular LSP was broken, or on the time when bypass LSP started to be used. The time for which a bypass LSP is used depends on the time between starting to use the bypass and the first LSP re-establishment triggered by the timer.

3.11 Conclusions

Ethernet over MPLS point-to-point (VPWS) and multipoint (VPLS) services were tested in a multi-vendor network built from Juniper Networks and Cisco Systems routers. The tests aimed to verify the interoperability of the two vendors' Ethernet over MPLS implementations in three areas:

- The basic interoperability of the service implementation, allowing the provisioning of Ethernet over MPLS services in a multi-vendor network.
- Use of Ethernet OAM mechanisms (Continuity Check Protocol, loopback, linktrace) to monitor and troubleshoot an Ethernet over MPLS service in a multi-vendor network.
- Fast recovery of an Ethernet over MPLS service by using the Fast Reroute mechanisms provided by the MPLS transport plane in a multi-vendor network.

The results of the tests described in this chapter show that the implementations of the tested services available on Juniper Networks and Cisco Systems routers are interoperable and that routers from the two vendors can be used together in a single network that provides Ethernet over MPLS services.

One interoperability issue was encountered in the configuration of BGP-signalled VPLS services. Cisco devices require that all routers participating in a VPLS instance use the same Route Distinguisher for the common VPLS instance, while Juniper allows each router to use different Route Distinguisher value. This may cause a problem when connecting a Cisco router as a new PE router in a VPLS instance provisioned on Juniper routers. The problem may be greater when using automatically generated Route Distinguishers on Juniper routers, as the algorithm used for generating Route Distinguishers on Juniper routers will generate different Route Distinguisher values on different routers in the same VPLS instance.

The service recovery tests proved that Fast Reroute mechanisms can be used to achieve sub-50 ms recovery times for Ethernet over MPLS services.

The Ethernet OAM tests confirmed that Continuity Check Protocol can successfully monitor the state of an Ethernet over MPLS service in both the provider domain (between provider edge devices) and the customer domain (between customer devices connected to the provider's network) and provide information about loss and recovery of connectivity to remote devices. The other OAM protocols (linktrace and loopback) did not work correctly over the Ethernet over MPLS services. Use of the linktrace and loopback protocols in a multi-vendor network with Ethernet over MPLS services requires further investigation, after all vendors implement full protocols.

Some terminology differences between the terms used by Cisco Systems and Juniper Networks have been encountered and are worth noting:

- The term "L2VPN" is used by Cisco for both variants of the Ethernet over MPLS services (point-to-point VPWS and multipoint VPLS service) while Juniper uses the same term in the wider sense of point-to-point VPWS services with BGP signalling. The exact name of the tested VPWS service (with LDP signalling) in Juniper configurations and documents is "l2circuit"; Cisco Systems uses the term "L2VPN cross connect" (l2vpn xconnect) for the same type of service.
- In addition, the terms used by the two manufacturers for Label Switched Paths are different. Juniper refers to them as "LSPs", while Cisco uses the name "TE tunnel".

4 Time-Sensitive Transport Technologies

4.1 Introduction

Transport and metro networks should be capable of supporting transport of all types of services. For supporting real-time applications, such as high-quality videoconferencing, delay and packet delay variation (PDV) are key parameters that must be kept to a minimum. Furthermore, packet-layer synchronisation, using, for example, the IEEE 1588 Precision Time Protocol (PTP), relies heavily on a low network PDV for achieving high precision [Ferrant]. Mobile base stations require synchronisation signals from the mobile backhaul network for coordinating handover [Cosart]. With the strong growth in mobile networking and the migration from circuit-switched SDH/SONET to packet-based mobile backhaul networks, there is therefore an emerging need for transporting synchronisation information at the packet layer across Ethernet networks [Briggs].

This chapter addresses data transport for very time-sensitive applications and services, such as banking, interconnection of data centres, videoconferencing for eHealth, etc. It is not within the scope of the document to identify such services, however; their existence is assumed.

The requirements for such time-sensitive transport (TST) might be satisfied using transport only with the circuit-switched transport layer, using technologies such as OTN or SDH, or using different flavours of packet switching, ranging from MPLS-TP, through Provider Backbone Bridge Traffic Engineering (PBB-TE), to traditional best-effort IP transport.

In the chapter, the physical layer, the transport layer and the packet layer are described, with a focus on the delay-contributing functions in each. The physical layer includes the Wavelength-Division Multiplexing (WDM) optical layer, including physical interfaces, etc. The transport layer includes mapping into frames and any delay related to this. Finally, the packet layer includes the necessary buffering, taking into account the QoS functionalities available. For the packet layer, MPLS-TP and, mainly, PBB-TE have been chosen.

The theoretical analysis is supported by measurements from experimental setups.

Hence, the chapter is organised as follows:

- In Section 4.2, the physical layer and WDM are described, with a focus on how these contribute to the delay.
- Section 4.3 addresses OTN and includes how functions in this layer contribute to the overall delay.
- Section 4.4 covers the same areas for the packet technologies.

- Section 4.5 discusses a hybrid circuit-packet solution using Fusion switches.

Each of the sections listed above includes both a theoretical part and a description of the tests carried out. Finally, in Section 4.6, a short conclusion is provided.

4.2 Physical Media Dependent Layer

4.2.1 Technology Overview

The physical layer comprises the optical transport medium and the electrical and optical interfaces in the nodes. The physical media dependent layer includes the processing (analog or digital) of the binary signal to whatever modulation format is needed.

Data transport over long distances requires at least a basic set of transmission equipment, which today, in its simplest form, is deployed by Dense Wavelength-Division Multiplexing (DWDM). Regardless of the kind of overlay network (transport networks (e.g. SDH) or data networks (e.g. IP)), the basic DWDM is required as the most fundamental networking building block. It is therefore important to be aware of its limitations regarding latency. The aim of this section is to examine the latency introduced by DWDM itself, in order to understand the overall latency.

4.2.1.1 *Latency Sources*

Latency sources in a DWDM network include the transmission time in the fibre itself, the optical components, and the opto-electrical components needed to carry out the function of optical networks from source to destination. Figure 4.1 shows a typical DWDM system. The optical components in the figure are mainly the In-Line Amplifiers (ILAs), Dispersion Compensation Modules (DCMs, either Dispersion Compensation Fibre (DCF)-based DCM or Fibre Bragg Grating (FBG)-based DCM), and Optical Add-Drop Multiplexers (OADMs).

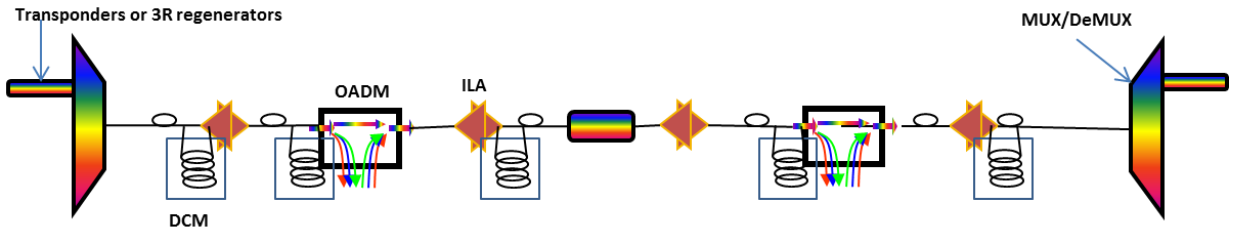


Figure 4.1: DWDM transmission components

Fibre

Light in vacuum travels at 299792458 m/s, which, in terms of latency, translates to 3.33 µsec/km. However, the optical signal propagation speed through a fibre cable is lower than the speed of light because of the cable's so-called refractive index property. This speed reduction increases the latency to approximately 5.0 µsec/km.

For this source of latency, it is important to note also that the fibre is not always installed on the shortest physical route between two points, which would also add to the latency.

With regard to the optical fibre itself, nothing can be done to improve the latency of the already deployed fibre. However, there have been some demonstrations of “Hollow Core Fibre”, which would be able to reduce the fibre latency dramatically [Jay].

DCF

The optical component that introduces the most latency is the Dispersion Compensation Fibre (DCF). The DCF is used to remove the effects of the so-called chromatic dispersion of the fibre, which degrades the fibre’s performance and limits its maximum reachable distance. Only used in long-distance networks, a DCF is a long spool of special fibre that is usually 15% – 25% of the total fibre length, therefore adding 15% – 25% of latency. The longer the distance, the more DCFs need to be in place, adding up to a few milliseconds of latency.

In order to reduce the latency introduced by dispersion compensation modules, the compensation technique based on Fibre Bragg Grating (FBG) could be used. One other method, which can be used as a result of the introduction of 100 Gbit/s channel capacity, is to use coherent detection technique and Digital Signal Processing (DSP) at the receiver side, which eliminates the need for dispersion compensation along the fibre path, moves all dispersion compensation to the receiver and achieves compensation digitally. More complex receivers at transponders will add some latency introduced by the transponder itself, but the overall latency will be reduced by 100s µs (compared with DCF).

Amplifiers

A second optical component is the optical amplifier. Amplification of the signal is needed in long-distance networks, since the light source fades as it travels through the fibre. Optical amplifiers also remove the need for optical-electrical-optical (OEO) conversion, which of course removes latency. The most common in-line optical amplifier is the Erbium-Doped Fibre Amplifier (EDFA). An EDFA adds about 10 m of extra fibre. Each EDFA introduces only tens of nanoseconds of delay. This is negligible compared to DCF, but it needs to be taken into account since there are tens of EDFA amplifiers along the transmission link. Raman amplifiers are another kind of amplifier that amplify the signal along the transmission path and that (unlike EDFA) do not require additional fibre. Theoretically, Raman amplifiers do not introduce latency.

Opto-Electrical Components

The opto-electrical components include transponders and muxponders. Transponders/muxponders are in charge of colour conversion (traffic signals are converted from grey to a specific light colour or wavelength), mapping and OTN frame adaption. Also, Forward Error Correction (FEC) is performed in transponders/muxponders. Such a device introduces up to 100 µs depending on the mapping type, OTN frame adaption and FEC mechanism used.

4.2.2 Physical Media Latency Experiment

4.2.2.1 *Test Objective*

The purpose of the test was to examine the latency on DWDM.

4.2.2.2 *Test Setup*

The test used UNINETT's DWDM path between Trondheim and Oslo, as shown in Figure 4.2(A) The DWDM path is about 639 km long, and consists of 7 DCMs (5 based on DCF and 2 based on FBG), 11 EDFA amplifiers, 4 (de)multiplexers, 1 Variable Optical Attenuator (VOA) and 1 interleaver.

As long as the components are pure optical components (no O-E and E-O conversion involved), an upper limit of the latency can be estimated by the dimensions of the element together with the velocity of light in matter of about 20 cm per ns in the glass. Assuming 15 cm as the worst-case assumption, the length of the path of light in any card including patch cords should not be longer than 90 cm, corresponding to 6 ns latency at maximum (exception: DCF spools contain many km of fibre and EDFA with about 10 m of fibre). All amplifiers together introduce no more than 1µs; it can be negligible. These numbers are therefore completely negligible compared to transmission fibres, DCF spools and electrical processing latency times. The test therefore considers latency contributions from DCMs and transponders only.

In addition to the DWDM, the testbed consists of Spirent equipment, which can generate traffic and analyse the performance. The Spirent tester is connected to a low latency switch and further to the DWDM link. The data traffic is looped back to DWDM by another switch at the far end. To find out the latency contribution from the transponders, two tests were performed: one with transponders and one without transponders, where the switches were connected to the DWDM as an alien wavelength.

As is shown in Figure 4.2(B), the latency contribution from the switches and Spirent tester was measured in advance in order to subtract this contribution from the overall latency measured according to Figure 4.2.

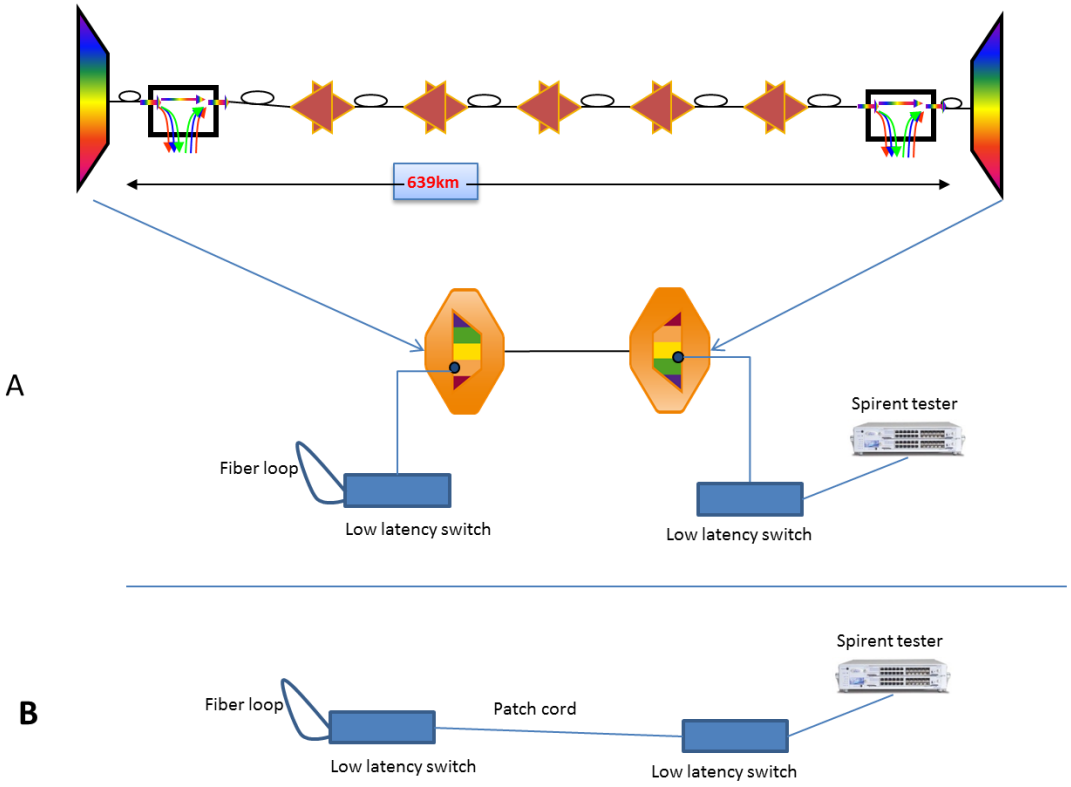


Figure 4.2: Field testbed scenarios. Latency over DWDM using transponders (A) and without transponders (B)

4.2.2.3 *Test Description*

The following test was performed:

- Latency contribution from the switches and Spirent tester as shown in Figure 4.2(B).

4.2.2.4 *Test Results*

The test results are summarised in Table 4.1.

Latency introduced by switches and Spirent [µs]	Total latency with transponder [µs]	Total latency without transponder [µs]
41	7849	7584

Table 4.1: Latency based on the three different scenarios as shown in Figure 4.2

Based on the results in Table 4.1, it was concluded that the latency contribution from the transponder (in this case) is 66 µs.

The total fibre path (including access fibre) is approximately 1290 km. Assuming 5 µs/km delay from the fibre, this leads to a total fibre latency of 6450 µs.

Item	Latency [µs]	Description
Fibre	6450	1290km @ 5 µs/km
DCM	390.5	10 fibre based DCM and 4 FBG based DCM. The data is according to vendor documentation.
Transponders	265.5	Calculated from the test results on DWDM with/without transponders
Switches and Spirent tester	40.65	Based on Figure 4.2
Optical components	100	The contribution from all optical components which are considered to be negligible, but we assumed to be not more than 100 µs
Total calculated	7246.65	The sum of above values
Measured by test	7848.96	The result by testing over DWDM
Difference between measured and expected latency	+602.31	We measure about 300 µs more latency in each direction than we expected.

Table 4.2: Latency sources of DWDM path

As shown in Table 4.2, approximately 300 µs more latency was measured than was expected. It is believed that this is due to the unreliability of the fibre length assumptions.

4.2.3 Conclusions for the Physical Layer

The main source of latency in a transport network is the transmission fibre. Assuming a DWDM length of 1000 km fibre, the latency of the fibre will be 80% – 90% of the total latency in the network. The second largest latency contribution comes from fibre-based DCMs. The new DWDM networks, which have deployed coherent detection and DSP, will have a design that is free of in-line dispersion. This will eliminate delay from DCMs.

The third biggest latency contribution comes from transponders (and regenerators), especially if a more complex Forward Error Correction is used. The difference between FEC and Signal processing-based FEC (SFEC) could lead to additional latency of more than 100 µs.

4.3 Optical Transport Network (OTN)

4.3.1 Technology Overview

The Optical Transport Network (OTN) transport technology has been developed and described in ITU-T Recommendation G.709 (2003). Since 2003 the standard has been reviewed several times to adapt it to new market needs, especially in order to adapt to different types of Ethernet bit rates and physical interfaces. The technology is client-signal agnostic in the sense of transporting any kind of client signal by encapsulating an existing frame of data, regardless of the native protocol, to create an optical data unit (ODU). It is also supported by Operation, Administration, Maintenance and Provisioning (OAM&P) functionality to support OAM&P in optical carriers such as DWDM. Some of other main functionalities are:

- Using forward error correction (FEC) mechanism that improves error performance on noisy links. This introduces increased processing delay.
- ODU-switching becomes more and more important due to its ability to switch different client signals efficiently and fast through the network.
- Powerful protection and restoration mechanisms.
- Ability to monitor, detect and locate faults in a multi-domain and multi-vendor network.

OTN technology and its developments in recent years have been described in depth in the first deliverable of JRA1 Task 1, “Deliverable DJ1.1.1: Transport Network Technologies Study” [DJ1.1.1].

4.3.1.1 *Frame Delay Measurement (FDM)*

In order to better support transport of delay-critical data like time transport and Fibre Channel, it is important to have control over end-to-end delay of the service path in the transport network. It is also important to record changes in delay which will occur in the case of link error and protection switching.

It is possible to measure the round trip delay of any ODUk using a specific field of ODU overhead. The text below from [DJ1.1.1] shows the delay measurement mechanism in OTN transport networks.

The Delay Measurement Message (DMM) and the Delay Measurement Reply (DMR) include time stamps that are used to calculate the frame delay. Delay measurement can be accomplished on either a one-way or round-trip basis. One-way delay measurement, by one DMM message, requires synchronised clocks between the transmitting and receiving Maintenance End Points (MEPs) to achieve an accurate measurement. The one-way delay is calculated by the following equation:

Frame Delay = RxTimef – TxTimeStampf

where:

- RxTimef = time that the 1DM PDU was received.
- TxTimeStampf = time that the 1DM PDU was sent.

Two-way delay measurement avoids the clock synchronisation issue, but could incur inaccuracy due to the DMM to DMR processing in the target MEP. Consequently, Y.1731 allows two options in the measurement of two-way delay. If the target MEP turnaround delay is not considered significant, then the round-trip delay can be calculated by the following equation:

Frame Delay=RxTimeb-TxTimeStampf

where:

- RxTimef= time that the DMR protocol data unit (PDU) is received by the initiating MEP.

A more accurate two-way delay measurement can be achieved if the target MEP turnaround delay is subtracted out. In this case, the round-trip delay can be calculated as follows:

Frame Delay=(RxTimeb-TxTimeStampf)-(TxTimeStampb-RxTimeStampf)

where:

- TxTimeStampb= time that the DMR PDU is sent by the target MEP.
- RxTimeStampf= time that the DMM PDU is received by the target MEP.

This second option requires that the DMR PDU include two additional time stamps: TxTimeStampb and RxTimeStampf.

4.3.1.2 *Vendor Support of FDM*

In the past two years, ODU switches have received a lot of attention in the optical transport networking community. The first generations of ODU switches have been ready since 2011, and new functionalities are becoming available on these platforms. Almost all major transport network vendors have ODU switches in their product portfolio, but there are some differences between them regarding support of different OTN functionalities. Support of FDM is only implemented by a few vendors, but almost all of them have this functionality in their roadmap.

4.3.2 OTN Delay Experiment

4.3.2.1 *Test Objective*

The purpose of the OTN measurements test was to quantify the delay caused by the OTN processing induced mainly by the FEC mechanism.

4.3.2.2 *Test Setup*

Three Alcatel Lucent 1830 PSS DWDM/OTN switches have been made available; two of these are located in Copenhagen (PSS1 and PSS2) and one in Hamburg (PSS3). The measurements are divided into two measurement series.

First, the two co-located switches (PSS1 and PSS2) are used to measure the delay from the OTN processing; the propagation delay can be ignored as only 10–15 meters of fibre is used. Second, all three switches are included and one to three bi-directional loops are created between all three switches, including the fibre delay. The latter measurements are used as a sanity check and verification.

The delay is measured using a Xena Bay Ethernet tester located at Technical University of Denmark (DTU) with a connection to the PSS1 through the Danish National Research and Education Network DeIC.

4.3.2.3 *Loop Internally in PSS1: Test Description and Results*

In order to measure the delay between the test location at DTU and PSS1, an internal loop was created within PSS1. A reduced RFC 2544 for latency was conducted, with the results given in Table 4.3.

64 bytes	512 bytes	1518 bytes	9600 bytes
2324,00	2357,36	2431,269	3033,246

Table 4.3: Internal loop in PSS1 to isolate the delay to the tester. All delays in µsec.

Due to the inherent store and forward processing in the Ethernet equipment in between DTU and the PSS1 location at NORDUnet, the frame delay is heavily dependent on the frame size. The results are, however, linear with the frame size, and for a “zero-length” reference frame the delay is thus 2319 µsec.

4.3.2.4 *Delay in PSS1 and PSS2 without Propagation Delay*

Test Description

The Gigabit Ethernet connection from the tester is mapped to an ODUO in PSS1. This signal is then transmitted to PSS2 using an OTU connection which includes the OTN FEC capability. It should be noted that standard FEC is enabled according to ITU G.975.1 App. option I.9. The setup is illustrated in Figure 4.3.

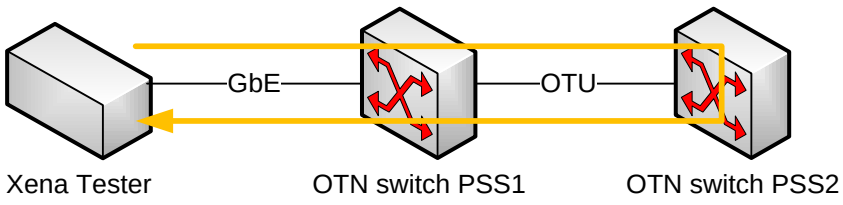


Figure 4.3: Loop in PSS2 - 2 times OTN processing

As a verification of the OTN processing delay observed through the above measurement, an extra loop is created looping back to PSS1, as shown in Figure 4.4. Hence two extra uni-directional processings are necessary.

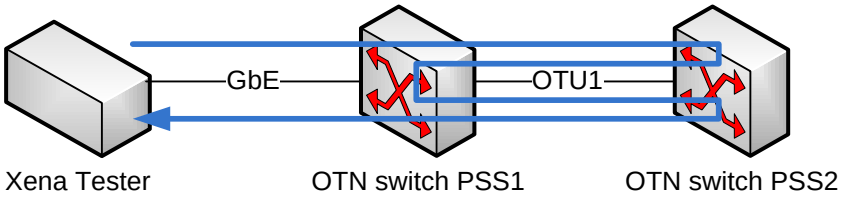


Figure 4.4: Looping through PSS2 to PSS1 – 4 times OTN processing

Test Results

The measurements are given in Table 4.4, where the delay is measured and, for each frame size, the values from Table 4.3 are subtracted. It is seen that the OTN processing delay, defined as the processing from an OTU receiver to an OTU transmitter, is independent of the Ethernet frame size, which is as expected as the Ethernet signal is just a client signal for the OTN. The uni-directional OTN processing delay for the Alcatel PSS OTN switch is thus 40 µsec. According to ITU G.709 Appendix A, the interleaving of the 16 FEC rows into an OTU row requires a FEC of 4080 bytes, which, for an ODU0 rate of 1G, approximates to a delay of 33 µsec for the decoding, which is in line with the observed results.

Measurement	64 bytes	512 bytes	1518 bytes	9600 bytes
Extra delay 2 OTN processing	80,466	80,612	80,795	80,838
Extra delay 4 OTN processing	161,521	161,198	161,245	161,314

Table 4.4: Extra delay for 2 and 4 OTN processing functions. Delays in µsec.

4.3.2.5 *"Tours" in the OTN Triangle*

Test Description

The results in the previous section are further used in a long-distance setup including the three Alcatel OTN switches PSS1, PSS2 and PSS3 located in Copenhagen, Copenhagen and Hamburg, respectively.

The delay in the OTN triangle is measured using 1, 2 and 3 bi-directional loops, as illustrated in Figure 4.5.

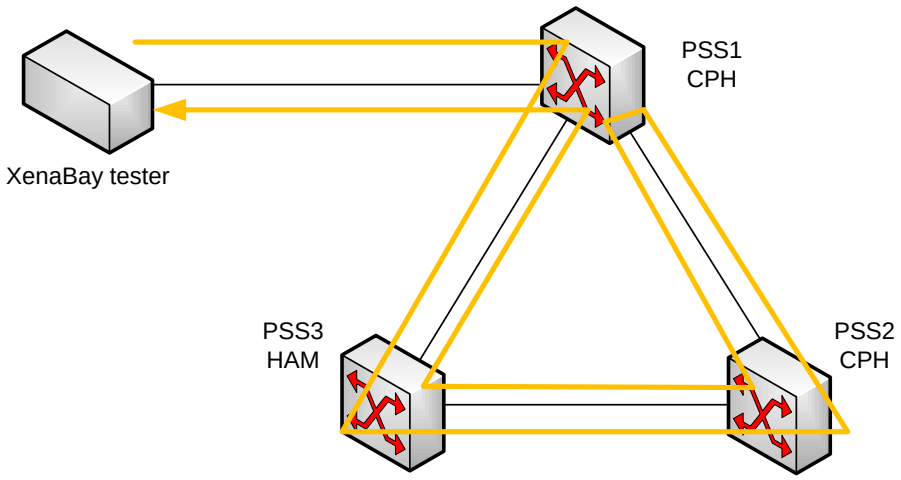


Figure 4.5: OTN triangle test

The OTN switches are located in Copenhagen (PSS1), Copenhagen (PSS2 and Hamburg (PSS3). Here, a one bi-directional loop is illustrated. For 2 and 3 loops, the connection is continuing through PSS2 and PSS3 to PSS1 before turning around.

In these tests the data loops around in the OTN triangle from one to several times.

The delay in the loops (0/1/2/3) can be quantified as:

- Reference delay (previously measured): GbE transmission and propagation delay from the XenaBay test solution at DTU through DeIC (Danish NREN) to NORDUnet's premises. This includes various Ethernet trans- and muxponders. Furthermore it includes the mapping from GbE into the ODU0.
- Propagation delay: propagation delay in the loop between OTN switches. Each loop will have 2 propagation delays.
- OTN processing delay: OTN switching delay (as measured previously).

Table 4.5 shows the delay contributions for 0, 1, 2 and 3 bi-directional loops.

Loops in triangle	Delay contribution
0	Reference delay
1	Reference delay + 2 propagation delay + 6 OTN processing delays
2	Reference delay + 4 propagation delay + 12 OTN processing delays
3	Reference delay + 6 propagation delay + 18 OTN processing delays

Table 4.5: OTN measurement scenarios and delay factors

As previously described, the distance between PSS1 and PSS2 is insignificant. However, the distances between PSS1 and PSS3 and between PSS2 and PSS3 are 670.4 and 406.6 km, equalling a total fibre length of 1077 km. It was not possible to obtain exact figures on the delay in the dispersion compensatoig modules.

If, however, the propagation delay is approximated to the fibre transmission length of the single mode fibre, then the expected delays can be calculated, as shown in Table 4.6.

Loops	Delay contributions	Expected delay
0	2319 µsec + 2 x 1077km x 5 µsec/km + 6 x 40 µsec	2319 µsec
1	Reference delay + 2 propagation delay + 6 OTN processing delays	13329 µsec
2	Reference delay + 4 propagation delay + 12 OTN processing delays	24339 µsec
3	Reference delay + 6 propagation delay + 18 OTN processing delays	35349 µsec

Table 4.6: OTN measurement scenarios and calculated delay

For each loop configuration a reduced RFC2544 delay measurement was conducted with packet sizes of 64, 512, 1518 and 9600 bytes. The measurements were done using a GbE tester from Xena Networks. Each of the tests were done with 50% and 99% load; however, only the results with 99% load are used in the further analysis.

Test Results

The results of the measurements for 99% load are given in Table 4.7.

Loops	64 bytes	512 bytes	1518 bytes	9600 bytes
0	2324,00	2357,36	2431,269	3033,246
1	13877,06	13910,57	13984,55	14586,51
2	25282,26	25315,57	25389,44	25991,43
3	36798,49	36832,45	36906,68	37508,73

Table 4.7: Measured delay for OTN loops

The measurements are illustrated in the chart in Figure 4.6.

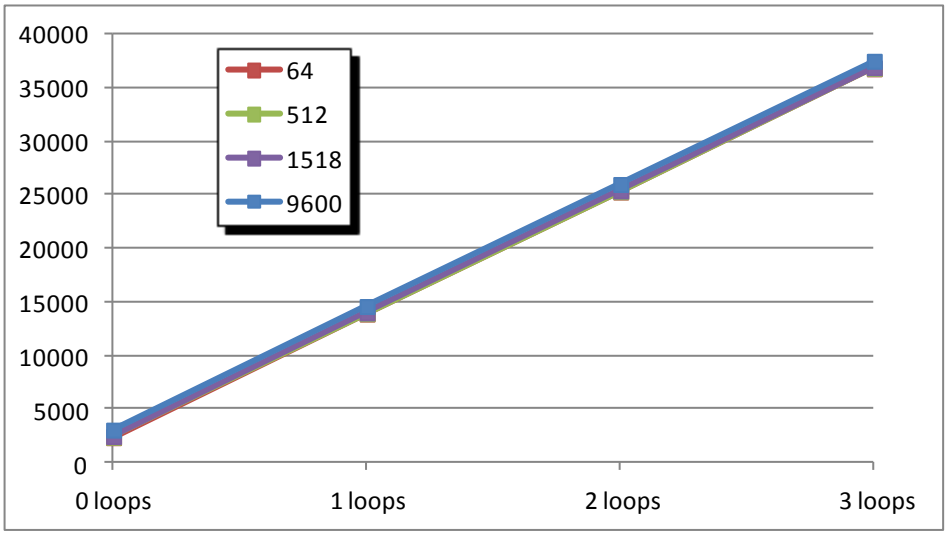


Figure 4.6: OTN measurements: delay in µsec as function of the number of loops for different frame sizes

It is seen from the measurements that the difference in packet size only leads to a fixed offset independent of the number of loops. Thus the offset (710 µsec) is only dependent on the Ethernet-based connection from DTU to NORDUnet, where the OTN switches are located.

Based on the results in Table 4.7, the delay for zero length Ethernet frames can be defined. Furthermore, it is calculated how much the expected values differ from the measured values, normalised to a single uni-directional loop (1077 km).

Loops	0 bytes measured	0 bytes expected	Difference pr unidirectional loop
0	2319 µsec	2319 µsec	0
1	13872 µsec	13329 µsec	271 µsec
2	25277 µsec	24339 µsec	234 µsec
3	36794 µsec	35349 µsec	240 µsec

Table 4.8: Zero-length Ethernet frame delay and difference from expected delays

On average a difference per uni-directional loop of approximately 250 µsec is observed, which corresponds to a fibre length of 50 km. With a combination of low dispersion transmission fibres and dispersion compensation modules, such delay is within the realistic scope. However, no firm confirmation of the cause of the delays can be provided.

4.3.3 Conclusions for the OTN Layer

The OTN transport technology has evolved from a core digital wrapper for SDH signals, providing forward error correction, into a generic transport technology also allowing Ethernet clients. While the OTN OAM allows in-line delay measurements, it was decided to use dedicated test equipment for the delay measurements in these tests.

Using three Alcatel 1830 PSS OTN switches, the delay of the OTN processing has been measured in a lab scenario where the propagation delay was insignificant. Here the OTN processing delay was measured to 40 µsec, of which 32 µsec is assumed to be used by the FEC.

In addition, loops were done in a 1077 km triangle between all three OTN nodes and this study was used to verify the obtained results. This verification did not completely validate the results. However, the differences were explainable and were due to the dispersion compensation delay.

4.4 Packet Layer

This section addresses the delay contributions from the packet layer. The chosen technologies are Provider Backbone Bridge Traffic Engineering (PBB-TE) and Multi-Protocol Label Switching Transport Profile (MPLS-TP) since they are both suitable as Layer 2 protocols, handling the data traffic as packets, even though neither uses the more complex Layer 3 IP Longest Prefix Match (LPM) address lookup.

The technologies are explained and, for PBB-TE, delays are measured through tests.

4.4.1 Technology Overview

4.4.1.1 *MPLS TP*

Overview

Multi-Protocol Label Switching Transport Profile (MPLS-TP) is a new technology developed jointly by the ITU-T and the IETF. The key motivation is to add OAM functionality to MPLS in order to monitor each packet and thus enable MPLS-TP to operate as a transport network protocol.

It is useful to understand some of the key differences between MPLS-TP and classic MPLS.

- Uses network management instead of signalling protocols for determinism (know where your packets are).
- Offline routing.
- Removes functionality that interferes with OAM:
  - Penultimate Hop Popping (PHP) – last hop IP meaning no MPLS OAM end to end.
  - Equal Cost Multi-Path (ECMP) – makes OAM difficult.

- Label Switched Path (LSP) merging – makes OAM difficult.

Applications

MPLS-TP is a packet transport protocol with capabilities that traditionally belong to transport networks such as SONET/SDH and OTN. The intention with MPLS-TP is to be able to replace such legacy networks as SONET/SDH, though not OTN, while still keeping the advantages of packet transport.

Due to its extensive feature catalogue, MPLS-TP is very flexible and can be used for many different applications. One main advantage is that it uses Pseudowire (PW) as a transport entity. PWs are able to encapsulate any type of traffic, such as Asynchronous Transfer Mode (ATM), Frame Relay, Point-to-Point Protocol (PPP), Ethernet, etc.

MPLS-TP is applicable to situations where reliability, QoS and OAM are the main requirements.

MPLS-TP can be operated/controlled via network management or a control plane, the latter being a main advantage when dynamic provisioning is required. Moreover, MPLS-TP is fully compatible with IP/MPLS networks, which presents many possibilities for network solutions that demand MPLS/MPLS-TP interworking.

Delay Contribution

In addition to the functionalities on the lower layers, the MPLS-TP layer contributes to the delay with the following functions:

- Label push.
- Label swap.
- Label pop.
- Label index lookup and retrieval of port and output label information.

The MPLS-TP technology does not require that the data packet is stored in order to do a Cyclic Redundancy Check (CRC); however, if the physical interfacing is based on Ethernet, this will be included anyway.

Mapping to Lower Layer

MPLS-TP can run over Ethernet, SONET/SDH (G.783) and OTN (G.709, G.872) using Generic Framing Procedure (GFP). In these studies OTN was chosen as the underlying layer and thus the delay from the GFP should be included for the MPLS-TP or the OTN layer.

4.4.1.2 *PBB-TE*

Overview

Provider Backbone Bridge Traffic Engineering (PBB-TE) is the third (and latest) standard developed by the IEEE with the aim of giving providers a Layer 2 carrier-grade transport based on classical Ethernet, also known as Carrier Ethernet Transport (CET[1]). The first two technologies of the CET family are Provider Bridges (PB)

[1] Sometimes CET is also used for Ethernet over Multi-Protocol Label Switching (EoMPLS) transport, but this section will use the acronym for the PB/PBB/PBB-TE family only.

and Provider Backbone Bridges (PBB). As PBB-TE (which was developed from the Nortel proprietary technology Provider Backbone Transport (PBT)) re-uses some features of PB and especially of PBB, this section starts with a brief description of these technologies.

PB separates provider VLAN tags from customer tags while PBB goes further and separates MAC addresses as well, by encapsulating a user Ethernet frame into a provider frame. Figure 4.7 shows these two tags and how Ethernet has evolved from classic Ethernet to VLAN tagging and Q in Q to MAC in MAC.

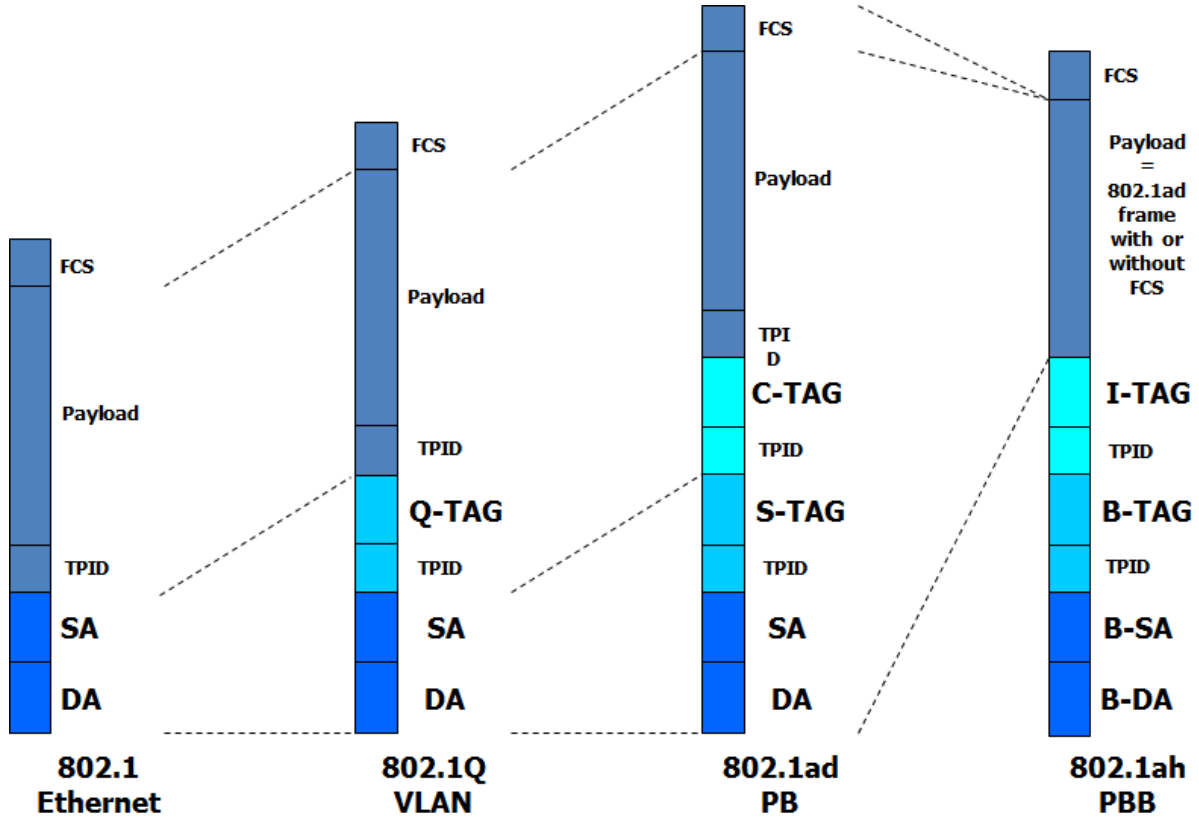


Figure 4.7: Ethernet evolution from classic Ethernet to PBB

The inner tag can be used to define (up to 4094) VLANs in the customer domain. The outer tag can be used for identifying (up to 4094) VLANs and services in the provider domain. These two distinct tags allow a clear separation of the type of services offered in both domains and their transparency.

The Provider Backbone Bridges technique, which is standardised as part of IEEE 802.1ah [IEEE802.1ah], was designed to overcome the main drawbacks of PB. PBB, unlike PB, encapsulates the complete frame of the customer. It adds a tag that includes not just a provider domain VLAN ID (VID), but also a new set of destination and source MAC addresses along with a new service identification tag called I-SID. This makes the provider domain completely separate from the customer domain. As a consequence, protocols such as the spanning tree in the customer network do not interfere with those in the provider domain. Furthermore, not all nodes in the provider domain have to learn all the MAC addresses of the customer network.

Applications

Two types of PBB-TE-based services are defined by IEEE standard 802.1Qay [IEEE802.1Qay]:

- Point-to-point (E-Line type).
- Point-to-multipoint (E-Tree type).

A point-to-point service is supported by two point-to-point Ethernet Switched Paths (ESPs) where the ESPs' end points have the same backbone MAC addresses.

A point-to-multipoint TE service instance is supported by a set of ESPs that comprises one point-to-multipoint ESP from a root to each of *n* leaves plus a point-to-point ESP from each of the *n* leaves to the root. This type of service uses Layer 2 multicast and is aimed at multimedia applications.

It is also possible to build any-to-any TE service based on sets of point-to-point ESPs if the VPLS approach is followed. In such a case, these point-to-point ESPs and point-to-point I-SIDs should create full-mesh connectivity between Backbone Edge Bridges that use this connectivity to send copies of ingress customer frames to a required destination. The split horizon technique, which is used by VPLS, could be applied to avoid loops. However, the 802.1Qay standard doesn't describe any-to-any services, leaving those to proprietary implementations.

PBB-TE services are Layer 2 VPN services; they connect customer sites (pairs or groups) without taking into account any IP information.

PBB-TE offers access and aggregation domains with the same bandwidth and QoS guarantees of a big provider network with an MPLS core. MPLS is seen as a better option for the core networks than PBB-TE for several reasons, the main ones being maturity, tight integration with the IP protocol suite, a diverse control plane, and feature richness.

An example application is mobile backhaul, where PBB-TE offers a cheap and effective way to connect base stations to a central station. (This is of potential interest to the GÉANT community, since the NRENs' ability to offload mobile aggregation is one of the focus points of GN3plus.) PBB-TE is usually seen as a major replacement of TDM-based solutions for mobile operators' backhaul as it provides deterministic paths with several classes of service and bandwidth guarantees, which are very important for transferring delay-sensitive voice traffic.

Delay Contributions

PBB-TE contributes to the delay with the following functions:

- Tagging and Ethernet header lookup. This is the delay experienced by the lookup function that uses the destination address (outer MAC/inner MAC) in combination with the VLAN ID.
- CRC calculation at sender and receiver. This will make the scheme dependent on the frame size for the data flow in question. Long frames require buffering of the full frame for determining whether the frame should be forwarded or dropped depending on the results of the error detection function.

4.4.2 PBB-TE Delay Experiment

In order to verify the delay performance of PBB-TE, access to PBB-TE switches from two vendors is possible, namely TPACK (reference design) at DTU and Ciena at Janet.

4.4.2.1 *Test Objective*

The purpose of the PBB-TE analysis is to include the Layer 2 frame processing in the delay measurements. This is done with and without background traffic. Furthermore, the interoperability studies highlight how the OAM functionalities interwork.

4.4.2.2 *Simple PBB-TE Only Test*

Test Setup

This setup is entirely located at DTU where the N2X test equipment is connected to a TPACK PBB-TE switch which is virtualised into five independent PBB-TE switches. These are then connected through patch cables; the maximum number of switches is constrained by the number of output ports. The setup is illustrated in Figure 4.8.

As a result the delay through a PBB-TE (vendor specific) switch can be quantified.

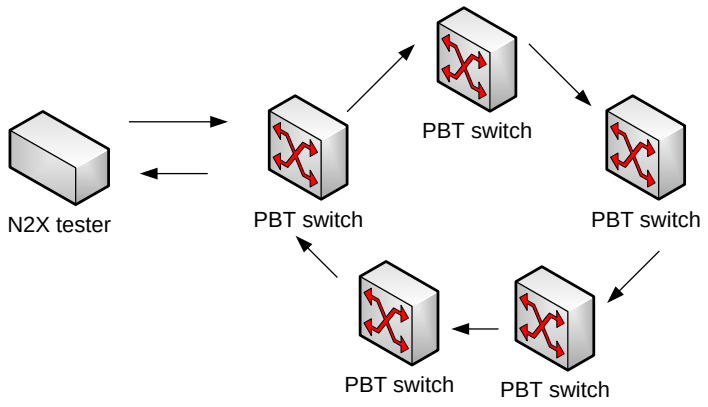


Figure 4.8: PBB-TE at DTU – all-in-one device virtualised into several switches

The PBB-TE measurements were done using an Agilent 1GbE N2X test solution and only insignificant fibre delay between the Ethernet ports.

Test Results

The results are shown in Figure 4.9 for 64, 512 and 1500 bytes Ethernet frames. In addition, results were obtained with 9600 bytes Jumbo frames; the delay was in linear proportion to that shown in the figure, thus a delay of 80 µsec was observed per switch for Jumbo frames, slightly depending on the load.

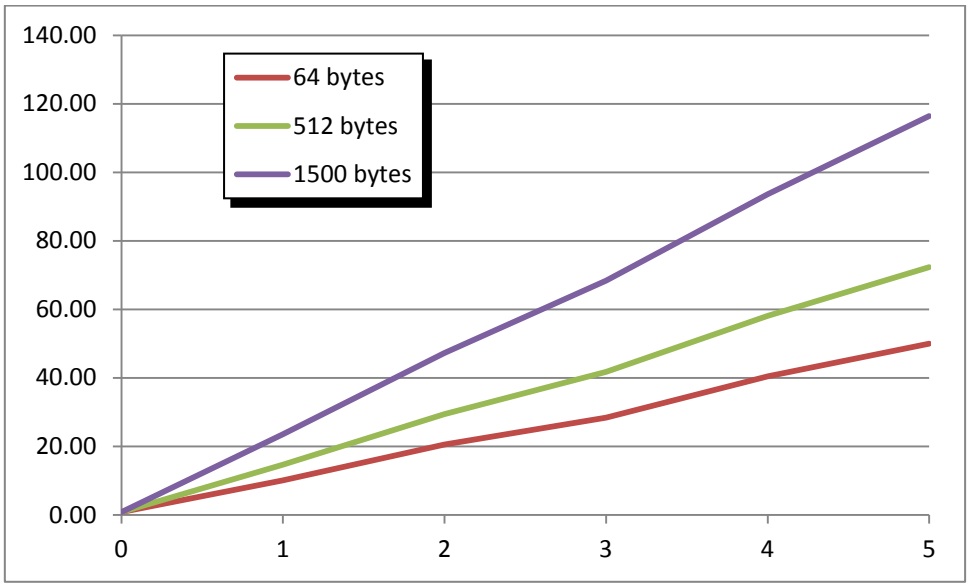


Figure 4.9: PBB-TE delay measurements showing delay in µsec as function of the number of PBB-TE switches

As expected, the delay depends heavily on the number of PBB-TE switches and the frame size. This is due to the store and forward architecture of Ethernet. These figures are expected to be reduced by a factor of approximately 10 if 10GbE is used.

4.4.2.3 *Ciena PBB-TE Test*

Test Objective

The purpose of the PBB-TE trial on the Janet Ciena-based testbed was to find out to what extent PBB-TE operations increase the latency of Ethernet streams compared to plain traditional Ethernet switching based on a single VLAN tag.

The extra PBB-TE operations (encapsulation/decapsulation into/from the extended PBB-TE frame), which contribute to the longer switching time of a PBB-TE frame compared to the plain switching of a single-tagged Ethernet frame, were considered under *Delay Contributions* in Section 4.4.1.2. However, the absolute values of these extra operations could be quite different and, if they are a small percentage of the whole frame forwarding time, then PBB-TE could be considered as a good candidate for the transport of time-sensitive applications. Otherwise, PBB-TE's advantages, such as fast protection switching and traffic engineering, would not outweigh the additional delays introduced by the extra operations.

Test Setup

A testbed built on two Ciena 5305 switches was used to investigate whether the extra operations of PBB-TE result in the significant time difference between PBB-TE and plain Ethernet switching (Figure 4.10). The Ciena 5305 model implements PBB-TE-related operations (frame encapsulation/decapsulation) in the hardware (in the port's FPGA), so it was expected that the PBB-TE latency overheads would not be very high and that the PBB-TE switching time would not exceed plain Ethernet switching time significantly.

To measure switching latency, the SunRise RxT tester was used. The tester supports the Y.1564 specification for Ethernet services assurance testing, which includes measurement of the frame round-trip time (RTT) for different frame sizes and stream rates. To exclude delays caused by congestion, the stream rate was 100 kbps in all tests.

To evaluate the effect of extra PBB-TE operations with a good level of precision, the two switches and the tester were connected and configured in such a way that a frame travels 8 hops on its route through the testbed. As a result, variations of RTTs were multiplied by a factor of 8 and hence were easier to measure with appropriate precision.

Two patches (between ports 7/21 and 7/22 of the Reading and London switches) and the VLAN ID translation technique helped to loop traffic from the tester around the switches, so that to reach Port 3 of the tester Ethernet frames had to make two full round trips through the switches' ports.

Two different ranges of VLAN IDs were used for transmitting PBB-TE and plain VLAN traffic: range 1901–1904 for PBB-TE, and range 1801–1804 for VLAN traffic. Accordingly, two separate sets of virtual switches were created to forward PBB-TE and VLAN frames: switches vs 1901 to vs 1904 for PBB-TE, and switches vs 1801 to vs 1804 for VLAN traffic.

The dashed coloured lines in Figure 4.10 illustrate the path of PBB-TE traffic between tester ports. The red line shows the initial segment of the path when single-tagged frames with VLAN ID 1901 traverse port 7/7 of the Reading switch and are forwarded by virtual switch vs 1901 of the Reading switch to the encapsulating end of the PBB-TE tunnel between the Reading and London switches. After adding the PBB-TE header field, the frames are carried along the tunnel (on the basis of the I-SID field value playing the role of VLAN ID for PBB-TE traffic). The decapsulating end of the PBB-TE strips all the fields of the PBB-TE header and sends the original frame to virtual switch vs 1901 of the London switch, which translates the VLAN ID 1901 to 1902 and sends the frame through a physical patch to virtual switch vs 1902, as the green line shows. The further frame transformations are similar to those described above and, as a result, the frame arrives at Port 3 of the tester after crossing the PBB-TE tunnel 4 times, undergoing 4 operations of the PBB-TE encapsulation and 4 operations of the PBB-TE decapsulation.

Port 3 of the tester is set up in the loopback mode (according to the 802.3ah standard) and hence re-directs incoming frames back by swapping MAC and IP destination and source addresses. As a result, the PBB-TE frames do another 4 trips around the Ciena switches and finally arrive at Port 2 of the tester having traversed the PBB-TE tunnel 8 times.

The path of the VLAN-based frames is similar but with the difference that virtual switches vs 1801 to vs 1804 direct frames to physical ports 7/2 and 7/1 instead of to the encapsulation ends of the PBB-TE tunnel. Therefore traffic for VLAN 1801 is forwarded through the testbed along the same physical path as the PBB-TE frames but without the use of PBB-TE operations. The number of hops for VLAN-based traffic is the same, i.e. 8.

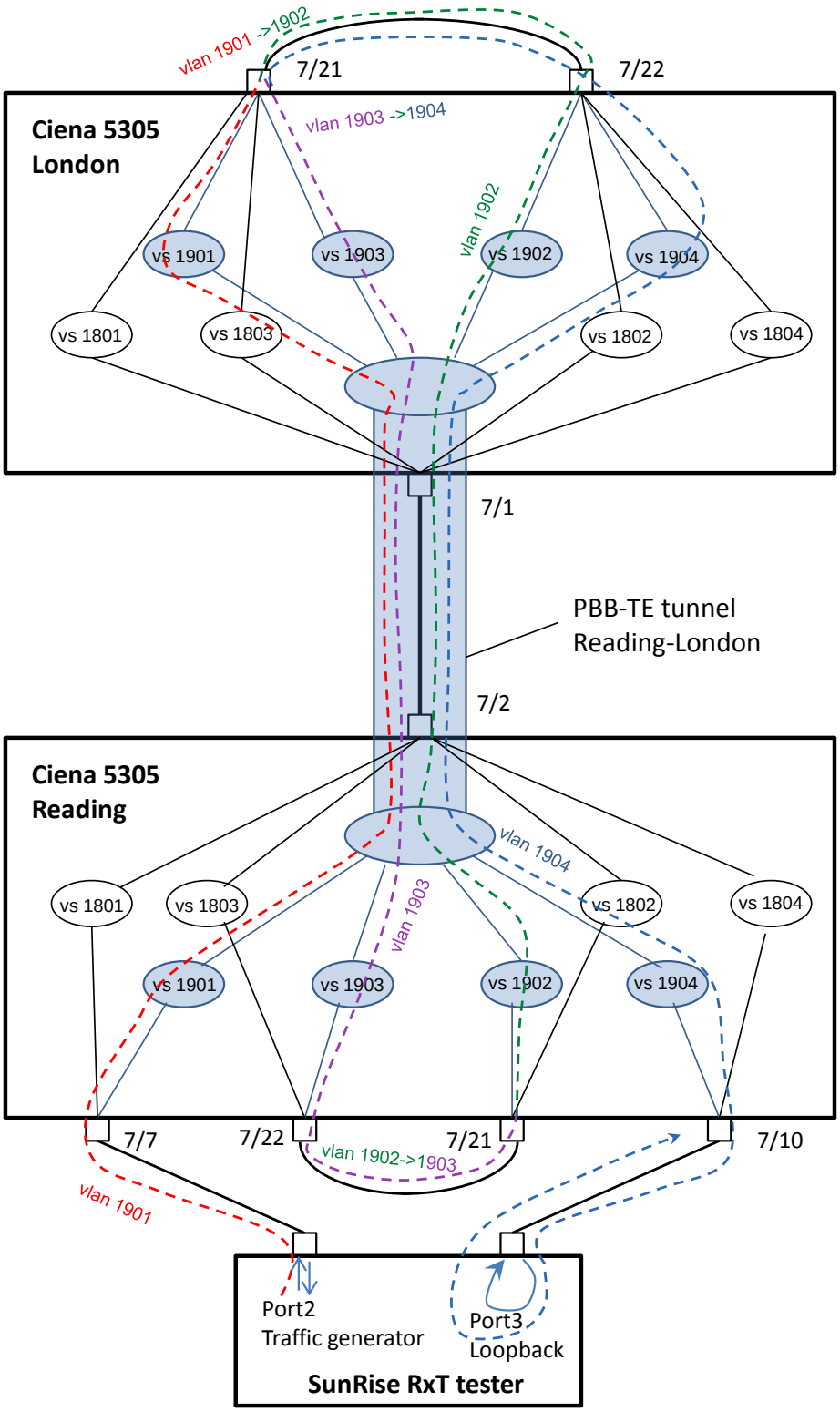


Figure 4.10: The Janet Ciena PBB-TE testbed

Test Results

The results of RTT times for both PBB-TE and VLAN-based traffic are presented in Table 4.9 and in Figure 4.11.

Frame size (Bytes)	PBB-TE (µs)	Plain VLAN switching (µs)
84	271	261
512	375	365
1200	523	514
1500	586	575
2000	695	685

Table 4.9: RTT times

These results were taken from RxT tester reports and hence reflect accumulated frame delays after 8 hops around the looped path; for the one-hop delay, the numbers should be divided by 8.

The RTT times for both PBB-TE and VLAN-based traffic were relatively stable, with the average delay variation no more than 2 µs, so the extra latency related to the PBB-TE operations can be evaluated as 10 µs for 8 hops or 1.25 µs per hop (switch). It is a relatively small PBB-TE contribution and could be considered as negligible compared to the overall delay values in range of (271–695) µs per 8 hops or (34–87) µs per switch.

The extra PBB-TE delay does not depend on the frame size, as the test results show; this is quite understandable, as PBB-TE operations are done on already-buffered frames.

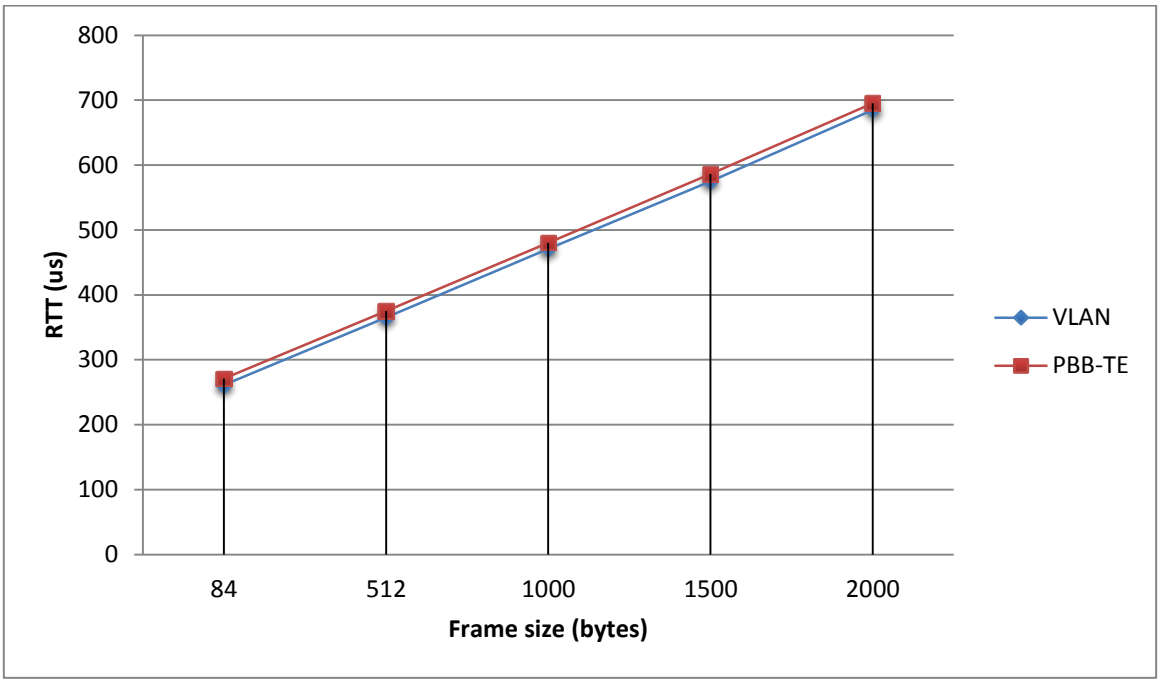


Figure 4.11: RTT of PBB-TE and VLAN-based frames for different frame sizes

4.4.3 Conclusions for PBB-TE

In this section, PBB-TE and, to a minor extent, MPLS-TP have been discussed with respect to the delay-contributing factors. As expected, the delay in PBB-TE switches is directly dependent on the frame length of the Ethernet frames due to the store and forward. It is expected the same behaviour would be present for MPLS-TP, as the physical interface is also usually Ethernet.

The measurements were done with equipment from two different vendors. In one measurement the delay was observed as a function of the number of nodes and the frame size and significant delays can be observed for Jumbo frames of 9600 bytes. Here the processing per node equals 4–5 km of fibre delay.

The overhead going from Ethernet VLAN tagging to PBB-TE tagging was also measured, using Ciena devices at Janet, and an additional delay of 10 µsec was found. This delay was independent of the frame size because the buffer is already stored for evaluating the frame check sequence.

4.5 Hybrid Solutions

4.5.1 Technology Overview

The fusion networking technology was first introduced in academic literature in 2003 under the name integrated hybrid networks. In 2012, the first commercial fusion product was launched in the market, the TransPacket H1.

The circuit transport in fusion is named Guaranteed Service Transport (GST) and the packet transport allows Statistical Multiplexing (SM) and is therefore called SM transport. The aim of the GST transport is to provide an Ethernet transport with performance comparable to the OTN (G.709) transport. The result is a network with a unique combination of properties: GST circuit paths with a low processing overhead, zero packet loss, low delay and minimised Packet Delay Variation (PDV) combined with packet paths with SM capability enabling the throughput efficiency of packet networks. That is, the network has:

- Predictable low and constant latency.
- No packet loss caused by contention.
- Full transparency also on timing, allowing packet layer synchronisation packet transport.
- Isolated sub-wavelength circuits enabling privacy and performance independence.

The SM transport provides an Ethernet transport comparable to Best Effort transport in a packet switched network, namely:

- High network throughput efficiency due to statistical multiplexing.
- Applications range from video and voice services to email and browsing as the Quality of Service level depends on dimensioning of the network and the total network load.

Integrated hybrid networks, also known as fusion networks [Palacios, Gauger], fuse the circuit and packet network in a time-interleaved manner without using time-slots. A statistical multiplexing (SM) best-effort packet scheduler looks into a guaranteed service transport (GST) circuit stream of packets, detecting vacant gaps between the GST packets. Whenever a gap is detected, the scheduler checks whether an SM packet is available that has a size that fits the gap. If so, the packet is inserted in the gap without affecting the timing of the packets in the GST stream [Palacios].

An example schematic diagram of a hybrid node is presented in Figure 4.12. The node has two 10 Gb/s Ethernet (10GE) interfaces for the wavelength transport channel. Four Gigabit Ethernet (GE) interfaces are applied to increase the channel utilisation by adding SM traffic. Packets entering a hybrid node are tagged with a VLAN ID indicating the type of service. Any SM packets received at a 10GE interface are dropped to one of the GE interfaces. GST packets received at a 10GE interface are passed through to the other 10GE interface with absolute priority. The bypassing packets are given a fixed delay δ corresponding to the duration of a maximum-length SM packet. This delay ensures that:

- Any ongoing scheduling of SM packets is allowed to finish without interfering/being pre-empted by the GST packets arriving at the input channel.
- The gap length between GST packets is detected.

The gap length is applied by the SM packet scheduler, which searches the head of queues Q1–Q4 for a packet smaller than the detected gap. If a packet is found, it is scheduled. If there is leftover space in the gap, the scheduler proceeds until the gap is full, no suitable SM packet is found, or the SM queues are empty.

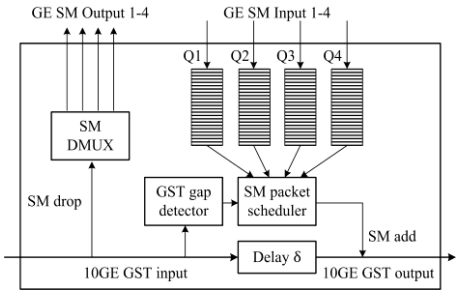


Figure 4.12: Schematic diagram of the uni-directional transport through the fusion add/drop muxponder when using four of the 10 inputs

The scheduling mechanism of the fusion network is illustrated in Figure 1.13. The figure shows how the GST packets are left untouched whilst the SM packets are inserted in spare slots between the GST packets, avoiding PDV and packet loss on the GST packets. Packets forwarded as GST are forwarded through the network along a circuit inheriting circuit performance properties: a low fixed latency and no packet loss due to contention. SM packets are scheduled onto the same line as the GST packets; however, an SM packet is only scheduled when a vacant gap with a suitable size is found between the GST packets. The GST packets, as well as the size of the gap between the GST packets, are then left untouched, enabling a full transparency on the GST paths, also with respect to timing. SM packets, on the other hand, wait in buffer queues until a suitable gap is found into which the SM packet can be scheduled. SM packets are therefore forwarded through the network with variable latency, and if the mean load is high due to a high extent of oversubscription, packets will be lost, just like in a packet switched network.

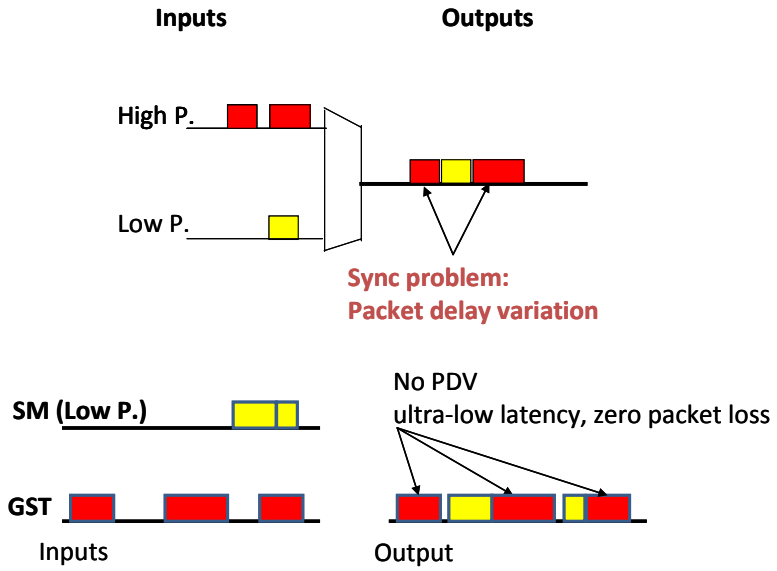


Figure 4.13: QoS and PVD in switches/routers vs fusion-based device (switch)

The upper half of Figure 4.13 shows strict priority QoS scheduling in packet switches and routers. If two packets of low and high priority arrive simultaneously, the low-priority packet will be scheduled as soon as the high-priority queue is empty. If a high-priority packet arrives during scheduling of a low-priority packet, it will be

delayed until the low-priority packet has been scheduled. Hence PDV occurs on high-priority packets. The lower half of the figure illustrates the fusion scheduling where low-priority SM packets are inserted only if there is a vacant gap between the high-priority GST packets. Hence, PDV and packet loss are avoided on GST packets.

4.5.2 Hybrid Solution Experiments

4.5.2.1 *Test Objective*

The purpose of the tests was to demonstrate that:

- The fusion network is able to emulate router offload and transport GST traffic with circuit QoS and independently of the insertion of statistically multiplexed SM traffic.
- The efficient add/drop of SM traffic on the lightpath used for router bypass of GST traffic increases the network throughput significantly.

4.5.2.2 *Test Setup*

The field-trial setup uses three prototype Fusion Ethernet nodes from TransPacket, in the carrier network of UNINETT. Two experiments (connectivity scenarios) were set up.

Experiment 1

The setup for the first experiment is shown in Figure 4.14. In order to measure the GST traffic performance, one Gigabit Ethernet (GE) GST stream was generated from the Spirent packet generator/tester. The traffic:

1. Was added at node 1 (H1_1).
2. Was forwarded towards node 2 (H1_2).
3. Bypassed H1_2.
4. Was received on node 3 (H1_3).

The packet length was proportional to the Ethernet frame length and was uniformly distributed between 64 and 1518 bytes. The average GST load on the GE interface was set to 0.99 by fixing the inter-packet length. The GST performance was measured for different network loads by changing the load of nine added SM streams up until network congestion. The measurements were conducted up to a total load of 0.99 on the 10 Gigabit Ethernet (10GE) lightpath.

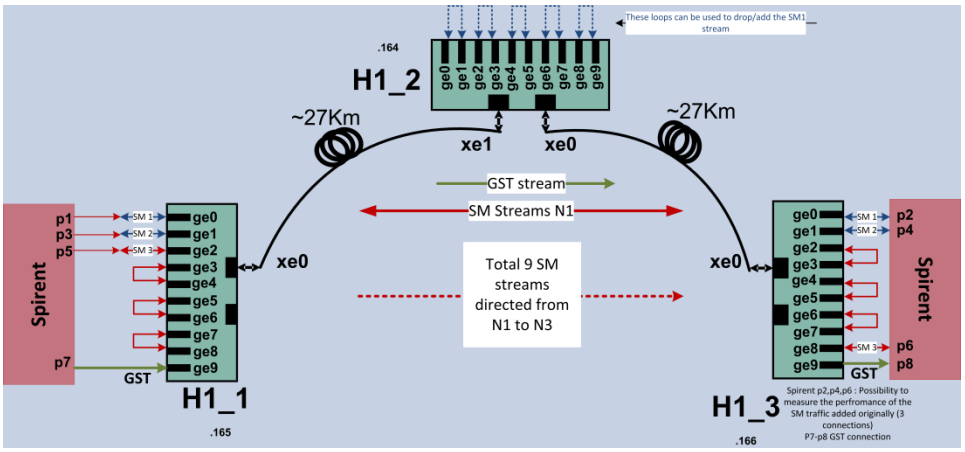


Figure 4.14: Setup for experiment 1

The 1Gb/s Ethernet GST stream is generated by the Spirent SPT-2000 packet generator/tester, added at node H1_1, bypassing H1_2 and dropped at H1_3. The stream is then measured by the tester. The SM traffic is added on the other GE interfaces of the nodes without affecting the GST stream and with a total lightpath load up to saturation.

Experiment 2

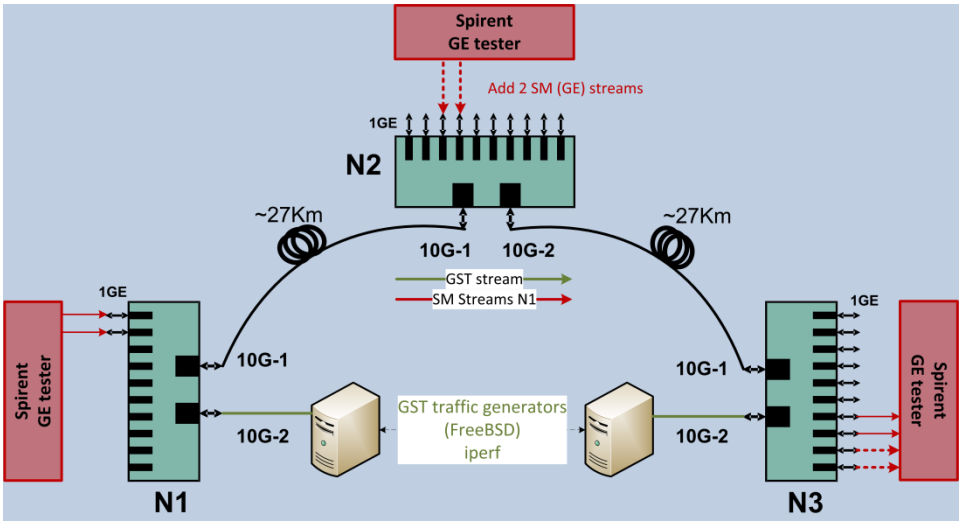


Figure 4.15: Setup for experiment 2

Figure 4.15 shows the fusion network setup for the second experiment with: three prototype Fusion nodes from TransPacket, two FreeBSD servers with Iperf traffic generators and a Spirent SPT-2000 packet generator/tester. Xe0 and xe1 are 10GE interfaces while ge0-ge9 are GE interfaces. Streams 1 and 2 are added on the first node and destined for node 3. The second node treats these streams as GST; hence they bypass the intermediate node. The second node adds two streams 3 and 4 with destination node 3.

4.5.2.3 *Test Results*

The results for the average packet delay and packet loss ratio (PLR) for GST and SM traffic are shown respectively in Table 4.10 and Table 4.11. The system load is normalised for 1 Gb/s links.

System Load	Average delay	Average jitter	Max jitter	Packet loss
[0.99, 9.9]	311.66 (µs)	30 (ns)	140 (ns)	0

Table 4.10: GST performance measurements

System Load	Avg. delay (µs)	PLR
5.49	303.36	0
7.29	304.58	0
9.09	307.75	0
9.09	307.75	0
9.54	311.5	0
9.63	313.71	0
9.675	319.31	0
9.693	30351.34	3.64e-4
9.702	30853.01	1.89e-3
9.72	31186.66	5.07e-3
9.81	60450.43	1.77e-2
9.9	59566.98	3.33e-2

Table 4.11: SM performance measurements

The experiment results confirm that:

- The GST traffic is transported through the network with absolute priority: no GST packet losses are observed and the average GST delay remains constant regardless of the network condition/congestion. Hence GST has circuit QoS and is not affected by the SM insertion.
- SM insertion increases the 10GE lightpath utilisation up to 97% without any losses.
- The SM PLR is 1e-02 at the total lightpath load of 0.99. Hence, the network performs as a saturated statistical multiplexing packet network with high utilisation while providing a service with circuit QoS properties.

The propagation delay on the two fibre links connecting the nodes was measured to a total of 266.11 µs. The nodal processing delay of the SM packets added was measured to a reference value of 2.37 µs. The GST

average end-to-end delay was 311.66 µs through all the measurements. Hence, the added delay by the nodes on the GST path is 45.55 µs. The packet delay variation (PDV) peak value is 130 ns and on average 30 ns. The experiment demonstrated for the first time 1GE to 10GE GST traffic insertion and the measured delay at the load 0.99 was 26 µs. The results show that the fusion network enables transmission of packet layer synchronisation traffic in GST, while enabling lightpath utilisation of 96.9% without any packet losses when inserting statistically multiplexed SM traffic.

4.5.3 Conclusions for Hybrid Networks

The UNINETT field trial demonstrated the feasibility of integrated hybrid networking through a network of Fusion Ethernet-based prototype nodes from TransPacket. The results demonstrate a circuit quality of service performance for the GST transport: zero packet loss, a node delay much lower than the fibre transmission delay, and with a packet delay variation in the nanosecond range. The circuit transport (GST) has a low processing overhead and enables efficient transparent router bypass. Its quality is higher than for routers, enabling transport of time-sensitive traffic and packet-layer synchronisation information. High throughput efficiency was demonstrated by adding packet-switched statistical multiplexed traffic on the common circuit/packet lightpath. Circuit traffic was then not affected, even at maximum lightpath utilisation of 99% with 10% circuit traffic.

4.6 Conclusions

In this chapter, a selection of transport technologies has been evaluated for transporting time-sensitive services. The objective is to quantify the inherent delays in the technologies so that a user can use the results to estimate delays in a given transport network environment. The evaluated technologies are OTN, PBB-TE and a hybrid packet-circuit solution.

As a first step, to understand the overall network latency, the latency introduced by the physical layer of a DWDM network was examined. The main sources of latency in a transport network are the transmission fibre, fibre-based DCMs and transponders (and regenerators).

Using three Alcatel 1830 PSS OTN switches, the delay of the OTN processing has been measured in a lab scenario where the propagation delay was insignificant. Here the OTN processing delay was measured to 40 µsec of which 32 µsec is probably used by the FEC. In addition, loops were done in a 1077 km triangle between all three OTN nodes and this study was used to verify the obtained results.

Then the delay in PBB-TE switches was measured and was found to be directly dependent on the frame length of the Ethernet frames due to the store and forward. The measurements were done with equipment from two different vendors. The overhead going from the Ethernet VLAN tagging to PBB-TE tagging was also measured and an additional delay of 10 µsec was found.

Trials with Fusion Ethernet-based prototype nodes from TransPacket demonstrated a circuit quality of service performance for the guaranteed data: zero packet loss, a node delay much lower than the fibre transmission delay and with a packet delay variation in the nanosecond range. The circuit transport (GST) has a low

processing overhead and enables efficient transparent router bypass. The most significant advantage of such a device in the NREN environment is the ability to deliver sub-wavelengths like Ethernet services with low latency and packet-delay variation for time-sensitive applications.

5 Final Conclusions

With this report the work of JRA1 T1 in the GN3 project comes to an end. The Task's work has covered the study of the Transport Network Technologies most relevant to the NREN community and has provided information and test results that the JRA1 T1 team considers relevant for building future networks. The extensions for GN3 Y4 covered testing in the following areas, as described in this report:

- MPLS-TP/T-MPLS.
- EoMPLS interoperability.
- Time-sensitive transport technologies.

The MPLS-TP/T-MPLS testing was very useful to understand the current status of T-MPLS in the Alcatel-Lucent 1850 TSS-320 and the usability of the NMS to configure services, LSPs and pseudowires. The testing of T-MPLS/MPLS-TP architecture and Ethernet services gave positive results. The OAM features available in the release that was tested were based on the ITU-T Y.1731 standard, while other implementations in the market are based on IETF tools such as LSP ping and BFD. This duality of protocols makes interoperability impossible unless operators are given the option of choosing between both protocols.

Overall the testing was very useful for gaining experience with T-MPLS/MPLS-TP technology and for becoming familiar with protection and OAM functionality in packet-based networks. However, the implementations need to become more mature, with the introduction of a control plane and the full set of OAM features. A more agile NMS in combination with a CLI would be the best solution for allowing operators with different skill sets to choose the best option for their profile and their needs.

Finally MPLS-TP needs to be fully interoperable with MPLS in order to achieve seamless connectivity of services across domains.

Ethernet over MPLS point-to-point (VPWS) and multipoint (VPLS) services were tested in a multi-vendor network built from Juniper Networks and Cisco Systems routers. The tests aimed to verify the interoperability of the two vendors' Ethernet over MPLS implementations in three areas:

- The basic interoperability of the service implementation, allowing the provisioning of Ethernet over MPLS services in a multi-vendor network.
- Use of Ethernet OAM mechanisms (Continuity Check Protocol, loopback, linktrace) to monitor and troubleshoot an Ethernet over MPLS service in a multi-vendor network.

- Fast recovery of an Ethernet over MPLS service by using the Fast Reroute mechanisms provided by the MPLS transport plane in a multi-vendor network.

Apart from the minor issues mentioned in the report, the results of the tests listed above show that the implementations of the tested services available on Juniper Networks and Cisco Systems routers are interoperable and that routers from the two vendors can be used together in a single network that provides Ethernet over MPLS services.

Selected time-sensitive applications within telemedicine, edutainment and finance put severe requirements on the transport technologies in the network. Hence, a number of technologies for Layer 1 and 2 have been evaluated to identify the inherent sources of delay. The technologies include DWDM for the physical layer, OTN as a circuit-switched transport technology, PBB-TE for the packet layer, and a special hybrid packet-circuit solution.

The delays imposed by the technologies were measured experimentally, and for medium- to long-range communication the fibre propagation delay is – as expected – the dominating factor. The standard delay in OTN was measured for a specific vendor, and a delay of 40 µsec was recorded due to the FEC processing. For PBB-TE the delay varied from 10 µsec to 80 µsec depending on the frame length, and a 10 µsec delay was observed going from VLAN tagging to PBB-TE tagging. The hybrid test showed circuit quality service performance for the prioritised data.

JRA1 T1's study of transport network technologies has identified technologies and features that might be used in the GÉANT and NREN community. 100G technology has become a reality and there are already implementations in many European networks, including GÉANT [GN3PRCERN100G, GN3PRxatlantic100G]. OTN will provide the switching and multiplexing capabilities needed to groom 10G and lower capacity into 100G pipes, a concept applied in the GÉANT network, for example [InfineraPRGÉANT]. Moreover OTN integrates OAM capabilities in combination with a control plane, allowing restoration and dynamic provisioning. IP/MPLS networks will be introducing MPLS/TP features as they become available, which will allow NRENs to provide better and more reliable services. Finally, Ethernet OAM will allow NRENs to provide end-to-end services with better troubleshooting and monitoring capabilities.

References

[Briggs] P. Briggs, R. Chundury, J. Olsson, “Carrier Ethernet for mobile backhaul”, IEEE Communication Magazine, Vol. 48(10), pp. 94–100 (2010)

[CONNECT10Jan2013] A. Colmenero, “Shaping the Future of NRENs”, CONNECT, Issue 10 January 2013 http://issuu.com/danteprm/docs/connect_january_2013?mode=window&proSidebarEnabled=true&embedId=6131560/1220277

[Cosart] L. Cosart, “NGN packet network synchronization measurement and analysis”, IEEE Communication Magazine, Vol. 49(2), pp.148–154 (2011)

[DJ1.1.1] A. Colmenero, R. Corn, M. Garstka, J. Kloots, U. Monaco, V. Olifer, J. Radil, K. Stanecki, S. Tyley, “Deliverable DJ1.1.1: Transport Network Technologies Study” www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-09-224-DJ1-1-1v1-0_Transport_Network_Technologies_Study_Read_Only.doc

[DJ1.1.2] K. Bozorgebrahimi, M. Channegowda, A. Colmenero, A. Manolova Fagertun, M. Garstka, R. Lund, V. Olifer, J. Radil, “Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing” http://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-12-100_DJ1-1-2_Transport-Network-Technologies-Study-and-Testing.pdf

[Ferrant] J. L. Ferrant, M. Gilson, S. Jobert, M. Mayer, M. Ouellette, L. Montini, S. Rodrigues, S. Ruffini, “Synchronous Ethernet: a method to transport synchronization,” IEEE Communication Magazine, Vol. 46(9), pp. 126–134 (2008)

[Gauger] C. M. Gauger, J. P. Kuhn, E. V. Breusegem, M. Pickavet, P. Demeester, “Hybrid optical network architectures: bringing packets and circuits together,” IEEE Comm. Magazine, Vol 44(8), pp. 36-42 (2006)

[GN3PRCERN100G] “CERN gets first GÉANT 100Gbps link for Christmas”, Press Release, Press Release, 21 December 2012 http://geant3.archive.geant.net/Media_Centre/News/Pages/CERN_first_user_of_GEANT_terabit_network.aspx

[GN3PRxatlantic100G] “Internet2, NORDUnet, ESnet, SURFnet, CANARIE, and GÉANT to build world’s first 100G intercontinental transmission links for research and education community”, Press Release, 24 April 2013 http://www.geant.net/MediaCentreEvents/news/Pages/first-transatlantic-link.aspx

[IEEE802.1ah] IEEE Standard for Local and metropolitan area networks — Virtual Bridged Local Area Networks. Amendment 7: Provider Backbone Bridges

[IEEE802.1Qay] IEEE Standard for Local and metropolitan area networks — Virtual Bridged Local Area Networks. Amendment 10: Provider Backbone Bridge Traffic Engineering

[IETF-RFC6671] IETF RFC 6671, M. Betts, “Allocation of a Generic Associated Channel Type for ITU-T MPLS Transport Profile Operation, Maintenance, and Administration (MPLS-TP OAM)”, November

	2012 http://tools.ietf.org/html/rfc6671
[InfineraPRGÉANT]	"DANTE Selects the Infinera DTN-X Platform for GÉANT Multi-Terabit Pan-European Research and Education Network", Press Release, 23 May 2012 http://www.infinera.com/j7/servlet/NewsItem?newsItemID=306
[ITU-T G.8113.1]	G.8113.1/Y.1372.1 (ex G.tpoam G.mplstpoam) "Operations, Administration and Maintenance mechanism for MPLS-TP in Packet Transport Network (PTN)" http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=7196
[ITU-T G.8113.2]	G.8113.2/Y.1372.2 (ex G.tpoam G.mplstpoam) "Operations, administration and maintenance mechanisms for MPLS-TP networks using the tools defined for MPLS" http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=8148
[ITU-T Newslog1]	Unopposed approval of MPLS-TP carrier network standards, 21 November 2012 http://www.itu.int/ITU-T/newslog/Unopposed+Approval+Of+MPLSTP+Carrier+Network+Standards.aspx#.UTdAshlneN0
[Jay]	John A. Jay, "Low Signal Latency in Optical Fiber Networks", Corning Optical Fiber
[MPLS-TP_facts]	http://www.itu.int/ITU-T/newslog/MPLSTP+The+Facts.aspx
[NORDUnet2012TSTT]	H. Wessing, "Technology evaluation for time-sensitive data transport", NORDUnet Conference, September 2012 https://events.nordu.net/display/ndn2012web/Technology+evaluation+for+time+sensitive+data+transport
[Palacios]	J. P. Fernández-Palacios, L. Perez, J. Rodriguez, J. Dunne, M. Basham, "IP off-loading over multi-granular photonic switching technologies," in Proc. of European Conference and Exhibition on Optical Communication (ECOC), 2010
[TechAnnex]	GN3 Project "Annex I – Description of Work" [restricted access] \\chfile01\GEANT\GN3\GN3-2009-001-099\GN3-09-004v3.5 GN3 Technical Annex - DoW - Annex I.pdf
[TransPacket]	http://www.transpacket.com/
[WTSA12Blog]	Unopposed approval of MPLS-TP carrier network standards, 21 November 2012 http://wtsa12.wordpress.com/2012/11/21/unopposed-approval-of-mpls-tp-carrier-network-standards/

Glossary

AIS	Alarm Indication Signal
ALU	Alcatel-Lucent
APS	Automatic Protection Switching
AS	Autonomous System
ATM	Asynchronous Transfer Mode
BFD	Bi-directional Forwarding Detection
BGP	Border Gateway Protocol
BTV	Broadcast TV
CCDB	Continuity Check Database
CCM	Continuity Check Message
CCTNT	Carrier Class Transport Network Technology
CE	Customer Edge (the customer device facing the provider's network)
CET	Carrier Ethernet Transport
CFM	Connectivity Fault Management protocol
CLI	Command Line Interface
CRC	Cyclic Redundancy Check
CV	Connectivity Verification
DCF	Dispersion Compensation Fibre
DCM	Dispersion Compensation Module
DeIC	Danish e-Infrastructure Cooperation
DM	Delay Measurement
DMM	Delay Measurement Message
DMR	Delay Measurement Reply
DSP	Digital Signal Processing
DTU	Technical University of Denmark
DWDM	Dense Wavelength-Division Multiplexing
ECMP	Equal Cost Multi-Path
EDFA	Erbium-Doped Fibre Amplifier
E-LAN	Ethernet LAN
E-Line	Ethernet Line
EoMPLS	Ethernet over MPLS
EPL	Ethernet Private Line
EP-LAN	Ethernet Private LAN
EP-Tree	Ethernet Private Tree
ESP	Ethernet Switched Path
E-Tree	Ethernet Tree

EVC	Ethernet Virtual Connection
EVPL	Ethernet Virtual Private Line
EVP-LAN	Ethernet Virtual Private LAN
EVP-Tree	Ethernet Virtual Private Tree
FDM	Frame Delay Measurement
FGB	Fibre Bragg Grating
FDI	Forward Defect Indication
FEC	Forward Error Correction
FIB	Forwarding Information Base
FPGA	Field-Programmable Gate Array
FW	Firmware
G	Gigabit
Gbit/s	Gigabit per second
GE	Gigabit Ethernet
GFP	Generic Framing Procedure
GMPLS	Generalised Multi-Protocol Label Switching
GST	Guaranteed Service Transport
IANA	Internet Assigned Numbers Authority
ID	Identifier
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
ILA	**I**n-Line Amplifier
IOS	Internetwork Operating System
IP	Internet Protocol
I-SID	Ethernet Service Instance Identifier
IS-IS	Intermediate System to Intermediate System protocol
ITU-T	International Telecommunication Union – Telecommunication Standardisation Sector
JRA1	GN3 Joint Research Activity 1, Future Network
JRA1 Task 1	JRA1 Task 1, Carrier Class Transport Network Technologies
JRA2	GN3 Joint Research Activity 2, Multi-Domain Network Service Research
JRA2 Task 3	JRA2 Task 3, Monitoring
L2VPN	Layer 2 Virtual Private Network
LAN	Local Area Network
LB	Loopback
LDP	Label Distribution Protocol
LOC	Loss of Continuity [or Connectivity – waiting for confirmation]
LOS	Loss of Signal
LPM	Longest Prefix Match
LSP	Label Switched Path
LT	Linktrace
MA	Maintenance Association
MAC	Media Access Control
MD	Maintenance Domain
MEF	Metro Ethernet Forum
MEP	Maintenance End Point
MEP ID	Maintenance association End Point Identifier

MIP	Maintenance association Intermediate Point
MPLS	Multi-Protocol Label Switching
MPLS-TP	MPLS Transport Profile
m/s	Metres per Second
NMS	Network Management System
NNI	Network-to-Network Interface
NREN	National Research and Education Network Glossary
OADM	Optical Add-Drop Multiplexers
OAM	Operation, Administration and Maintenance
OAM&P	Operation, Administration, Maintenance and Provisioning
ODU	Optical Data Unit
ODUk	ODU of any of the signal types 1/2/2e/3/3e2/4
OEO	Optical-Electrical-Optical
OMS	Optical Management System
OTN	Optical Transport Network
P	Provider router (a router in the core of the provider's network)
PB	Provider Bridges
PBB	Provider Backbone Bridges
PBB-TE	Provider Backbone Bridge Traffic Engineering
PBT	Provider Backbone Transport (Nortel proprietary technology from which the standard PBB-TE technology was developed by the IEEE)
PDU	Protocol Data Unit
PDV	Packet Delay Variation
PE	Provider Edge router (the provider's router to which customer equipment is connected)
PHP	Penultimate Hop Popping
PLR	Packet Loss Ratio
PPP	Point-to-Point Protocol
PSC	Protection State Coordination
PTN	Packet Transport Network
PW	Pseudowire
PTP	Precision Time Protocol
QoS	Quality of Service
RDI	Remote Defect Indicator
RSVP	Resource Reservation Protocol
RTT	Round-Trip Time
SA	GN3 Service Activity
SDH	Synchronous Digital Hierarchy
SE	Synchronous Ethernet
SFEC	Signal processing-based FEC
SM	Statistical Multiplexing
SONET	Synchronous Optical Networking
TDM	Time Division Multiplexing
TE	Traffic Engineering
T-MPLS	Transport MPLS
TSS	Transport Service Switch
TST	Time-Sensitive Transport

TTL	Time to Live
UNI	User Network Interface
VB	Virtual Bridge
VID	VLAN ID
VLAN	Virtual LAN
VOA	Variable Optical Attenuator
VoIP	Voice over IP
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VPWS	Virtual Private Wire Service
WAN	Wide Area Network
WDM	Wavelength-Division Multiplexing
WTR	Wait to Restore
WTSA-12	World Telecommunication Standardisation Assembly, 20–29 November 2012