25-05-2012

Deliverable DJ1.1.2: Transport Network Technologies – Study and Testing

Deliverable DJ1.1.2

Contractual Date:	31-03-2012
Actual Date:	25-05-2012
Grant Agreement No.:	238875
Activity:	JRA1
Task Item:	T1
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	NORDUnet
Document Code:	GN3-12-100
Authors:	K. Bozorgebrahimi (NORDUnet/UNINETT), M. Channegowda (Essex University), A. Colmenero (NORDUnet), A. Manolova Fagertun (NORDUnet/DTU), M. Garstka (PSNC), R. Lund (NORDUnet), V. Olifer (JANET), J. Radil (CESNET)

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 Carrier Class Transport Network Technologies (CCTNTs) documented in DJ1.1.1, this report describes the extensive testing performed by JRA1 Task 1. The tests covered EoMPLS, Ethernet OAM, Synchronous Ethernet, PBB-TE, MPLS-TP, OTN and GMPLS, and the CGE technology proxy developed for AutoBAHN.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

Following on from the theoretical research into Carrier Class Transport Network Technologies (CCTNTs) carried out by GN3 Joint Research Activity 1 Future Network, Task 1 Carrier Class Transport Network Technologies (JRA1 Task 1) during the first phase of the GN3 project, documented in “Deliverable DJ1.1.1: Transport Network Technologies Study”, this report describes the extensive testing performed by the Task, and presents the test results. The main objective of the tests was to verify whether the technologies, as implemented in the equipment used, meet the criteria of CCTNT; evaluate the current status of standardisation and implementation; 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; 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, CCTNT-relevant tools, technologies and transport network architectures discussed in this report are Ethernet over Multi-Protocol Label Switching (EoMPLS); Ethernet Operation, Administration and Maintenance (OAM); Synchronous Ethernet (SE); Provider Backbone Bridge Traffic Engineering (PBB-TE); Multi-Protocol Label Switching – Transport Profile (MPLS-TP); and Optical Transport Network (OTN) and Generalised Multi-Protocol Label Switching (GMPLS). The report also covers the cross-Activity tests conducted on the prototype Carrier Grade Ethernet (CGE) technology proxy (TP) designed and implemented for AutoBAHN.

EoMPLS

For the purpose of this document, the term EoMPLS means Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS).The main goal of the tests was to verify whether the EoMPLS technology, as implemented on Juniper Networks routers, fulfils three key CCTNT requirements:

- Service management – Ethernet OAM.
- Reliability – protection and restoration.
- Scalability – multi-domain implementation.

Other tested features were multicast in VPLS and EoMPLS over MPLS Transport Profile (MPLS-TP). Another goal of the tests was to learn the current status of implementation of the EoMPLS technology on routers that are now available to research networks.

The results show that the OAM mechanisms defined for Ethernet (continuity check protocol, link trace and MAC ping) work properly for the EoMPLS services in both customer and provider domains, and can be used to monitor the major service parameters, including availability and packet delay. The major limitation for the mechanisms is that the backbone routers of the provider network lack Layer 2 interfaces and therefore are not visible for the Ethernet-based mechanisms. In case of an outage the Ethernet OAM can correctly indicate the domain in which the outage is located but is not able to find the precise location of the outage in the provider network; other mechanisms are therefore required. The equipment tested (i.e. the hardware implementations as opposed to the technology itself) does not yet support all the OAM functionality defined in the relevant standard (Y.1731).

The protection and restoration mechanisms available for EoMPLS services are offered by the underlying MPLS technology, which supports very robust mechanisms for service recovery in case of an outage in the provider network. The test results verified that the mechanisms ensure service recovery in sub-50 ms times.

The multi-domain test results verified the scalability offered by the EoMPLS technology, mainly in terms of connecting different administrative domains. The tests show that service restoration between multiple domains works properly, offering sub-50 ms restoration times, and that Ethernet OAM mechanisms can be successfully used to monitor the state of an EoMPLS service in a multi-domain environment.

Other results showed that multicast services can be successfully provided in an EoMPLS service instance and that EoMPLS services can be implemented in an MPLS-TP network.

The test results prove that EoMPLS offers manageable and scalable transmission services with robust protection and restoration mechanisms. The technology can be used in research networks as well as carriers' networks to support data transmission services for the research community and other users.

Ethernet OAM

Ethernet was originally created purely as a LAN technology, without any of the OAM features typical of carrier-grade transport technologies. However, considerable efforts by standards bodies and vendors have been made recently to overcome this drawback and to convert the technology into Carrier Ethernet, making it more suitable for wide-area deployment of Ethernet connections. The main objective of the test was to gain experience in monitoring and troubleshooting wide-area point-to-point Ethernet connections on an end-to-end and per-segment basis using the Ethernet OAM functions available in the latest multi-vendor equipment (from Ciena, Extreme Networks, Cisco, Brocade, Overture, Accedian Networks and Cyan, Inc.).

The results of the tests dedicated to monitoring state, rate, delay, jitter and loss of Ethernet services between participants' testbeds on an end-to-end basis mainly proved to be stable and satisfactory. Two problems were discovered with regard to delay measurement of services involving one of the testbeds; these were investigated in subsequent tests.

The results of the investigation into CyPortal, a cloud-based web portal that stores and visualises measurement data, confirmed its usefulness for providing managed wide-area Ethernet servicers, while the results of the investigation into the feasibility of establishing two-level hierarchical sessions to monitor an Ethernet service from both customer and provider perspectives showed that the Ethernet OAM features of the latest network equipment can provide seamless monitoring in a multi-vendor environment.

Synchronous Ethernet

Synchronous Ethernet (SE) is a new standard to provide synchronisation between Ethernet interfaces. It was tested on Cisco Systems equipment. Additional testing was performed with Precision Time Protocol (PTP). The main goal of the tests was to gain practical experience with the new technology, and to assess the current status of implementation of SE (as represented by Cisco equipment) on different series of routers.

The results of the first network clocking tests showed that PTP and SE worked as expected, mainly because of the rather simple setup. More configuration and debugging challenges can be expected in more complex environments.

The results of the second test to provide synchronisation using SE, Ethernet Synchronisation Messaging Channel (ESMC) and PTP in a more complex environment were positive. Maximum Time Interval Error (MTIE) measurements were performed for both SE and PTP, but problems were encountered and the results were not predicative.

It is likely that the R&E community will prefer PTP over SE, as PTP can provide both frequency and time distribution and no additional hardware is required, while SE can distribute accurate frequency only and requires new hardware. Future requests from the R&E community are needed to confirm this expectation.

PBB-TE

The main objectives of the PBB-TE trial were to gain practical experience with and investigate the manageability of PBB-TE in a multi-domain, multi-technology environment, and to evaluate the state of standardisation of PBB-TE, in order to make that knowledge and information available to the GÉANT community. A series of tests were conducted over a multi-domain, multi-technology testbed consisting of the trial participants' local testbeds plus a core testbed. The testbed included equipment from Ciena, Extreme Networks, Juniper, Foundry Networks / Brocade and Dell.

The PBB-TE core tests investigated network-to-network connectivity, resilience, OAM and Ethernet Service Manager functionality. The test results were positive, i.e. the expected result – and indicator of a successful outcome – was achieved. Two problems (relating to frame dropping and traffic selection) were encountered and successfully resolved during the network-to-network connectivity test.

The local Essex University tests investigated PBB-TE connectivity, resilience, and Connectivity Fault Management (CFM) and PBB-TE protection and resilience. The results were positive; no problems were encountered.

The local JANET Lumen House tests investigated unprotected tunnels, CFM, performance monitoring, path protection and traffic policing. The results were positive. Two constraints were encountered in the unprotected tunnels test (regarding the type of port on which PBT tunnels can be created, and changing a B-VID value), and one problem was encountered in the CFM test (CFM for virtual switches stops working after rebooting); a fix was found and the problem has been reported to the vendor.

Generally, the tests conducted showed that the PBB-TE functionality of the equipment used corresponded to expectations, and demonstrated proper behaviour in the key areas of manual establishment of TE point-to-point tunnels; separation of customer and provider address spaces; Continuity Check Message (CCM) monitoring of state of tunnels; fast protection switching; and per-VLAN traffic policing.

The participants concluded that PBB-TE as a transport technology is better suited to single- than multi-domain applications; that PBB-TE deployed in one domain can smoothly interoperate with EoMPLS deployed in other domains; and that EoMPLS should be used in the core networks while PBB-TE could be used in access networks and large campus networks.

MPLS-TP

MPLS-TP is seen by some as a new technology and by others as simply the old, MPLS technology with additional features. It has proved divisive in other ways, too, leading to two different tracks of MPLS-TP OAM standards: IETF and ITU-T. The evolution and standardisation process is still ongoing. The objectives of the trial were to test MPLS-TP and its features; identify the current status of the implementations, particularly with regard to OAM and control plane solutions; and identify areas for further study. The aspects tested were MPLS-TP architecture, services, OAM, protection and control plane. Testing took place at the lab premises of Alcatel-Lucent (ALU), Ciena and Nokia Siemens Networks (NSN).

The MPLS-TP architecture test successfully demonstrated the ability of the equipment to perform basic MPLS-TP functions, e.g. Label Switch Router and Label Edge Router functionality, and tunnel/Label Switched Path (LSP) and Pseudowire (PW) configuration, according to the IETF and ITU-T standard definitions. However, some functionality (for example, ring protection) is still missing due to the early stage of the implementations.

The objective of the services testing was to verify support of the three service types defined by the Metro Ethernet Forum: E-Line, E-LAN and E-Tree. Port-based services of the three service types were successfully configured and verified. To ensure that all service-type combinations are tested it is recommended that testing is continued in GN3 Y4 to address VLAN-based services and to test the different protection schemes for every possible service.

The objective of the OAM tests was to verify the operation of both IETF and ITU-T versions of MPLS-TP OAM tools. The functionality of the IETF tools ping and traceroute was successfully demonstrated in both LSPs and PWs and their importance for on-demand monitoring in production networks shown. A lab demonstration verified the functionality of the ITU-T tools Connectivity Verification (CV), Loopback Measurement and dual-ended Delay Measurement, both proactive and on demand.

The objective of the protection tests was to verify 1+1 protection using Bi-directional Forward Detection (BFD), the IETF mechanism, and CV and Degraded Signal Defect (dDEG), which are ITU-T mechanisms. The results confirmed that the equipment supports 1+1 protection using all three mechanisms: ALU supports the ITU-T-based tools while Ciena and NSN support the IEFT-based tools.

The equipment’s control plane capabilities – namely, topology discovery, topology updates, and LSP/PW creation and deletion – were demonstrated successfully, showing that MPLS-TP is able to operate with and without a Network Management System (NMS) and that the use of the control plane is optional, complying with one of the main requirements of the MPLS-TP framework. The demonstration also proved that the implementation of key protocols (Open Shortest Path First, Resource Reservation Protocol – Traffic Engineering and Label Distribution Protocol) is mature enough for production environments.

The working sessions with the different vendors and the different demonstrations showed that MPLS-TP is a major focus in the industry. It is seen by some vendors as the preferred technology to interface between OTN and the upper layers to deliver packet-based services providing transport features, and many vendors are integrating it into their transport equipment, albeit at different rates, with different priorities and with different (IETF or ITU-T) OAM solutions.

Areas identified for future work include MPLS/MPLS-TP interoperability; continued monitoring and investigation of new MPLS-TP developments; further services testing; and a more detailed investigation of control plane capabilities.

OTN and GMPLS

The transport technology OTN was developed around ten years ago by the ITU-T and has been reviewed several times during recent years to adapt to new market needs. In particular, equipment vendors have been developing new hardware known as OTN switches. The main goals of this set of tests were to establish proof of concept, to learn as much as possible about the technology so that the information can then be made available to the GÉANT community, to identify how NRENs can use it to advantage in their transport networks, and to trigger new and more specific discussions about OTN in the NREN community. The following areas were covered during the testing: Optical Data Unit (ODU) switching; survivability; OAM; and control plane. OTN technology was tested in collaboration with ADVA Optical Networking, Ciena and Nokia Siemens Networks (NSN).

The results of the ODU switching tests confirmed that the equipment was able to map the high-speed (10 GE and 40 GE) client signals tested according to standard (G.709, G.Sup43) specifications; options defined in the standard but not available at the time of the test are in the roadmap and will be available in the near future. The equipment was also able to map 1 GE signals and signals with a bandwidth between 1 GE and 10 GE by using ODU0 and ODUflex, and supports ODU switching functionality and multi-stage ODUk mapping, which offer NRENs a more flexible and integrated way to perform circuit switching (connection-oriented) in a large-scale, cross-domain, IP-over-DWDM network environment, making it possible to deliver dedicated capacity with a high level of Quality of Service to the user. Taking into account that the minimum granularity is ODU0 and the possibility for ODUflex, the multiplexing possibilities of OTN give a lot of flexibility when providing connections in the network.

The survivability tests confirmed that the equipment supports all the standard-defined (G.873.1) protection architectures tested. Hybrid restoration, allowing 50 ms switching, could be of interest to NRENs carrying important traffic or networks with a complicated physical fibre infrastructure, because it allows extra protection compared to traditional Sub-Network Connection Protection (SNCP); however, it requires up-front bandwidth reservation. In control plane restoration without mesh restoration, the new route can be chosen based on either the administrative cost or the total latency of the different links. Although the only way to observe dynamic protocol interaction in case of a link failure was via the management system logs, in the scenario tested the GMPLS procedure in the restoration event was clear.

In the Tandem Connection Monitoring (TCM) testing, the equipment behaved as expected in its defect detection and alarm handling. While the higher monitoring levels and more flexible surveillance possibilities offered by TCM make it potentially very beneficial in the NRENs' multi-domain, multi-vendor environment, its deployment needs thorough planning and close collaboration between network operators. ITU-T G.709

introduces two additional potentially useful functions: the Fault Type, Fault Location (FTFL) channel, which helps to pinpoint the exact fault location when used in conjunction with the TCM functionalities, and delay measurement of ODUk path (DMp), which can be used to select the appropriate route in case of service restoration.

The addition of a control plane to the OTN architecture provides new functionality compared to the initial OTN implementation – topology discovery and update, automated provisioning and decommissioning, and automated restoration – that covers the requirements for carrier-grade transport networks. The control plane/GMPLS tests confirmed that the equipment supports all three areas of functionality, with the observed protocol interaction largely following the standard; three instances of unexpected behaviour require further investigation.

OTN offers networks more powerful switching, mapping and survivability functionality in the digital domain compared to legacy networks like Synchronous Digital Hierarchy (SDH). In addition, it brings seamless integration to the optical domain and provides a common vehicle for mapping, switching and transporting all types of client signals. The testing of existing OTN platforms shows that the products are reaching market maturity. The most important functionalities, such as switching on different ODU levels and survivability based on different Sub-Network Connection (SNC) parameters, including TCM, are already available. ODUflex is not fully implemented yet, but the tests confirmed that the basic functionalities to make ODUflex possible are working well. The integration of a control plane into the OTN technology adds important functionality and intelligence and opens up possibilities for dynamic provisioning tools integration, which is a major requirement for transport technologies in NRENs' transport networks.

Cross-Activity Work

JRA1 Task 1 participants conducted tests on the prototype Carrier Grade Ethernet (CGE) technology proxy (TP) designed and implemented for AutoBAHN (JRA2 Multi-Domain Network Service Research, Task 2 Hybrid Network Provisioning). The TP is a module that is able to communicate with the AutoBAHN software and reserve resources for the Bandwidth on Demand (BoD) service. It was developed to support CGE technology and was based on the implementations of the relevant standards and technologies by Extreme Networks, in particular on the BlackDiamond® 12804 switches, running ExtremeXOS Network Operating System version 12. The work was supported by Essex University, and used the Essex University testbed.

The TP was successfully able to configure the testbed switches, set up a PBB-TE tunnel and enable Layer 2 connectivity between the desired end points. The compatibility of CGE technology with the BoD service and with the AutoBAHN tool in particular was therefore verified. However, extensive configuration needs to be performed on each underlying device, as there is currently no NMS that could abstract these operations. In addition, not all of the BoD parameters were available on the selected equipment, notably bandwidth limits for the created paths. As a result, support for BoD in a CGE domain has to be combined with over-provisioning of the service in terms of capacity. Some test stages, including circuit creation between testbed edge clients, are still ongoing; further results will be available in due course.

Conclusions

The tests carried out proved that these technologies and their capabilities are able to meet the NRENs' requirements for Carrier Class Network Technologies capable of delivering reliable and flexible transport services. It is up to the individual NRENs to decide – with this report and DJ1.1.1 to help them – whether and how to integrate these technologies in their respective infrastructures and how to operate them. There seems to be a general tendency in the NREN environment towards EoMPLS which, with the new added features such as OAM and VPLS multicast, is a perfect match for delivering carrier-class services. However, for some applications, OTN might be the right solution for achieving seamless integration between the optical and digital layer. Since next-generation core networks are going to be built with links of 100 Gbit/s and beyond, there is an obvious need for grooming and multiplexing of 1 GE and 10 GE links. The integration of features like Ethernet OAM and OTN TCM would require carefully thought out planning and design, and close collaboration between NRENs to achieve the smooth operation of their networks. OPEX reduction and more stable and reliable services are the main benefits claimed by the equipment providers, although this needs to be validated by real experiences in real networks.

PBB-TE is a good choice for single domains and campus networks, and several NRENs in Europe have PBB-TE implemented in their networks. The future of MPLS-TP is still uncertain and it will take some time before the penetration of this technology in the market can be evaluated. MPLS-TP is a good choice for those organisations that have a strong transport tradition and culture in their operations. Meanwhile GMPLS will benefit NRENs by adding intelligence to their networks and allowing survivability functions (automated restoration) as well as integration with BoD applications.

There is still work to be done with regard to the integration of these technologies in multi-domain environments, including defining common operational procedures and best practices. In some cases the technology is still in development phase, which means that investigation needs to be continued in Y4 and future GÉANT projects. Areas of focus identified for Y4 include further testing of MPLS-TP; further testing of Ethernet OAM and Service Assurance testing; cross-Activity work with JRA2 Task 3 for perfSONAR extensions to support Ethernet OAM; time-sensitive data applications study; and OpenFlow.

The Task will continue to disseminate the results of its work via the publication of white papers and conference participation, and will conclude at the end of GN3 Y4.

1 Introduction

1.1 Carrier Class Transport Network Technologies

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. Carrier Class Transport Network Technologies (CCTNTs) are defined as technologies designed to provide transport for network services and protocols. "Carrier Class" denotes that the technologies are extremely reliable, support a wide range of speeds up to the current industry maximum, and are well-tested and proven in their capabilities.

A transport network technology must meet the following criteria to qualify as a CCTNT:

- Effectively support diverse types of traffic such as the elastic traffic of data applications and time-sensitive multimedia traffic.
- Effectively support all popular customer services such as Internet access, Virtual Private Networks (VPNs), Voice over IP (VoIP), IPTV and others.
- Be manageable by providing diverse and feature-rich Operation, Administration and Management (OAM) functionality.
- Be reliable by providing resilience and fast restoration for transport connections.
- Be scalable to support numerous customer connections through a carrier network.
- Be able to provide Quality of Service (QoS) and bandwidth guarantees when necessary.
- Provide separation of customer and provider networks in terms of operation and configuration parameters such as address spaces, connection IDs and others.
- Be cost-effective. This is a major requirement for service providers. Network costs are extremely high and a good argument for service providers to change their legacy technologies is the ability to provide better and more flexible services at the same or, if possible, lower cost.
- Deliver high bandwidth and performance up to the current industry limit (i.e. up to 100 G today and 1 T in the nearest future).
- Conform to the appropriate standards.
- Be multi-protocol. A CCTNT should be capable of transporting any kind of customer traffic and so should support different if not all existing protocols.

The CCTNTs, CCTNT-relevant tools, technologies and transport network architectures discussed in this report are as follows:

- Ethernet over Multi-Protocol Label Switching (EoMPLS).
- Ethernet Operation, Administration and Maintenance (OAM).
- Synchronous Ethernet (SE).
- Provider Backbone Bridge Traffic Engineering (PBB-TE) – the IEEE 802.1Qay standard.
- Multi-Protocol Label Switching – Transport Profile (MPLS-TP).
- Optical Transport Network (OTN) and Generalised Multi-Protocol Label Switching (GMPLS).

Further information about CCTNTs is provided in “Deliverable DJ1.1.1: Transport Network Technologies Study” [DJ1.1.1].

1.2 Purpose and Audience

Following on from the theoretical research into CCTNTs carried out by JRA1 Task 1 during the first phase of the GN3 project and documented in [DJ1.1.1], this report describes the extensive testing performed by the Task, and presents the test results. The objective of the tests was to:

- Verify whether the technologies, as implemented in the multi-vendor equipment used in the tests, meet the criteria of CCTNT.
- Evaluate the current status of implementation of the technologies in the equipment used.
- Evaluate the current status of standardisation of the technologies.
- 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.
- Identify areas for further study.

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 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. 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 Approach

The process for carrying out the work in this phase was as follows. First, it was important to identify the main areas to be tested and how they could be tested – test equipment is not always available at NREN labs and it is very expensive. In order to have access to the newest hardware and its newest features, the Task started a round of meetings with vendors to secure their involvement. Vendors were approached locally by each NREN, which proved to be an effective solution as the Task secured the interest and collaboration of several different vendors. Vendors were always properly informed about the GN3 project. Initially, it was the intention to borrow hardware from the vendors to be able to test it. However, this proved to be very complicated for various reasons. A second option was to visit the vendors' lab premises to perform demonstrations and proof of concept testing. In some cases it was possible to obtain hardware and software from the vendor, such as for the Ethernet OAM testing (described in Chapter 3 on page 46).

The work was divided so that each NREN was responsible for a specific area. In some cases several NRENs were able to contribute to the same test. (A list of trial participants is provided in Appendix A.)

1.4 Assumptions

The following assumptions have been made during the testing and evaluation phase:

- The Task undertaking the work is part of a Research Activity, which means that it should investigate new technologies, or new developments in well-established technologies, to help identify and realise their potential benefits for the GÉANT NREN community and deliver innovation to the GÉANT and NREN infrastructure.
- The Task should support the GN3 Service Activities (SAs), for example, SA1 Network Build and Operations. The SAs should be able to use the results of JRA1 Task 1. JRA2 Multi-Domain Network Service Research will also focus on which type of transport technology should be used and what the de-facto control plane will be.
- The term “Carrier Class Transport Network Technologies” is defined by the JRA1 Task 1 team as outlined in Section 1.1 above.

1.5 Constraints

In carrying out the work described in this report, the Task encountered the following constraints:

- In order to have access to the newest hardware and its newest features, many of the tests were performed at vendor premises, sometimes in the form of demonstrations. Nevertheless, the Task considers the results to be very relevant and useful, as it is hoped that this report will demonstrate.
- Some of the technologies tested are still in the process of being developed and standardised by the standardisation bodies. For this reason the report only provides a snapshot of the situation at the time of testing, with indications of where the technologies are heading. Some of the information given in the report might change in the near future or new information become available that could have been of major relevance to the testing. The Task will continue to monitor developments during Year 4 of the GN3 project, and make new or updated information available as necessary.

1.6 In this Document

The rest of this document is structured as follows:

- Chapters 2 to 7 discuss the following CCTNTs, CCTNT-relevant technologies and transport network architectures:
    - Chapter 2: EoMPLS.
    - Chapter 3: Ethernet OAM.
    - Chapter 4: Synchronous Ethernet.
    - Chapter 5: PBB-TE.
    - Chapter 6: MPLS-TP.
    - Chapter 7: OTN and GMPLS.

  Given the length of some of the chapters, each starts with an overview, summarising its contents, which can be read instead of the complete chapter. (Because of this, there is a degree of repetition between the overview and the rest of the chapter.) Each chapter ends with a set of overall conclusions. In general, therefore, 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 Infrastructure, Setup and Configuration.
    - Test Description.
    - Expected Results.
    - Test Results.
    - Test Conclusions.
    - Conclusions.

- Chapter 8 describes the cross-Activity tests conducted by the Task on the prototype Carrier Grade Ethernet technology proxy designed and implemented for AutoBAHN (JRA2 Multi-Domain Network Service Research, Task 2 Hybrid Network Provisioning). It follows the same structure given above.
- Chapter 9 *Conclusions* offers an assessment of the tests and recommendations, and outlines future work.
- Appendix A gives the name and NREN/institution of the participants in each of the trials, together with the trial leader.
- Appendix B presents the technology proxy framework configuration referred to in Chapter 8 *Cross-Activity Work*.

1.7 Acknowledgements

The authors would like to thank all the suppliers who provided experts, equipment and test facilities. It would not have been possible to achieve the results presented in this document without their help and assistance during the test period. The suppliers involved were:

- Accedian Networks.
- ADVA Optical Networking.
- Alcatel-Lucent (ALU).
- Ciena.
- Cisco.
- Cyan, Inc.
- Extreme Networks.
- Juniper Networks.
- Nokia Siemens Networks (NSN).
- Overture.

1.8 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 Ethernet over MPLS

2.1 Overview

This chapter presents the results of testing selected features of the Ethernet over MPLS implementation on Juniper Networks routers. For the purpose of this chapter the term “Ethernet over MPLS” (EoMPLS) means Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS).

As mentioned in the *Introduction*, the GN3 project has identified several requirements for a Carrier Class Transport Network Technology (CCTNT) as well as for a few technologies that may be considered carrier class. The results of this work were published in “Deliverable DJ1.1.1: Transport Network Technologies Study” [DJ1.1.1]. The main goal of the EoMPLS tests was to verify whether the Ethernet over MPLS technology fulfils the following three CCTNT requirements:

- Service Management – Ethernet Operation, Administration and Maintenance (OAM) functionality in VPWS and VPLS.
- Reliability – protection and restoration in VPWS and VPLS.
- Scalability – Multi-domain implementation of VPWS and VPLS.

Other tested features of the EoMPLS technology were multicast in VPLS, a new feature of this technology which seems very interesting for research networks with their multicast services, and EoMPLS over MPLS Transport Profile (MPLS-TP), which makes a link with another carrier class technology identified in [DJ1.1.1] – MPLS-TP.

Another goal of the tests was to learn the current status of implementation of the EoMPLS technology on routers that are now available to research networks. The tests were executed on Juniper Networks MX-series routers. Juniper Networks is one of the leading manufacturers of routers for carrier networks. The MX family consists of routers that offer a very wide range of functionality, including several new developments of the EoMPLS technology, and yet offering low per-port prices compared to other series of routers with support for EoMPLS, which makes them an ideal platform for the research networks and an ideal product for the tests.

The first three tests show that the OAM mechanisms defined for Ethernet (continuity check protocol, link trace and MAC ping) work properly for the EoMPLS services in both customer and provider domains, and can be used to monitor the major service parameters, including availability and packet delay (packet delay measurement is currently supported in the customer domain only). The major limitation for the mechanisms is that the backbone routers of the provider network lack Layer 2 interfaces and therefore are not visible for the Ethernet-based mechanisms. In case of an outage the Ethernet OAM can correctly indicate the domain in which the outage is located but is not able to find the precise location of the outage in the provider network. Other mechanisms, mainly IP-based tracing mechanisms, must be used by the operator to pinpoint the cause of the problem. As Test 3 shows, the Y.1731 Ethernet Alarm Indication Signal (ETH-AIS) mechanism is not yet supported by Juniper Networks routers, so it cannot be used between provider and customer equipment to indicate an outage in the provider network.

The next two tests verify the protection and restoration mechanisms available for EoMPLS services. The mechanisms are offered by the underlying MPLS technology, which supports very robust mechanisms for service recovery in case of an outage in the provider network. The mechanisms ensure service recovery in sub-50 ms times, which was the standard in the legacy Time Division Multiplexing (TDM) technologies (SDH, SONET) and is still required by some applications.

The multi-domain tests verify the scalability offered by the EoMPLS technology, mainly in terms of connecting different administrative domains, which is very important in the research networking environment, where GÉANT and NRENs work together to offer the user an end-to-end service. The tests show that service restoration between multiple domains works properly, offering sub-50 ms restoration times, and that Ethernet OAM mechanisms can be successfully used to monitor the state of an EoMPLS service in a multi-domain environment. The Ethernet OAM mechanisms are limited to the customer and provider domains because there is no Layer 2 interface at the domain boundaries that can serve as a measurement point for OAM mechanisms. The MPLS Trace mechanism or IP-based mechanisms can be used in the operator domain to monitor the state of the domain instead of Ethernet OAM mechanisms.

Test 8 shows that multicast services can be successfully provided in an EoMPLS service instance and Test 9 shows that the EoMPLS services can be implemented in an MPLS-TP network, which is also within the scope of the GN3 JRA1 Activity.

The test results prove that EoMPLS offers manageable and scalable transmission services with robust protection and restoration mechanisms. The technology can be used in research networks as well as carriers' networks to support data transmission services for the research community and other users.

2.2 Introduction

This chapter presents the results of testing selected features of the Ethernet over MPLS implementation on Juniper Networks routers. For the purpose of this chapter the term "Ethernet over MPLS" (EoMPLS) means Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS).

The GN3 project has identified several requirements for a Carrier Class Transport Network Technology (CCTNT) as well as for a few technologies that may be considered carrier class. The results of this work were published in "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1]. The main requirements for a Carrier Class Transport Network Technology include scalability, reliability and service management. Most of the EoMPLS tests were designed for those three aspects of a carrier class technology, the tests' main goal being to verify whether the Ethernet over MPLS technology fulfils the following three requirements:

- Service Management – Ethernet Operation, Administration and Maintenance (OAM) functionality in VPWS and VPLS.
- Reliability – protection and restoration in VPWS and VPLS.
- Scalability – Multi-domain implementation of VPWS and VPLS.

Other tested features of the EoMPLS technology were multicast in VPLS, a new feature of this technology which seems very interesting for research networks with their multicast services, and EoMPLS over MPLS Transport Profile (MPLS-TP), which makes a link with another carrier class technology identified in [DJ1.1.1] – MPLS-TP.

Another goal of the tests was to learn the current status of implementation of the Ethernet over MPLS technology on routers that are now available to research networks. The tests were executed on Juniper Networks MX-series routers. Juniper Networks is one of the leading manufacturers of routers for carrier networks. The MX family consists of routers that offer a very wide range of functionality, including very robust implementation of EoMPLS, and yet are comparatively cheap, which makes them a perfect platform for the research networks and a perfect product for the tests.

The chapter describes nine tests executed on Juniper Networks routers which aimed to verify the features listed above:

- Tests 1, 2 and 3 verify the service manageability in terms of OAM functionality.
- Tests 4 and 5 verify the protection and restoration mechanisms available for EoMPLS services.
- Test 6 deals with both manageability and scalability as it verifies OAM mechanisms in a multi-domain network.
- Test 7 combines scalability with protection and restoration as it verifies restoration mechanisms in a multi-domain network.
- Test 8 deals with multicast in a VPLS service.
- Test 9 is designed to verify EoMPLS operations over an MPLS-TP infrastructure.

2.3 Technology Briefing

Ethernet over IP/MPLS (Internet Protocol/Multi-Protocol Label Switching) is one of the technologies that can be used to transport Ethernet frames over a provider’s backbone network. Although MPLS is said to be complex, it is also popular in big providers’ networks and over the years of deployment it has proved its reliability and scalability.

The Metro Ethernet Forum (MEF) has defined three types of services that can be delivered through Metro Ethernet. All three services types can be delivered by different services provisioned on top of the MPLS transport:

- E-Line (point-to-point service type) can be provisioned as Virtual Private Wire Service (VPWS).
- E-LAN (multipoint-to-multipoint service type) can be provisioned as Virtual Private LAN Service (VPLS).

- E-Tree (point-to-multipoint service type) can be provisioned as Virtual Private Multicast Service (VPMS).

The three services listed above are referred to as Layer 2 Virtual Private Networks (L2VPNs).

Figure 2.1 shows the reference model for L2VPNs. The model identifies the customer equipment (CE) which is connected to the provider network, provider edge equipment (PE) to which customer access links are connected, and the MPLS-enabled core of the provider network. The provider network can support multiple instances of L2VPNs which are logically separated from each other – customer traffic will never be forwarded between L2VPNs.

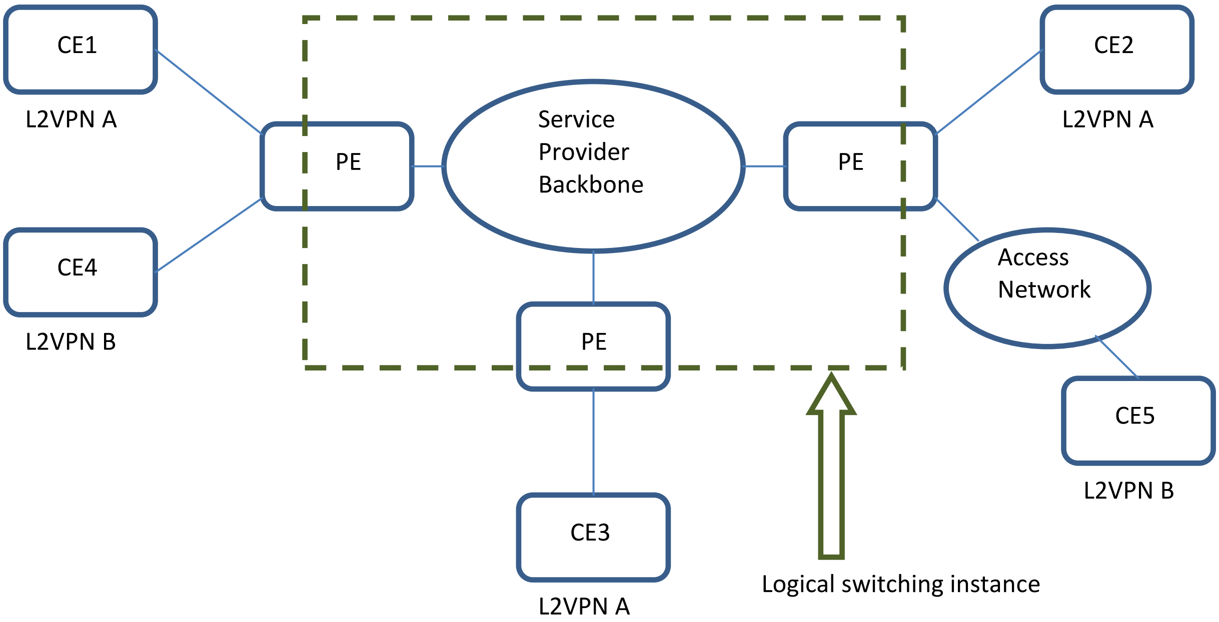


Figure 2.1: Reference model for L2VPNs (source: [RFC 4665])

All the services operate over pseudowires (PWs) defined by the IETF PWE3 working group. A pseudowire means an emulated point-to-point connection over a packet-switched network (PSN) that allows the interconnection of two nodes with any Layer 2 technology. The important feature of the pseudowire technology is that the service-specific functions are located only on the edge of the provider network, while in the core only MPLS label-switched paths (LSPs) are established. This way the core routers do not need any knowledge about the services provided by the network, which increases the scalability of the services and reduces the complexity of the core routers.

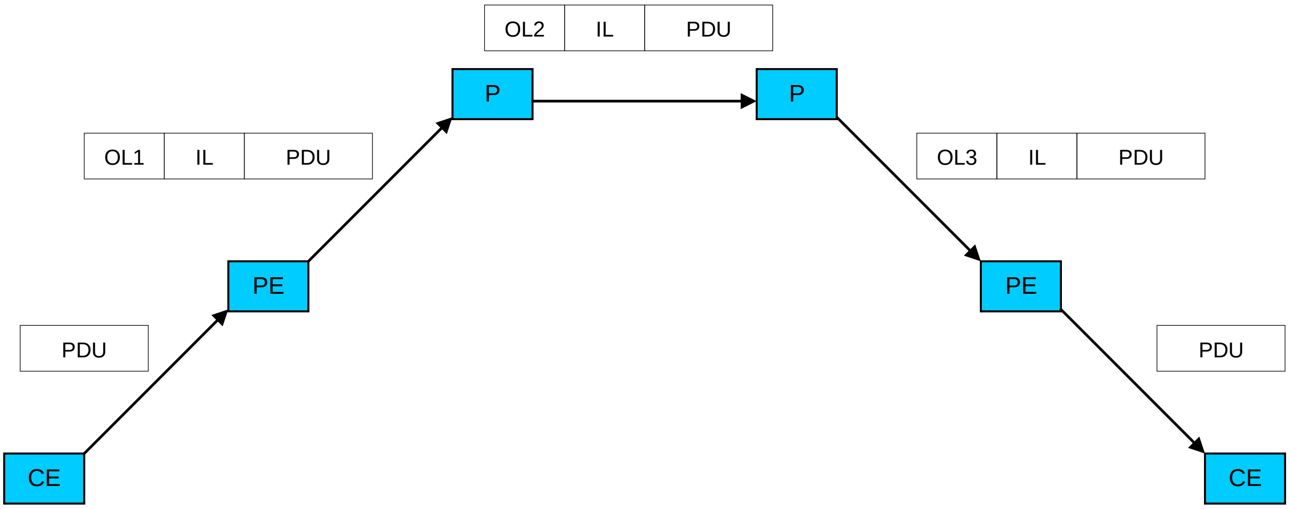


Figure 2.2: MPLS L2VPN forwarding

Figure 2.2 shows how customer data is forwarded in a L2VPN network. First an Ethernet frame (PDU) is received by a Provider Edge (PE) router from customer equipment (or Customer Edge – CE). The ingress PE router adds two MPLS labels to the frame. The outer label (OL) identifies the egress PE router – the destination of the packet in the MPLS network. The outer label is used in the MPLS network to switch the packet. The value of the label can be swapped by every core router (P) that forwards the packet (in the diagram, label OL1 is swapped to label OL2 and OL2 to OL3). The inner label (IL) identifies the VPN service to which the packet belongs. It is not used for forwarding in the MPLS network but it is used by the egress PE to select the proper outgoing link to the customer. The egress PE strips off both labels and sends the original Ethernet frame to the proper customer equipment.

Further details about the Ethernet over MPLS technology and the types of services it offers are provided in [DJ1.1.1].

2.4 Test Objective

The objective of the tests was to assess the status of implementation of selected VPWS and VPLS features on Juniper Networks routers. The selected features are either new or interesting for research and education networks. Those that are assumed to be interesting for NRENs include multi-domain implementations, because NRENs and GÉANT together offer end-to-end services that span several administrative domains, and multicast services, which can be used for broadcasting educational content and the results of scientific expeiments. The protection and restoration tests are of interest for both NRENs and commercial carriers as both must offer reliable services to their users.

The ultimate goal of the tests was to verify whether the Ethernet over MPLS technology fulfils selected requirements for Carrier Class Transport Network Technologies as identified in [DJ1.1.1].

2.5 Test Infrastructure

2.5.1 Description

The tests were conducted on Juniper Networks MX-series routers (MX960, MX480, MX240) with JUNOS software version 11.1R2.3.

IXIA (ex-Agilent) N2X router testers were used for traffic generation and analysis, and as measurement end points for the OAM test.

The testbed topology varied between tests and is shown in detail in the test descriptions. The general rule was that the five MX-series routers emulated a carrier's network (PE and P routers) with 10 GE interconnections as well as customer's edge routers (CE routers) in some tests, while the two testers emulated customers (CE routers).

2.5.2 Configurations

The configuration varied between tests. Information about the configuration is given in each test description.

2.6 Test 1: 802.1ag Connectivity Fault Management and MPLS Trace Test

2.6.1 Test Setup

The test should prove proper operation of the Connectivity Fault Management (CFM) in the customer domain (between CE routers) and provider domain (between PE routers). It should also prove that MPLS Trace can be used to find the location of a connectivity loss in an MPLS network.

The IXIA N2X analyser used as one of the CE routers will prove that the CFM implementation on MX-series routers is compliant with standards and can be used in a multi-vendor environment. In case it is not possible to use an analyser for this test, two MX-series routers can be used as CE routers to verify operations of CFM in the customer domain.

The test should be repeated for Label Distribution Protocol (LDP)-signalled VPLS, Border Gateway Protocol (BGP)-signalled VPLS and LDP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.3.

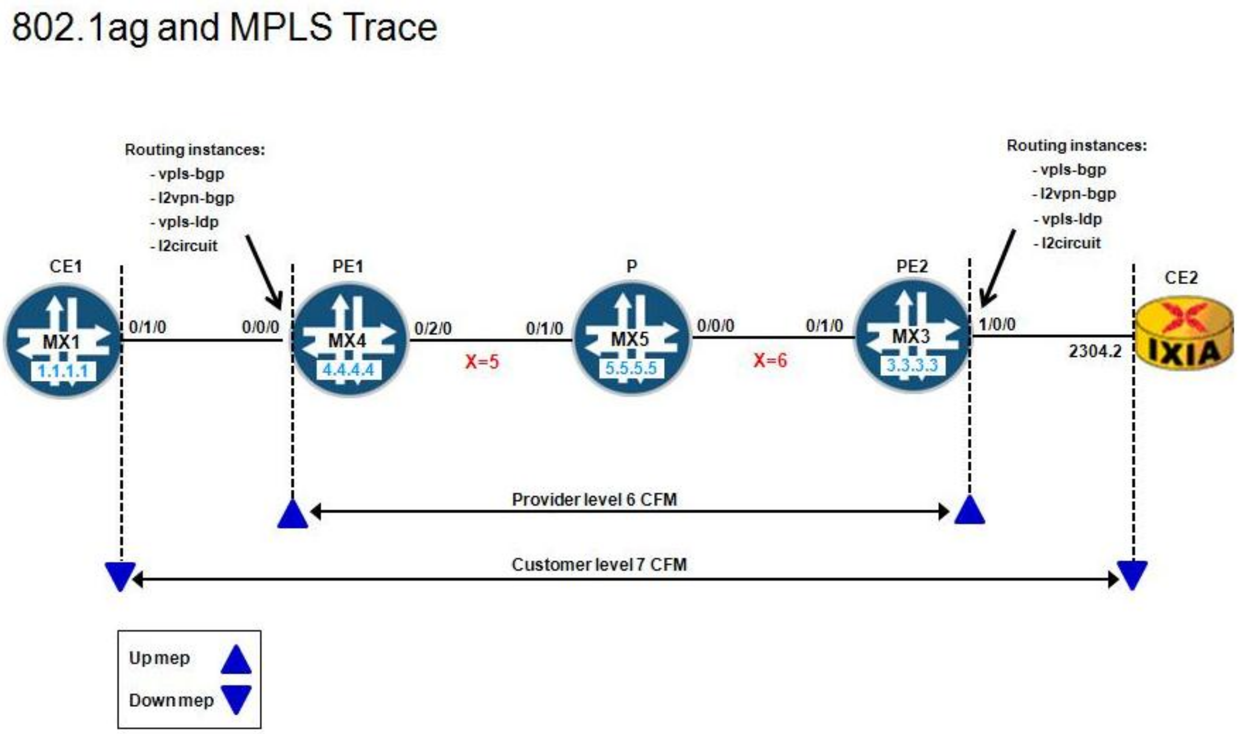


Figure 2.3: Testbed topology for Test 1 and Test 3

2.6.2 Configuration

Measurement end points (MEPs) should be located on CE devices (for the customer domain) and PE devices (for the provider domain). MEPs on PE routers should be associated with the VPLS or VPWS instance used for interconnecting CE devices. Measurement intermediate points (MIPs) for the customer domain should be located on PE routers.

The Level 6 MEPs implicitly create Level 7 MIPs for the customer domain located on customer-facing interfaces of the PEs. The MIPs are not showed in Figure 2.3.

2.6.3 Test Description

The test consisted of the following steps:

1. Verify operation of the continuity check protocol between PE1 and PE2.
2. Verify operation of the continuity check protocol between CE1 and CE2.
3. Verify operation of the linktrace protocol between PE1 and PE2.
4. Verify operation of the linktrace protocol between CE1 and CE2.

5. Change the admin status of one of the PE2-facing interfaces of the P router to DOWN in order to emulate a fibre cut between P and PE2.
6. Check the reaction of the CFM to the loss of connectivity between PEs.
7. Verify whether the linktrace protocol can find the location of the outage.
8. Use MPLS Trace from PE1 to find the location of the outage.
9. Change the admin status of one of the PE2-facing interfaces of the P router to UP.
10. Check the reaction of the CFM to the service restoration.

2.6.4 Expected Results

The test should prove proper operation of the continuity check protocol and the linktrace protocol in the customer and provider domains. It is expected that the two protocols will detect the connectivity failure (step 5) and restoration (step 9). Having MIPs for the customer domains on PEs should allow the location of the fault to be identified as being between the PEs.

The test should also demonstrate the ability of MPLS Trace to find the location of an outage in an MPLS network.

2.6.5 Results

- The continuity check protocol (CCP) ran properly in both provider (PE-PE) and customer (CE-CE) domains. For each MEP a remote MEP was successfully learned and its state was properly indicated as OK.

  Sample output:

  ```
  Remote MEP identifier: 300, State: ok
  MAC address: 00:1d:b5:41:11:4a, Type: Learned
  ```

- When the network was disconnected the CCP properly changed the states of remote MEPs to FAILED, indicating loss of connectivity with the remote MEP.

  Sample output:

  ```
  Remote MEP identifier: 300, State: failed
  MAC address: 00:1d:b5:41:11:4a, Type: Learned
  ```

- When the simulated network outage was fixed CCP changed the state of remote MEPs to OK, indicating restoration of connectivity. Of course, CCP reacted to the restoration of the VPLS or VPWS service, not the restoration of connectivity between PEs, which means some delay from the actual connectivity restoration. This is correct as CCP was used to monitor the services, not the network topology.
- The linktrace protocol was used to trace the connection in the provider domain. As there were no MIPs defined in the provider domain, only one hop (the remote MEP on the remote PE) was discovered.

Sample output:

```
Hop    TTL    Source MAC address        Next-hop MAC address
1      62     00:1d:b5:41:11:4a         00:00:00:00:00:00
```

- The linktrace protocol was used to trace the connection in the customer domain. It successfully reported two MIPs on both PE routers and a remote MEP on the remote CE.

  Sample output:

```
Hop    TTL    Source MAC address        Next-hop MAC address
1      61     00:00:65:01:01:02         00:00:65:01:01:02
2      62     00:1d:b5:41:11:4a         00:1d:b5:41:11:4a
3      63     00:1f:12:b8:78:00         00:00:00:00:00:00
```

- When the network was disconnected between PE1 and PE2 the linktrace protocol initiated from CE1 towards CE2 was not able to find the location of the outage. As explained by Juniper, this is because when the pseudowire goes down (as it does when connectivity between PEs is lost) all MAC addresses learned on the pseudowire are flushed, including the MAC addresses of the remote MIPs and MEPs.

- MPLS Trace was used to trace the connection in the provider domain. MPLS Trace successfully discovered all routers on the route (the P router and the remote PE) as well as showing the MPLS labels used for transmission (the outer label used for switching in the MPLS network).

  Sample output:

```
ttl    Label    Protocol    Address          Previous Hop     Probe Status
1      301632   RSVP-TE     192.168.5.5      (null)           Success
2      3        RSVP-TE     192.168.6.3      192.168.5.5      Egress
```

- When the connection between PE1 and PE2 was disconnected MPLS Trace was not able to find the location of the outage. This is because MPLS Trace traces an LSP in the MPLS network and when the connectivity is lost the whole LSP is removed so it cannot be traced using MPLS-based tools.

- As an addition to the initial test plan, MAC ping was used to verify the connectivity to the remote MEP. This test showed that Ethernet ping can be used for verifying and monitoring the status of a connection between CEs.

  Sample output:

```
jnpr@MX1# run ping ethernet 00:00:65:04:01:02 mep 403 maintenance-
      association Cust_l2circuit
PING to 00:00:65:04:01:02, Interface xe-0/1/0.103
60 bytes from 00:00:65:04:01:02: lbm_seq=0
60 bytes from 00:00:65:04:01:02: lbm_seq=1
60 bytes from 00:00:65:04:01:02: lbm_seq=2
60 bytes from 00:00:65:04:01:02: lbm_seq=3
^C--- ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
```

The tests were repeated for LDP-signalled and BGP-signalled VPLS as well as LDP-signalled and BGP-signalled VPWS, giving the same results in all cases.

2.6.6 Test Conclusions

- The test shows that the continuity check protocol and linktrace protocol work properly with Ethernet over MPLS services and can be used to monitor the state of such services by both customers and providers. MAC ping can also be used for online monitoring of the state of the service.
- Using an IXIA N2X tester as one of the CEs and a Juniper router as the other CE verified the interoperability of the continuity check and linktrace protocols between a Juniper implementation and other vendors (represented by IXIA).
- The linktrace protocol in the customer domain was able successfully to trace the connection, giving the user information about the MIPs located on the PE routers as well as the remote MEP on the remote CE. As it is not possible to have a MIP on a P router in the core of the provider network (for lack of Layer 2 interfaces on P routers), the linktrace protocol is not able to trace the route in the provider's network, making the topology of the network not visible to the customer.
- Using the linktrace protocol in the provider network shows only the remote MEP located on the remote PE, so the information given by linktrace is not very valuable. For lack of Layer 2 interfaces on the P routers in the core of the provider network, MEPs and MIPs cannot be located in the core and, as a result, the core routers (P routers) cannot be traced by the link trace protocol.
- MPLS Trace can be used to trace the route of the service in the provider network. This tool is available to the network administrator only and cannot be used by the customer.
- Neither linktrace nor MPLS Trace can be used to find the location of the outage in the provider network when connectivity between PE routers is broken. The linktrace protocol (initiated from a CE towards the local PE) can be used by the customer to verify the state of the CE-PE connection. Other tools, like IP traceroute available in an IP-MPLS network, can be used to facilitate locating transmission problems. The problem with linktrace and MPLS Trace does not affect the usefulness of the EoMPLS technology as alternative tools (e.g. IP traceroute) are available in packet networks and can be used in addition to linktrace and MPLS Trace in order to achieve all the desired functionality. MPLS networks consist of several layers and each layer offers some management tools which can be used together for network management and operations.

2.7 Test 2: Y.1731 Ethernet Frame Delay Measurement Test

2.7.1 Test Setup

The test should prove the ability of MX-series routers to measure the frame delay and frame delay variation for a VPLS and VPWS service in the customer domain (between CE routers) and provider domain (between PE routers).

The analyser used as one of the CE routers will prove that the implementation of the frame delay variation measurement on MX-series routers is compliant with standards and can be used in a multi-vendor environment. In case it is not possible to use an analyser for this test, two MX-series routers can be used as CE routers to verify operations of frame delay variation in the customer domain.

The test should be repeated for LDP-signalled VPLS, BGP-signalled VPLS and LDP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.4 (using two Juniper MX routers as CEs) and Figure 2.5 (using one Juniper MX router and one IXIA N2X tester as CEs).

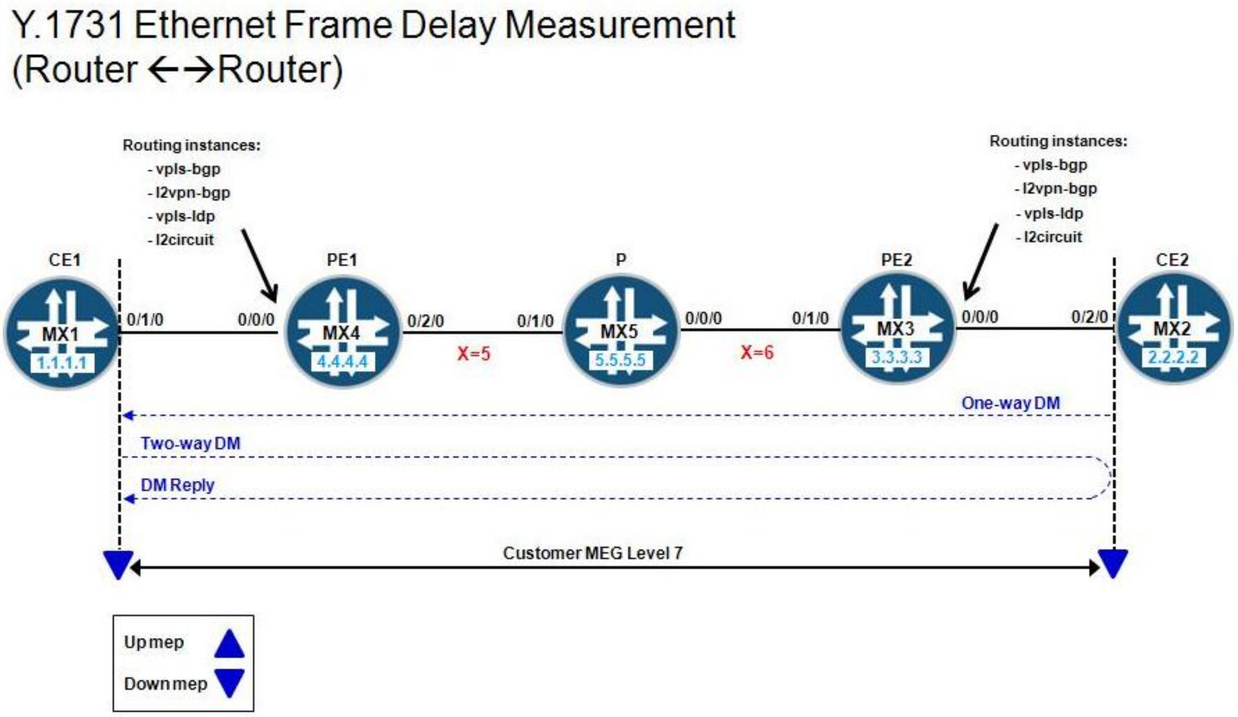


Figure 2.4: Testbed topology for Test 2 (Router↔Router)

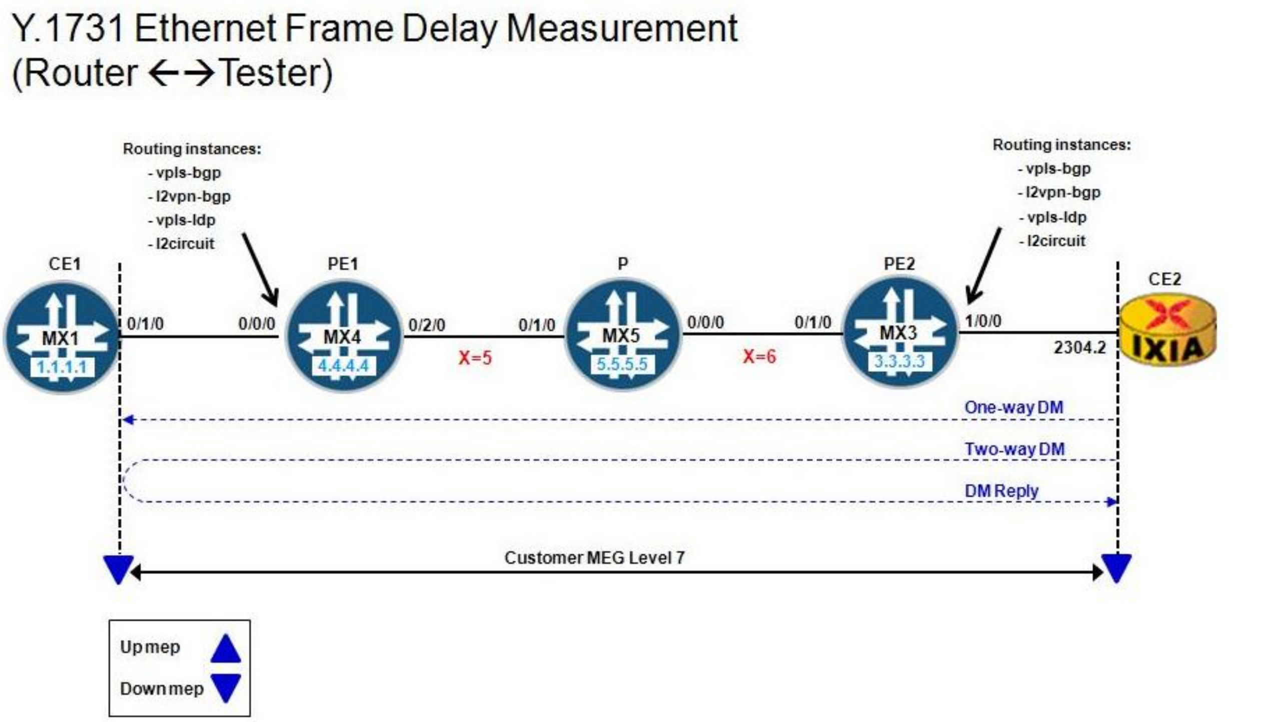


Figure 2.5: Testbed topology for Test 2 (Router↔Tester)

2.7.2 Configuration

Measurement end points (MEPs) should be located on CE devices (for the customer domain) and PE devices (for the provider domain). MEPs on PE routers should be associated with the VPLS or VPWS instance used for interconnecting CE devices.

One-way measurements require that the clocks of the sending and receiving devices are synchronised. If such synchronisation is not possible, only two-way measurements will be tested.

2.7.3 Test Description

The test consisted of the following steps:

1. Verify whether one-way frame delay and frame delay variation is measured between CE1 and CE2.
2. Verify whether two-way frame delay and frame delay variation is measured between CE1 and CE2.
3. Verify whether one-way frame delay and frame delay variation is measured between PE1 and PE2.
4. Verify whether two-way frame delay and frame delay variation is measured between PE1 and PE2.

2.7.4 Expected Results

The test should prove proper operation of the frame delay and frame delay variation measurement in the customer and provider domains.

2.7.5 Results

- One-way and two-way frame delay and frame delay variation were successfully measured in the customer domain (between CE devices). The tests were repeated using two Juniper MX routers as well as a single MX router and an IXIA N2X tester as CE devices.
- Frame delay and frame delay variation in the provider domain (between CE-facing PE interfaces) (i.e. steps 3 and 4 of the test) is not supported by Juniper.

The tests were repeated for LDP-signalled and BGP-signalled VPLS as well as LSP-signalled and BGP-signalled VPWS, giving the same results in all cases.

Sample output for one-way delay measurement on an MX router:

```
jnpr@MX2# run monitor ethernet delay-measurement one-way 00:1f:12:b1:00:29 mep
      100 maintenance-association Cust_l2vpn-bgp maintenance-domain Cust_l2vpn-
      bgp
One-way ETH-DM request to 00:1f:12:b1:00:29, Interface xe-0/2/0.100
1DM Frames sent : 10
--- Delay measurement statistics ---
Packets transmitted: 10
Average delay: NA, Average delay variation: NA
Best case delay: NA, Worst case delay: NA

jnpr@MX1# run show oam ethernet cfm delay-statistics maintenance-association
      Cust_l2vpn-bgp maintenance-domain Cust_l2vpn-bgp
MEP identifier: 100, MAC address: 00:1f:12:b1:00:29
Remote MEP count: 1

  Remote MAC address: 00:1f:12:b7:d9:4a
    Delay measurement statistics:
    Index  One-way delay  Two-way delay
              (usec)         (usec)
      1        259
      2        261
      3        261
      4        264
      5        263
      6        264
      7        263
      8        266
      9        264
      10       265
    Average one-way delay          : 263 usec
    Average one-way delay variation: 1 usec
    Best case one-way delay        : 259 usec
    Worst case one-way delay       : 266 usec
```

Sample output for two-way delay measurement on an MX router:

```
jnpr@MX1# run monitor ethernet delay-measurement two-way 00:1f:12:b7:d9:4a mep
      403 maintenance-association Cust_l2circuit maintenance-domain
      Cust_l2circuit
Two-way ETH-DM request to 00:1f:12:b7:d9:4a, Interface xe-0/1/0.103
DMR received from 00:1f:12:b7:d9:4a Delay: 205 usec Delay variation: 0 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 202 usec Delay variation: 3 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 202 usec Delay variation: 0 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 200 usec Delay variation: 2 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 206 usec Delay variation: 6 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 202 usec Delay variation: 4 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 198 usec Delay variation: 4 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 194 usec Delay variation: 4 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 195 usec Delay variation: 1 usec
DMR received from 00:1f:12:b7:d9:4a Delay: 192 usec Delay variation: 3 usec

--- Delay measurement statistics ---
Packets transmitted: 10, Valid packets received: 10
Average delay: 199 usec, Average delay variation: 2 usec
Best case delay: 192 usec, Worst case delay: 206 usec

jnpr@MX1# run show oam ethernet cfm delay-statistics maintenance-association
      Cust_l2circuit maintenance-domain Cust_l2circuit
MEP identifier: 103, MAC address: 00:1f:12:b1:00:29
Remote MEP count: 1

  Remote MAC address: 00:1f:12:b7:d9:4a
    Delay measurement statistics:
    Index  One-way delay  Two-way delay
              (usec)         (usec)
      1        435            205
      2        440            202
      3        437            202
      4        439            200
      5        441            206
      6         437            202
      7         433            198
      8         440            194
      9         438            195
      10        438            192
    Average one-way delay          : 437 usec
    Average one-way delay variation: 2 usec
    Best case one-way delay        : 433 usec
    Worst case one-way delay       : 441 usec
    Average two-way delay          : 199 usec
    Average two-way delay variation: 2 usec
    Best case two-way delay        : 192 usec
    Worst case two-way delay       : 206 usec
```

2.7.6 Test Conclusions

- One-way and two-way delay and delay variation measurements can be used in the customer domain to monitor the quality of an EoMPLS service. Using one-way measurements requires clock synchronisation between CEs. Such synchronisation requires additional mechanisms on the CEs, like Network Time Protocol (NTP). If such synchronisation is not possible, only two-way measurements are possible.
- Delay and delay variation measurements are not supported by Juniper in the provider domain so they cannot be used by the provider to monitor the quality of an EoMPLS service. (To date the authors have not investigated other products to know whether this is true of all vendors.) However, measurements can be done by the provider network operator using external probes which will be placed in the network as CE devices.

2.8 Test 3: Y.1731 ETH-AIS Test

2.8.1 Test Setup

The test should prove that loss of connectivity between PE routers in a VPWS or VPLS (detected by CFM) will trigger an Ethernet Alarm Indication Signal (ETH-AIS) towards CEs.

The analyser used as one of the CE routers will prove that the ETH-AIS implementation on MX-series routers is compliant with standards and can be used in a multi-vendor environment. In case it is not possible to use an analyser for this test, two MX-series routers can be used as CE routers to verify operations of ETH-AIS between PE and CE.

The test should be repeated for LDP-signalled VPLS, BGP-signalled VPLS and LDP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.3 on page 19.

2.8.2 Configuration

Measurement end points (MEPs) should be located on PEs and associated with the VPLS or VPWS service used for interconnecting CE devices.

2.8.3 Test Description

The test consisted of the following steps:

1. Verify connectivity of the VPWS or VPLS between CE devices (using ping).
2. Check whether one of the CE devices receives ETH-AIS or any other alarm from the corresponding PEs.
3. Change the admin status of one of the PE2-facing interfaces of the P router to DOWN in order to emulate a fibre cut between P and PE2.
4. Verify whether CFM has noticed the loss of connectivity between PEs.
5. Verify whether both CEs receive ETH-AIS from the corresponding PEs.
6. On both CEs, view the Chassis ID TLV and Port ID TLV in the received ETH-AIS frames.
7. Change the admin status of one of the PE2-facing interfaces of the P router to UP to restore connectivity in the VPWS or VPLS.
8. Verify that both CE devices stopped receiving ETH-AIS from the corresponding PEs.
9. Verify connectivity of the VPWS or VPLS between CE devices (using ping).

2.8.4 Expected Results

The test should prove proper generation of ETH-AIS by PE routers towards CE routers in case of connectivity failure between PEs (triggered by CFM).

2.8.5 Results

The test was not performed because the Juniper equipment used for the EoMPLS tests does not yet support Y.1731 ETH-AIS.

2.8.6 Test Conclusions

As Juniper does not yet support Y.1731 ETH-AIS, which is a relatively new development, this mechanism cannot be used on Juniper routers to signal the state of the service towards CE devices. (To date the authors have not investigated other products to know whether this is true of all vendors.) The test is described in full so that NRENs can perform it on other vendors' equipment.

2.9 Test 4: MPLS Fast Reroute in VPWS and VPLS Test

2.9.1 Test Setup

The test should prove that MPLS Fast Reroute can benefit VPWS and VPLS traffic in terms of faster restoration of connectivity in case of a link failure.

The traffic generator and analysers (IXIA testers) employed as CE devices will be used to measure the time needed for restoration of connectivity (time will be calculated on the basis of the number of packets lost during restoration).

The test should be repeated for LDP-signalled VPLS and LDP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.6.

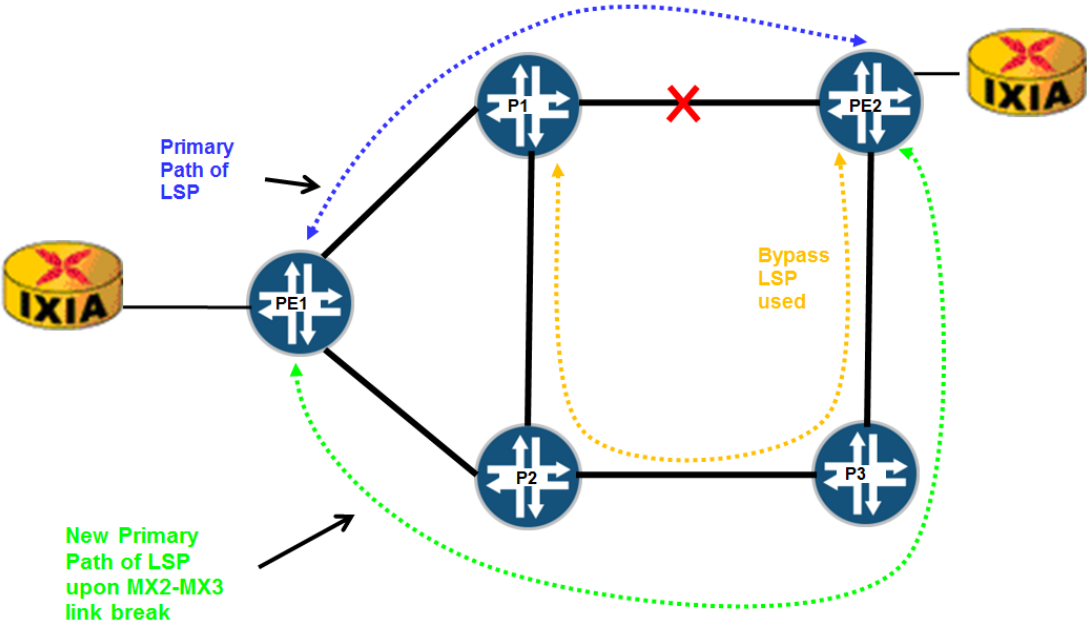


Figure 2.6: Testbed topology for Test 4

2.9.2 Configuration

The LSP between PE1 and PE2 should be configured with the primary path using the PE1-P1-PE2 route and the stand-by secondary path using the PE1-P2-P3-PE2 route. Make sure that the primary path is active.

For steps 1 to 6 Fast Reroute should not be active for the LSP. Fast Reroute (Link Protection) should be activated before step 8.

2.9.3 Test Description

The test consisted of the following steps:

1. Start generating traffic from one of the IXIA testers being used as CE devices and receiving at the other.
2. Disconnect the P1-PE2 link.
3. Watch the LSP being rerouted to the secondary path.
4. Calculate the time needed for service restoration based on the number of lost packets (difference between the number of packets sent by the source IXIA tester and received by the other).
5. Restore the link between P1 and PE2.
6. Make sure the LSP uses the primary path again.
7. Configure Fast Reroute (Link Protection) for the primary path to protect against P1-PE2 fibre cut.
8. Make sure the Bypass LSP for the P1-PE2 link uses the P1-P2-P3-PE2 route.
9. Disconnect the P1-PE2 link.
10. Watch the traffic being rerouted to the Bypass LSP.
11. Watch the LSP being rerouted to the secondary path.
12. Calculate the time needed for service restoration based on the number of lost packets (difference between the number of packets sent by the source IXIA tester and received by the other).

2.9.4 Expected Results

The test should prove that Fast Reroute can reduce the time needed for service restoration for VPWS and VPLS.

2.9.5 Results

The measured recovery times are shown in Table 2.1 below.

Test	Without Fast Reroute		With Fast Reroute	
	Number of Packets Lost	Recovery Time (ms)	Number of Packets Lost	Recovery Time (ms)
BGP-signalled VPWS	19002	190	2143	21
LDP-signalled VPWS	18736	187	2173	22
BGP-signalled VPLS	33408	334	2481	25
LDP-signalled VPLS	34550	346	2166	22

Table 2.1: Test 4 recovery times

The packet transmission rate was 100,000 packets per second (100 packets per millisecond).

2.9.6 Test Conclusions

For all the tested Ethernet over MPLS services Fast Reroute significantly decreases the time needed for service restoration when a transmission link in the provider network is broken and traffic must be rerouted to an alternate path. The restoration times when Fast Reroute is being used are much lower than the 50 ms guaranteed by TDM networks and required by some voice services.

This test proved that Ethernet over MPLS technology has a powerful protection and restoration mechanism that can offer restoration times no worse than traditional TDM networks. This is particularly important for voice and multimedia services which can use EoMPLS as their transport infrastructure. Having restoration times similar to those offered by legacy TDM networks allows services to be migrated from TDM infrastructure to MPLS without affecting Service Level Agreements (SLAs) and service reliability.

The tests were performed for a single LSP but the number of LSPs should not impact the recovery time. As the backup mode facility of Fast Reroute (link-protection keyword in the Juniper configuration) was used, a single bypass LSP can be used to protect multiple LSPs traversing an affected link, so the number of LSPs does not result in a complex configuration and should not impact the recovery times.

2.10 Test 5: BFD and MPLS Fast Reroute Test

2.10.1 Test Setup

The test should prove that Bi-directional Forwarding Detection (BFD) together with MPLS Fast Reroute can benefit VPWS and VPLS services in terms of faster restoration of connectivity in case of a link failure. For this test the failure will be emulated between two Layer 2 switches, so any loss of signal will not occur on routers and loss of connectivity must be detected with the use of other mechanisms (BFD in this test).

The traffic generator and analysers (IXIA testers) employed as CE devices will be used to measure the time needed for restoration of connectivity (time will be calculated based on the number of packets lost during restoration).

The test should be repeated for LDP-signalled VPLS and LDP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.7. S1 and S2 are simple Ethernet switches without IP or MPLS functionality, transparent for routing protocols and MPLS.

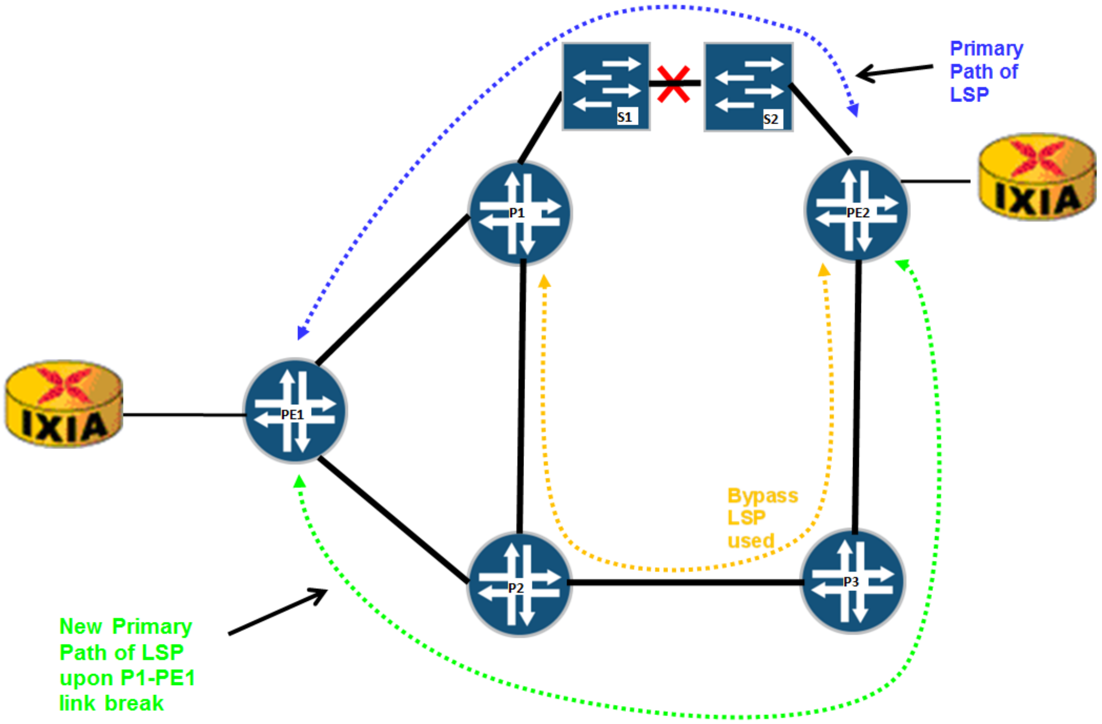


Figure 2.7: Testbed topology for Test 5

2.10.2 Configuration

The LSP between PE1 and PE2 should be configured with the primary path using the PE1-P1-S1-S2-PE2 route and the stand-by secondary path using the PE1-P2-P3-PE2 route. Make sure that the primary path is active.

Fast Reroute (Link Protection) should be enabled for this LSP to protect against connectivity failure between P1 and PE2.

BFD should not be enabled for steps 1 to 6. It should be enabled before step 8.

2.10.3 Test Description

The test consisted of the following steps:

1. Start generating traffic from one of the IXIA testers being used as CE devices and receiving at the other.
2. Emulate a connectivity failure between S1 and S2 (turn off the S2-facing interface of S1).
3. Watch the traffic being rerouted to the Bypass LSP.
4. Calculate the time needed for service restoration based on the number of lost packets (difference between the number of packets sent by the source IXIA tester and received by the other).
5. Restore the link between S1 and S2 (turn on the S2-facing interface of S1).
6. Make sure traffic is rerouted back to the PE1-P1-S1-S2-PE2 link.
7. Enable BFD between P1 and PE2 in order to trigger traffic rerouting to the Bypass LSP in case a connectivity failure between P1 and PE2 is detected by BFD.
8. Emulate a connectivity failure between S1 and S2 (turn off the S2-facing interface of S1).
9. Watch the traffic being rerouted to the Bypass LSP.
10. Calculate the time needed for service restoration based on the number of lost packets (difference between the number of packets sent by the source IXIA tester and received by the other).

2.10.4 Expected Results

The test should prove that BFD can reduce the time needed for service restoration for VPWS and VPLS in case a link failure does not cause loss of signal on routers' interfaces.

2.10.5 Results

The measured recovery times are shown in Table 2.2 below.

Test	Without Fast Reroute		With Fast Reroute	
	Number of Packets Lost	Recovery Time (Seconds)	Number of Packets Lost	Recovery Time (ms)
BGP-signalled VPWS	2066965	21	7194	72
LDP-signalled VPWS	2551521	26	7240	72
BGP-signalled VPLS	2037911	20	7615	76
LDP-signalled VPLS	2490858	25	6971	70

Table 2.2: Test 5 recovery times

The packet transmission rate was 100,000 packets per second (100 packets per millisecond).

2.10.6 Test Conclusions

- For all the tested Ethernet over MPLS services BFD significantly decreases the time needed for service restoration when a loss of signal event cannot trigger traffic rerouting. Such a case is quite common in an Ethernet network in which Layer 2 switches or optical devices are used between MPLS routers and a loss of connectivity inside such a sub-network does not trigger a loss of signal event on MPLS routers' interfaces.
- The problem with triggering traffic rerouting in such a case does not occur in TDM networks because the state of connectivity is signalled inside the containers (using AIS, etc.) and can trigger traffic rerouting. The test proved that Ethernet over MPLS services with the use of Fast Reroute (Link Protection) and BFD can offer restoration times close to traditional TDM networks, even when loss of signal cannot trigger rerouting (when rerouting is triggered by a loss of signal event the restoration times are even shorter, as proved in Test 4).
- It must be noted that the tests were conducted with rather extreme BFD parameters (minimum interval 10 ms and multiplier 2), with which BFD will not scale to a large number of sessions. When using less extreme BFD parameters, the restoration times can be longer than measured in the test. For example, if the minimum interval were extended to 100 ms (which means sending a BFD packet every 100 ms and waiting 100 ms for a reply packet before declaring the packet to be lost) and the multiplier value kept at 2 (which means that two consecutive BFD packets must be lost before the BFD session is brought down), the BFD would be unable to detect a failure in less than 200 ms.

2.11 Test 6: Multi-Domain 802.1ag and MPLS Trace Test

2.11.1 Test Setup

The test should prove proper operation of the Connectivity Fault Management and MPLS Trace in a multi-domain network in the provider domain and operator domain.

This test uses the so-called “Option C” for inter-Autonomous System (AS) connectivity – based on labelled unicast between ASs.

The test should be repeated for BGP-signalled VPLS, BGP-signalled VPWS and LDP-signalled VPWS between the PE routers.

Routers PE1, P1 and ASBR1 are located in AS1. Routers ASBR2 and PE2 are located in AS2.

The topology of the testbed is shown in Figure 2.8. Any kind of CE device can be used to ensure the state of the client interfaces of the PEs is UP.

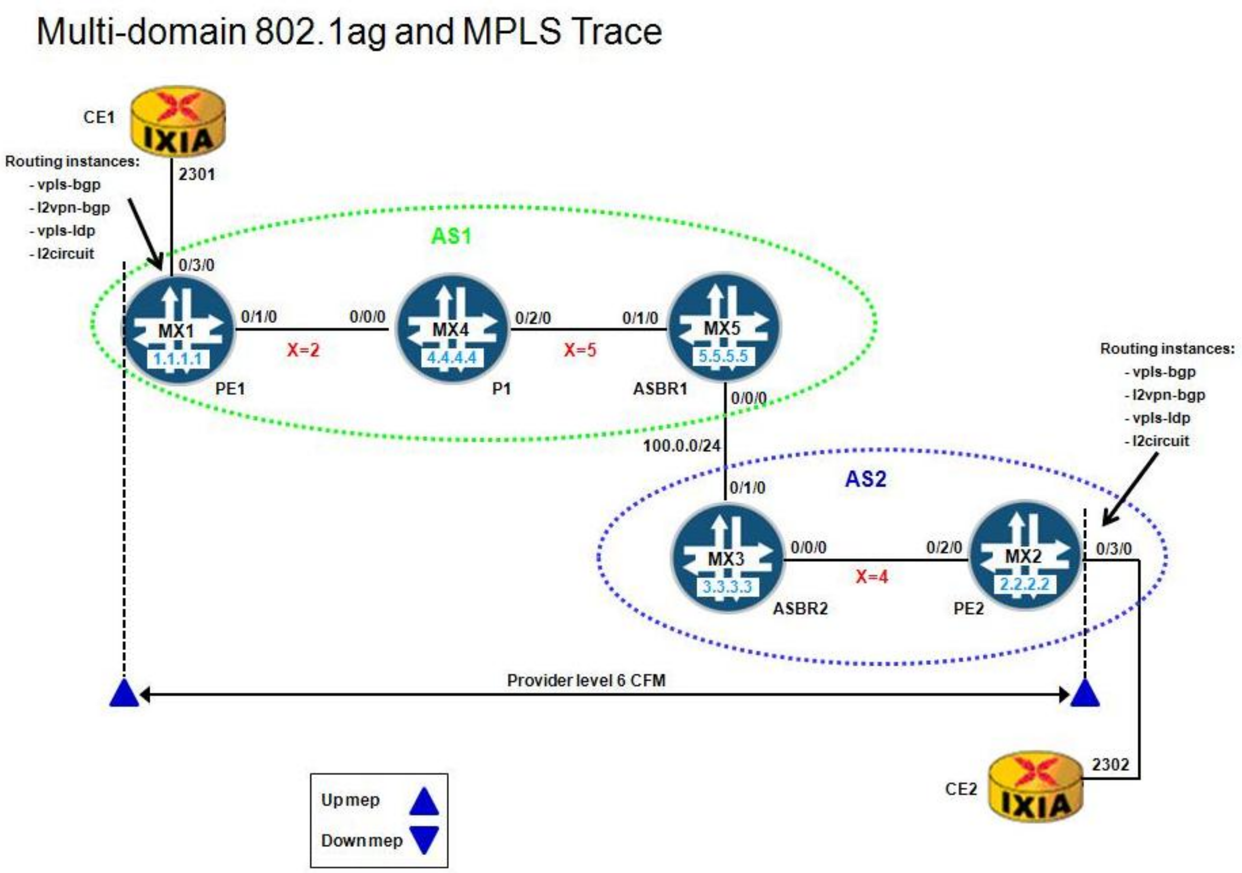


Figure 2.8: Testbed topology for Test 6

2.11.2 Configuration

Measurement end points (MEPs) should be located on PE devices (for the provider domain) and on PE1 and ASBR1 devices (for the operator domain). MEPs should be associated with the VPLS or VPWS instance between PE1 and PE2.

For the LDP-signalled VPWS test, static LSP stitching can be used on the border between ASs. For the BGP-signalled VPLS and BGP-signalled VPWS tests, external Border Gateway Protocol (eBGP) between Autonomous System Border Routers (ASBRs) can be used to signal labels.

2.11.3 Test Description

The test consisted of the following steps:

1. Verify operation of the continuity check protocol between PE1 and PE2 (provider domain) and PE1 and ASBR1 (operator domain).

2. Verify operation of the linktrace protocol between PE1 and PE2 (provider domain) and PE1 and ASBR1 (operator domain).
3. Check the result of MPLS Trace from PE1 to PE2 using the same LSP the VPLS or VPWS uses.
4. Change the admin status of one of the ASBR2-facing interfaces of the PE2 router to DOWN in order to emulate a fibre cut between ASBR2 and PE2.
5. Check the reaction of the CFM to the loss of connectivity.
6. Change the admin status of one of the ASBR2-facing interfaces of the PE2 router to UP.
7. Check the reaction of the CFM to the service restoration.

2.11.4 Expected Results

The test should prove the ability of the routers to provide multi-domain Ethernet over MPLS services and verify whether OAMs (801ag and MPLS Trace) can be used to monitor Ethernet over MPLS multi-domain services.

2.11.5 Results

- The continuity check protocol and linktrace protocol were properly established in the provider domain (between PE1 in AS1 and PE2 in AS2).
- The continuity check protocol properly indicated loss and restoration of connectivity between PE1 and PE2.
- The continuity check protocol and linktrace protocol could not be used in the operator domain (between PE1 and ASBR1) because there was no Layer 2 interface on ASBR1 on which a MEP could be located. The linktrace protocol was not able to find the location of the outage. BFD was used instead of linktrace to verify that the connectivity in the operator domain (AS1 – between PE1 and ASBR1) was not affected by the outage.

  Sample output:

  ```
  jnpr@MX1# run show bfd session extensive
                                                 Detect   Transmit
  Address              State    Interface        Time     Interval     Multiplier
  5.5.5.5              Up                        0.150     0.050           3
  ```

- MPLS Trace could not be used between PE1 and PE2 as it is not supported for inter-AS Option C yet. MPLS Trace was used between PE1 and ASBR1 to verify proper connectivity in the operator domain (AS1).

2.11.6 Test Conclusions

- For inter-AS Option C, continuity check protocol can be used to verify the state of the service in the provider domain (between PEs) but not in the operator domain. Other tools (like BFD) must be used to monitor the state of the service in the operator domain. This is due to the lack of a Layer 2 interface inside the EoMPLS service on the boundary between operator domains and is technology specific rather than vendor specific.

- MPLS Trace can be used in the operator domain (inside a single AS) but it is not supported in the provider domain (multiple ASs). This is due to the fact that separate LSPs are created in each domain and there is no a single LSP traversing both operator domains (the whole provider domain) which can be traced with MPLS Trace. This is a technology-specific issue and is independent of the vendor's implementation. Using MPLS Trace in the provider domain requires future standardisation and implementation efforts.

2.12 Test 7: Multi-Domain Restoration Test

2.12.1 Test Setup

The test should prove proper restoration of BGP-signalled multi-domain VPLS and BGP-signalled VPWS when a connection between domains fails. For the restoration, traffic should be rerouted to an alternative connection between the domains.

The test should be repeated for BGP-signalled VPLS and BGP-signalled VPWS between the PE routers.

The topology of the testbed is shown in Figure 2.9. Any kind of CE device can be used to ensure the state of the client interfaces of the PEs is UP.

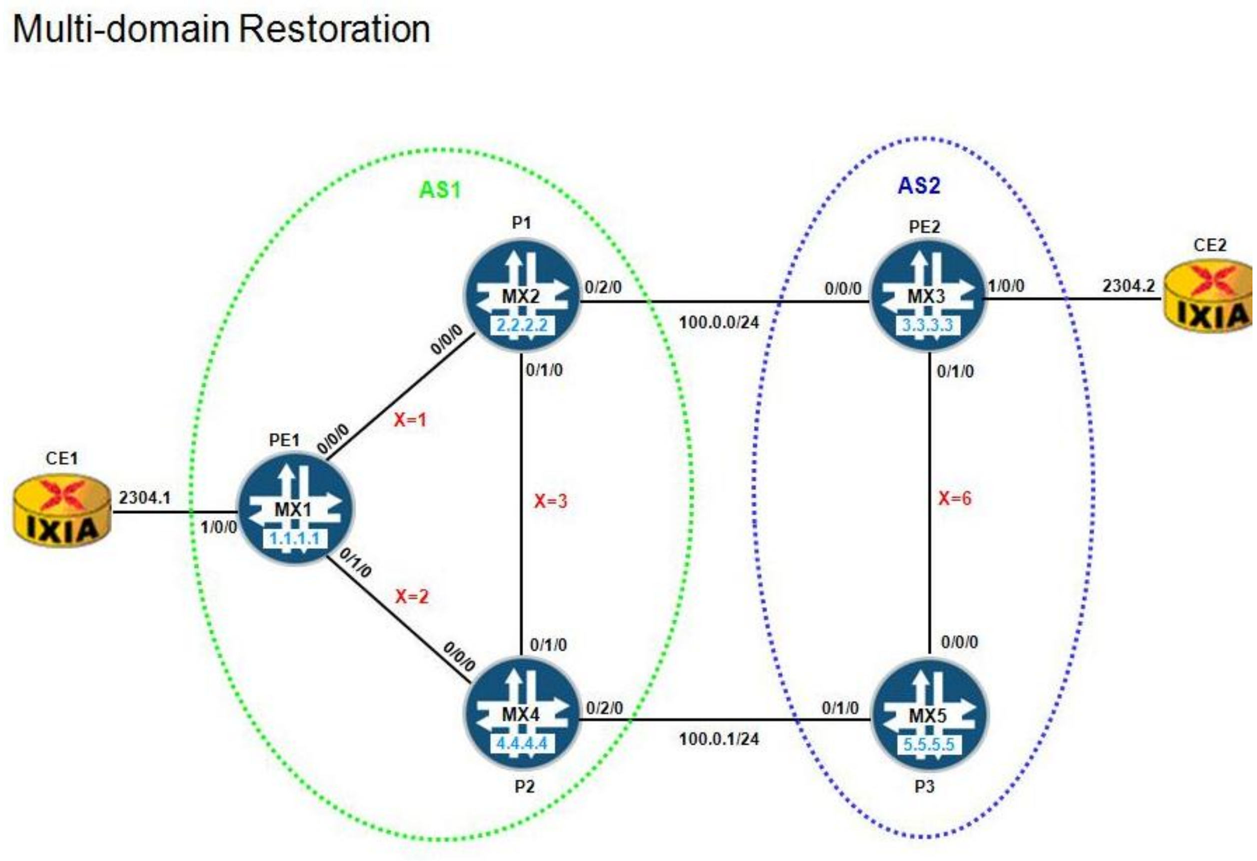


Figure 2.9: Testbed topology for Test 7

2.12.2 Configuration

Routers PE1, P1 and P2 are located in AS1. Routers PE2 and P3 are located in AS2. Routers P1, P2, P3 and PE2 serve as ASBRs. BGP route reflectors are located on routers P2 and P3.

The LSP between PE1 and PE2 for the VPLS or VPWS service should use links PE1-P1 and P1-PE2.

2.12.3 Test Description

The test consisted of the following steps:

1. Disconnect the P1-PE2 link (change the admin state of an interface to DOWN).
2. Demonstrate the LSP being rerouted to the P2-P3 link.

2.12.4 Expected Results

The test should demonstrate restoration of multi-domain VPLS and VPWS services. It is expected that when the P1-PE2 link is disconnected, the LSP is rerouted to an alternative path (using the P2-P3 link).

2.12.5 Results

The LSP was successfully rerouted to link P2-P3 when link P1-PE2 was disconnected. The time needed for service restoration was 31 ms.

2.12.6 Test Conclusions

Service restoration for inter-domain services was successfully demonstrated. EoMPLS services can be rerouted to alternative links when one connection between domains is broken. The restoration time is below 50 ms, which was required by some voice applications and was offered by the legacy TDM networks. Having such restoration time allows services to be migrated from legacy networks without affecting the SLAs.

2.13 Test 8: Multicast in VPLS Test

2.13.1 Test Setup

The tests should prove proper and efficient distribution of multicast streams in a VPLS instance.

The tests should be repeated for BGP-signalled VPLS and LDP-signalled VPLS between the multicast source and multicast destination.

The topology of the testbed is shown in Figure 2.10 and Figure 2.11. The multicast stream can be generated by any device attached to router R1. The multicast receivers are any devices attached to routers R3, R4 and R5, which allow validation of the received multicast stream.

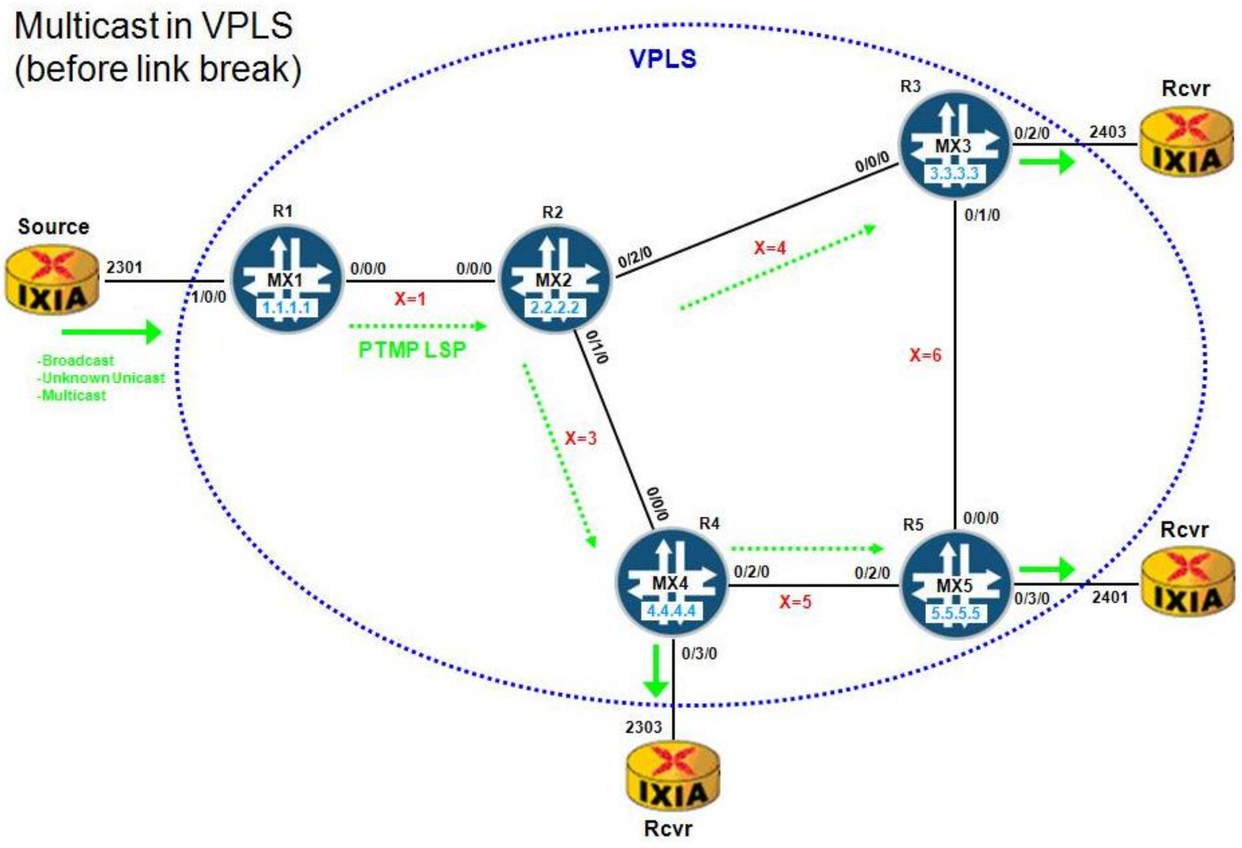


Figure 2.10: Testbed topology for Test 8 (before link break)

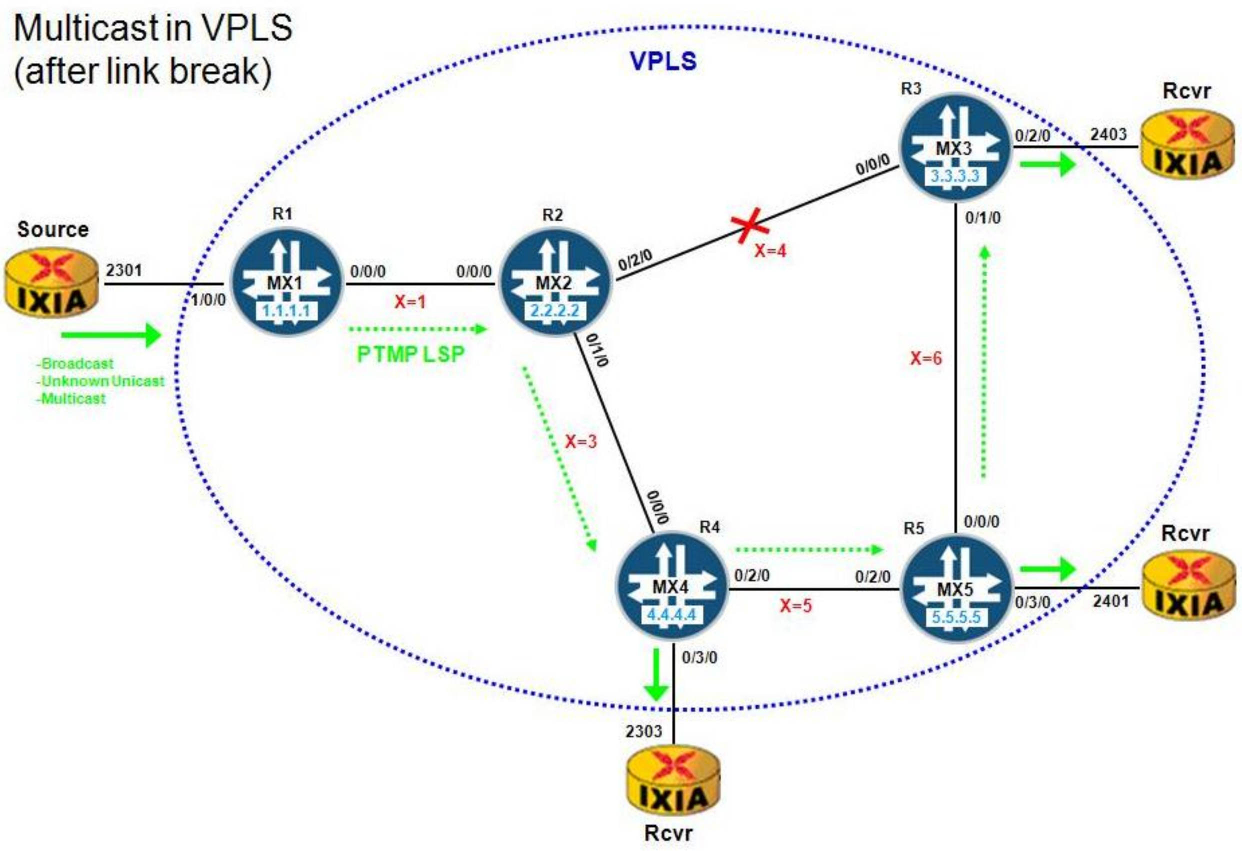


Figure 2.11: Testbed topology for Test 8 (after link break)

2.13.2 Configuration

The multicast transmission should use links R1-R2, R2-R3, R2-R4 and R4-R5.

2.13.3 Test Description

The test consisted of the following steps:

1. Verify that only one copy of the multicast stream is transmitted over links R1-R2, R2-R3 and R2-R4.
2. Verify that all three destination devices receive copies of the multicast stream generated by the source.
3. Disconnect the R2-R3 link (change the admin state of a router interface to DOWN).
4. Verify that the multicast service has recovered and all three destinations receive the multicast stream.
5. Verify that only one copy of the multicast stream is transmitted over the R2-R4 link.

2.13.4 Expected Results

The test should demonstrate efficient and reliable distribution of multicast streams in a VPLS instance. It is expected that all three destinations will receive the multicast stream generated by the source, while the stream is transmitted only once (one copy) on each link in the testbed. When one of the links fails the multicast stream should be rerouted to an alternative path.

2.13.5 Results

- It was verified that all the destination devices received a copy of the multicast stream generated by the source. Only one copy of the stream was transmitted over backbone links.
- The multicast service was successfully restored when the R2-R3 link was disconnected.

2.13.6 Test Conclusions

- The test proved that VPLS together with point-to-multipoint LSPs can be used for efficient distribution of multicast content in an Ethernet network. Point-to-multipoint LSPs allow replication of multicast frames on backbone routers – the routers where replication is really needed, saving the bandwidth on backbone links by not transmitting multiple copies of the same multicast packet over a single link (as in traditional VPLS with the ingress replication technique).
- The same router can be simultaneously used as core (sending a copy of the multicast stream over an MPLS LSP to the next router) and access (sending a copy of the same stream to the client device as native Ethernet).
- The protection and restoration techniques offered by MPLS can be used to guarantee high reliability of the multicast service by using MPLS functionality to restore the multicast services in the event of a link break in the network. As shown by the test, switching multicast traffic to an alternative path did not increase the load of the alternative path, which simplifies the traffic engineering in the network.

2.14 Test 9: VPWS/VPLS over MPLS-TP Test

2.14.1 Test Setup

The tests should demonstrate proper operation of VPWS and VPLS over MPLS-TP LSPs provisioned by a Network Management Station without use of Resource Reservation Protocol (RSVP) or LDP (or any similar protocol) for establishing LSPs.

The topology of the testbed is shown in Figure 2.12. CE devices are not shown in the diagram. Any kind of CE device can be used to ensure the state of the client interfaces of the PEs is UP.

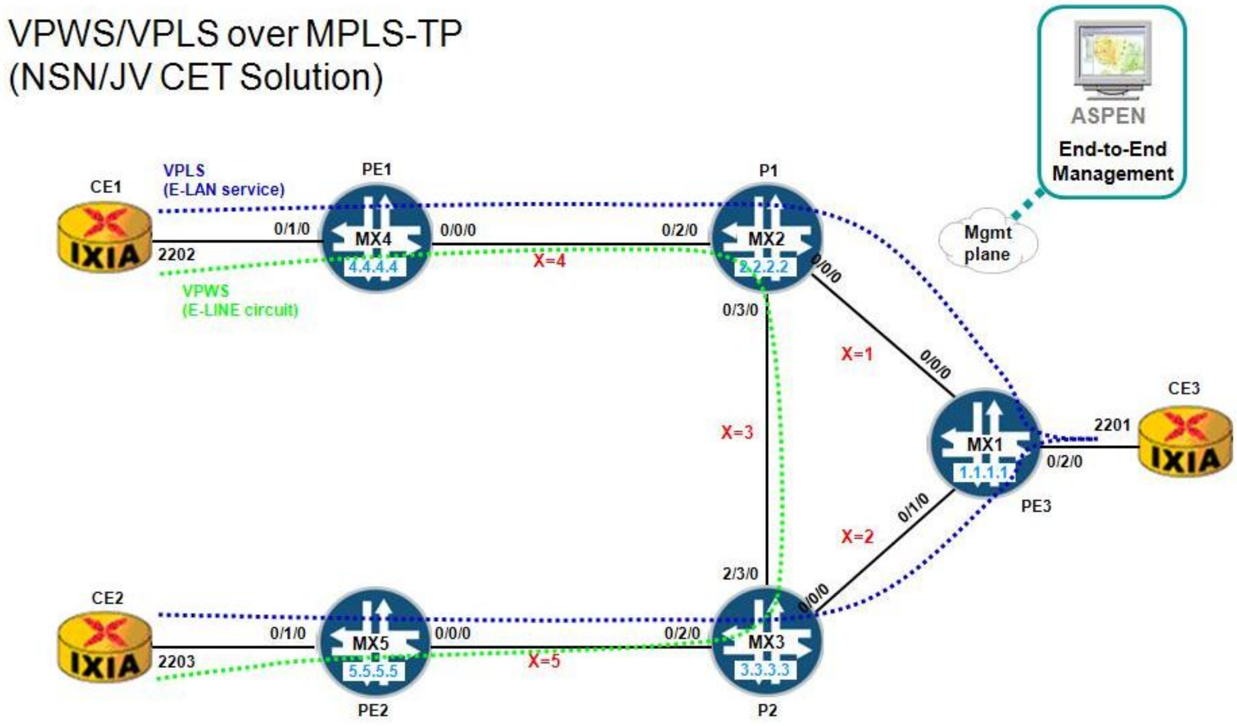


Figure 2.12: Testbed topology for Test 9

2.14.2 Configuration

All LSPs between PE routers should be provisioned by an external Network Management Station without any use of signalling protocols (RSVP, LDP or similar).

2.14.3 Test Description

The test consisted of the following steps:

1. Configure LDP-signalled VPWS between PE1 and PE2 to use an LSP provisioned in point 1.
2. Demonstrate proper operation of the VPWS service.
3. Configure BGP-signalled VPLS between PE1, PE2 and PE3 to use an LSP provisioned in point 1. The same LSP between PE1 and PE2 should be used simultaneously for VPWS and VPLS.
4. Demonstrate proper operation of the VPLS service.

2.14.4 Expected Results

The test should demonstrate proper operation of VPWS and VPLS in an MPLS-TP network with LSPs provisioned by an external Network Management Station without use of any signalling protocol.

2.14.5 Results

- All LSPs were provisioned by ASPEN Network Manager.
- The LSPs provisioned by ASPEN were used as transport for EoMPLS services.
- The test verified proper operation of EoMPLS services over LSPs provisioned by an external Network Management Station.

2.14.6 Test Conclusions

The test demonstrated that EoMPLS services can be supported by MPLS-TP networks in which MPLS tunnels (LSPs) are provisioned by an external Network Management Station.

2.15 Conclusions

The results of the tests described above show that the Ethernet over MPLS technologies and services (i.e. VPWS and VPLS) fulfil the selected requirements for Carrier Class Transport Network Technologies identified in Section 2.2, and offer mechanisms for OAM, multi-domain topologies, protection and restoration and efficient multicast distribution. In cases where the tested EoMPLS and MPLS techniques did not fully satisfy the requirements, alternative mechanisms were proposed to achieve the desired functionality.

The tests show that EoMPLS services can successfully use the mechanisms offered by MPLS for functions such as service restoration and multi-domain operations. This fact, combined with the wide use of MPLS in carriers' packet networks, is the main reason for the growing popularity of EoMPLS services.

The first three tests show that the OAM mechanisms defined for Ethernet (continuity check protocol, link trace and MAC ping) work properly for the EoMPLS services in both customer and provider domains, and can be used to monitor the major service parameters, including availability and packet delay (packet delay measurement is currently supported in the customer domain only). The major limitation for the mechanisms is that the backbone routers of the provider network lack Layer 2 interfaces and therefore are not visible for the Ethernet-based mechanisms. This is a technology-specific issue, independent of the vendor's implementation. The reason for the limitation is that the Layer 2 client traffic is tunnelled into a higher-layer technology and not converted back to Layer 2 on the intermediate routers.In case of an outage the Ethernet OAM can correctly indicate the domain in which the outage is located but is not able to find the precise location of the outage in the provider network. Other mechanisms, mainly IP-based tracing mechanisms, must be used by the operator to pinpoint the cause of the problem.

The next two tests verify the protection and restoration mechanisms available for EoMPLS services.

The multi-domain tests verify the scalability offered by the EoMPLS technology, mainly in terms of connecting different administrative domains, which is very important in the research networking environment, where GÉANT and NRENs work together to offer the user an end-to-end service.

Test 8 shows that multicast services can be successfully provided in an EoMPLS service instance and Test 9 shows that the EoMPLS services can be implemented in an MPLS-TP network, which is also within the scope of the GN3 JRA1 Activity.

The test results prove that EoMPLS offers manageable and scalable transmission services with robust protection and restoration mechanisms. The technology can be used in research networks as well as carriers' networks to support data transmission services for the research community and other users.

As shown by the tests, EoMPLS fulfils the major requirements for a Carrier Class Transport Network, including service reliability and multi-domain topologies in which several carriers co-operate to provision an end-to-end service spanning multiple networks. In addition, EoMPLS fulfills the CCTNT requirement of having mechanisms for service monitoring and maintenance by using mechanisms offered by different layers of an IP-MPLS network combined with Ethernet-specific tools.

3 Ethernet OAM

3.1 Overview

Wide-area Ethernet connections are gaining more and more popularity in the networking industry, including its academic branches. Such Ethernet connections are often replacing traditional Synchronous Digital Hierarchy (SDH), frame relay and T1/E1 circuits because of their lower cost and finer capacity granularity. However, manageability of wide-area Ethernet connections is still a challenge, as Ethernet was created purely as a Local Area Network (LAN) technology without any Operation, Administration and Maintenance (OAM) features typical of carrier-grade transport technologies such as SDH and Optical Transport Network (OTN), which help to monitor and troubleshoot connections. According to the first “Transport Network Technologies Study” [DJ1.1.1], until wide area Ethernet provides diverse and feature-rich OAM functionality, it cannot be considered a carrier-grade transport technology.

Fortunately, considerable efforts by standards bodies and vendors have been made recently to overcome this Ethernet drawback and to convert the technology into Carrier Ethernet. As a result, a core set of Ethernet OAM standards has been developed and ratified, and many vendors have already implemented these features in their equipment. However, experience in the use of Ethernet OAM functions among National Research and Education Network (NREN) network operation staff is still limited and, as a result, the available functions remain largely unused.

This chapter presents the results of a trial of Ethernet OAM functions over a dedicated wide-area testbed established by five NRENs in the context of JRA1 Task 1 activity. The trial was conducted to bridge the gap and gain some initial experience in monitoring and troubleshooting wide-area point-to-point Ethernet connections on an end-to-end and per-segment basis using the Ethernet OAM functions available in the latest equipment. The trial lasted for six months and confirmed that it is possible to monitor and visualise the health and performance of point-to-point Ethernet connections in a multi-domain environment using the Connectivity Fault Management (CFM) IEEE 802.1ag / Y.1731 functionality of network equipment.

The wide-area multi-site Ethernet testbed was established by five NRENs (JANET, NORDUnet, PSNC, SURFnet and CESNET) and one JANET customer (Essex University). Sources of OAM data were provided by existing equipment of participating NRENs in addition to dedicated OAM agents deployed for the trial. The existing NREN equipment was from different vendors, including Ciena, Extreme Networks, Cisco, and Brocade. As not all of the equipment available for the trial supported the required set of Ethernet OAM functions, small OAM agents on loan from vendors (Overture and Accedian Networks) were installed in the trial sites. The links between sites were of different types: all of them had an Ethernet User Network Interface (UNI) at both ends, but in between different combinations of Multi-Protocol Label Switching (MPLS), OTN, and tunnelling through IP were used. Monitoring data was stored and visualised by the cloud-based software system CyPortal from Cyan, Inc.

The trial objectives were:

- To monitor the state (up/down) of point-to-point Ethernet connections on an end-to-end and per-segment basis using the embedded OAM functionality of the equipment under test and of OAM agents.
- To monitor the performance (throughput, latency, and loss) of Ethernet connections.
- To evaluate the possibility of hierarchical state and performance monitoring of the connections under test.
- To evaluate the capabilities of management software systems to store and visualise monitoring data. This objective is very important as in the absence of a management system, a network administrator has no other option but to manually use appropriate command lines to invoke state and performance data periodically, which is not ideal.

To meet these objectives, a series of tests were conducted over the multi-domain testbed during the six-month period from May to October 2011. The chapter describes the testbed and the test scenario and results of each of the seven tests carried out.

Tests 1, 2, and 3 were dedicated to monitoring state, rate, delay, jitter and loss of Ethernet services between participants' testbeds on an end-to-end basis. The Command Line Interface (CLI) of OAM agents was used for these tests. Most of the results of these tests proved to be stable and satisfactory. Two problems were discovered with regard to delay measurement of services where the JANET testbed was involved; these were investigated in Tests 6 and 7.

Test 4 investigated the capabilities of CyPortal from Cyan, Inc. [CyanInc] – a cloud-based web portal that stores and visualises measurement data. CyPortal was used throughout the trial period and proved to be a useful system for providing managed wide-area Ethernet servicers. The company (Cyan Inc.) was keen to test their new service in academic environment suggesting a free trial.

Test 5 investigated the feasibility of establishing two-level hierarchical sessions to monitor an Ethernet service from both customer and provider perspectives. The provider monitoring session was established between network equipment of different vendors – Ciena and Extreme. The positive results of this test show that the Ethernet OAM features of the latest network equipment can provide seamless monitoring in a multi-vendor environment.

3.2 Technology Briefing

Ethernet Operation, Administration and Maintenance (OAM) functionality includes those capabilities that allow a service provider to create, monitor and troubleshoot Ethernet links and services in a standardised fashion. It helps service providers offer end-to-end service assurance across the IP/MPLS core, the Ethernet Metro, and to the customer premises. Ethernet OAM standards can be divided into three groups, which correspond to three stages of network maintenance:

- Service Assurance.
- Service Monitoring.
- Service Troubleshooting.

The standards in each group are described in the sections that follow.

3.2.1 Service Assurance Standards

These standards allow testing of whether a provisioned Ethernet service meets the requirements described in a Service Level Description (SLD). The standards can be split into two sub-areas:

- Service Definitions standards.
- Service Activation standards.

3.2.1.1 *Service Definitions Standards*

Service Definitions standards describe:

- A service type, e.g. point-to-point or point-to-multipoint.
- A service bandwidth profile, which includes such parameters as Committed Information Rate (CIR) and Excess Information Rate (EIR) for several classes of service.
- Service performance parameters, e.g. frames delay, jitter, and loss.

Service Definitions standards are of great importance as their absence can easily lead to confusion: for example, CIR might be measured for User Datagram Protocol (UDP) payload rate or for Ethernet frame rate and the results will be very different for the same service.

There are three major standards in this area – MEF 10.2 [MEF 10.2] and MEF 10.2.1 [MEF 10.2.1] from the Metro Ethernet Forum (MEF) and G.8011 [ITU-T G.8011] from the ITU-T – which give precise formal definitions of Ethernet service types and parameters. These standards are aligned with each other and give similar and non-contradictory definitions of Ethernet services.

3.2.1.2 *Service Activation Standards*

Ethernet OAM standards of this type describe how to check whether a provisioned Ethernet service complies with its SLD parameters. Until recently there were no standards in this area that took into account Ethernet-specific services; the most popular standard among tester vendors, RFC 2544, is an IP-centric standard and its use for Ethernet services could lead to ambiguous results. Fortunately, in spring 2011, the ITU-T approved the Y.1564 recommendation “Ethernet service activation test methodology” [ITU-T Y.1564], which bridges this gap.

Figure 3.1 below illustrates a simple disruptive on-demand procedure described in Y.1564 which tests connectivity and throughput up to CIR and EIR and policing limits by injecting traffic into an Ethernet connection. Traffic rate and performance parameters are measured according to Y.1563 [ITU-T Y.1563] definitions.

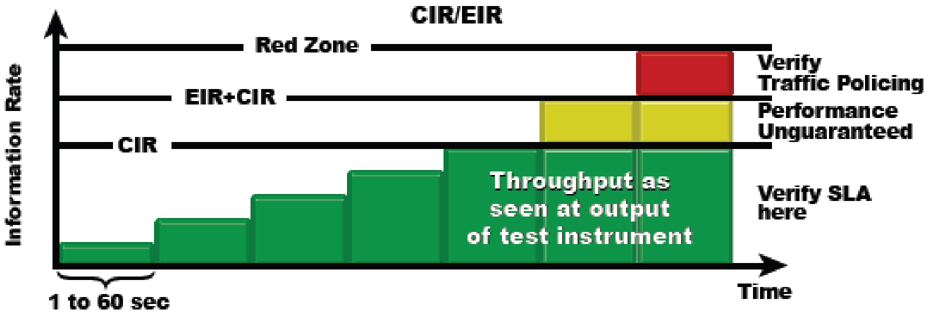


Figure 3.1: Throughput test according to Y.1564 recommendation

3.2.2 Service Monitoring Standards

Two major standards have been developed for this area:

- IEEE 802.1ag: “Connectivity Fault Management” [IEEE 802.1ag].
- ITU-T Y.1731: “OAM functions and mechanisms for Ethernet based networks” [ITU-T Y.1731].

For the purpose of monitoring Ethernet services, IEEE 802.1ag defines a hierarchy of maintenance sessions which allows the independent monitoring of different segments of the same service by different entities such as a customer, a service provider or an operator (Figure 3.2).

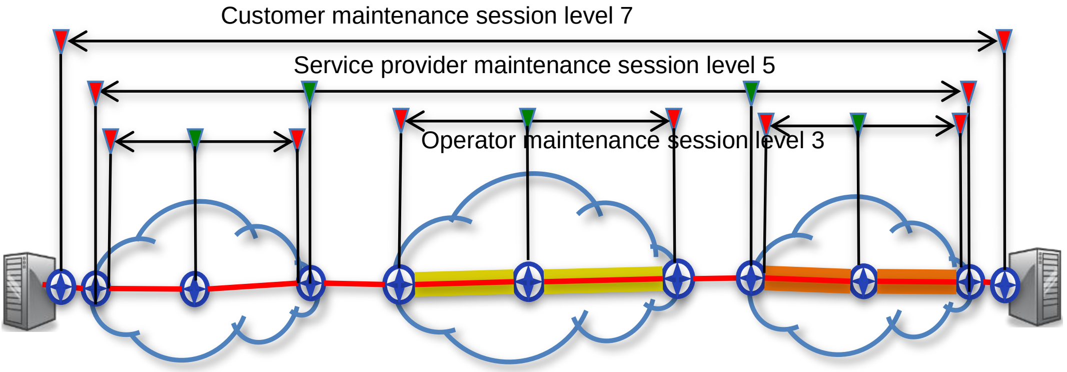

Customer maintenance session level 7
Service provider maintenance session level 5
Operator maintenance session level 3

Figure 3.2: A hierarchy of maintenance sessions according to IEEE 802.1ag specification

Each maintenance session uses a separate sequence of heartbeat messages called Continuity Check Messages (CCMs) to monitor the health of a service.

The Y.1731 recommendation extends the IEEE 802.1ag specification. It includes the description of the CFM CCM function (called ETH-CC in Y.1731) and adds to it several performance monitoring functions:

- Loss Measurement (LM),
- Delay Measurement (DM).

These functions allow active measurements of delay, delay variation and loss between connection end or intermediate points to be carried out.

Note: As Y.1731 covers all IEEE 802.1ag functionality and adds some extra functions, the name Y.1731 will be used in this chapter to refer to the IEEE 802.1ag / Y.1731 combined functionality when appropriate.

3.2.3 Service Troubleshooting Standards

Both the IEEE 802.1ag and Y.1731 standards describe two protocols (“functions” in Y.1731 terminology) that may be used for service troubleshooting:

- Linktrace Protocol – allows a path to be traced in an IP traceroute manner and can report on passing intermediate maintenance points along a service path.
- Loopback Protocol – allows connectivity to be checked with connection end and intermediate maintenance points in an IP ping manner.

In addition to these protocols, a number of signals carried in IEEE 802.1ag and Y.1731 frames have been introduced, which may be helpful in understanding the reason for a fault. These include the Remote Defect Indicator (RDI) and the Alarm Indication Signal (AIS).

3.3 Testbed Infrastructure

This section describes the general testbed setup, the setups of each participant’s testbed, and the links between the testbeds,

3.3.1 General Setup

Ethernet OAM tests were carried out on a multi-NREN testbed, shown in Figure 3.3 below.

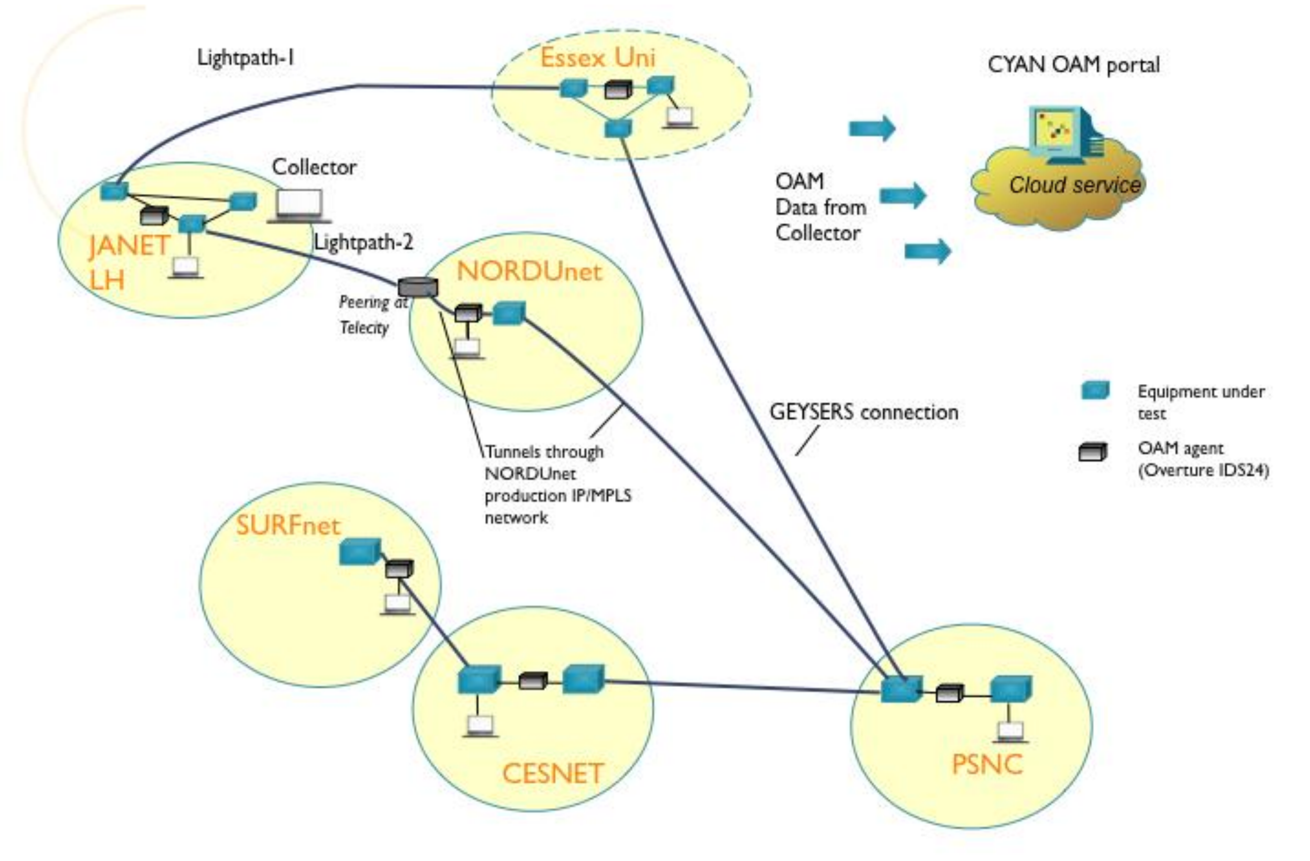


Figure 3.3: The multi-domain Ethernet OAM testbed

The testbed consisted of six partners' testbeds, belonging to five NRENs. The partners' testbeds were built on equipment from different vendors and connected by links of different types.

The testbeds also included Overture ISG24 switches (one per testbed, on loan), which support all major Ethernet OAM protocols and were used as reference OAM agents for OAM management end and intermediate points. The existence of such reference agents provided a useful opportunity to compare implementations of Ethernet OAM functions from different vendors.

A cloud-based web-portal (CyPortal from Cyan, Inc., [CyanInc] on loan) played the role of a central repository for the OAM data generated during the tests. Cyan tested CyPortal with Overture switches, so it was guaranteed that monitoring data from Overture ISG24 switches would be properly collected and visualised by the CyPortal system.

The CyPortal software received monitoring data not directly from ISG24 agents but through a collector, which sat in the JANET Lumen House (JANET LH) testbed. A Dell laptop with special software was used as the collector. The ISG24 agent boxes, the collector, and the CyPortal software interworked in the following way:

- Each ISG24 box performed an ftp push of fresh monitoring data to the collector every 15 minutes.
- The collector uploaded the data collected from all the ISG24 boxes to the CyPortal, also using ftp push, every 15 minutes.
- CyPortal stored the new data in its database and visualised it on request in the form of dynamic web pages.

3.3.2 Setup of Participants' Testbeds

The setup of each participant's testbed is described in the sections that follow.

3.3.2.1 *JANET Lumen House (JANET LH) Testbed*

The JANET LH testbed, shown in Figure 3.4 below, was built on two Ciena 311v switches, which support Ethernet, Provider Bridges (PB) and Provider Backbone Bridge Traffic Engineering (PBB-TE) as transport technologies, and Y.1731 as Ethernet OAM functionality. The key testbed elements are described in Table 3.1 below.

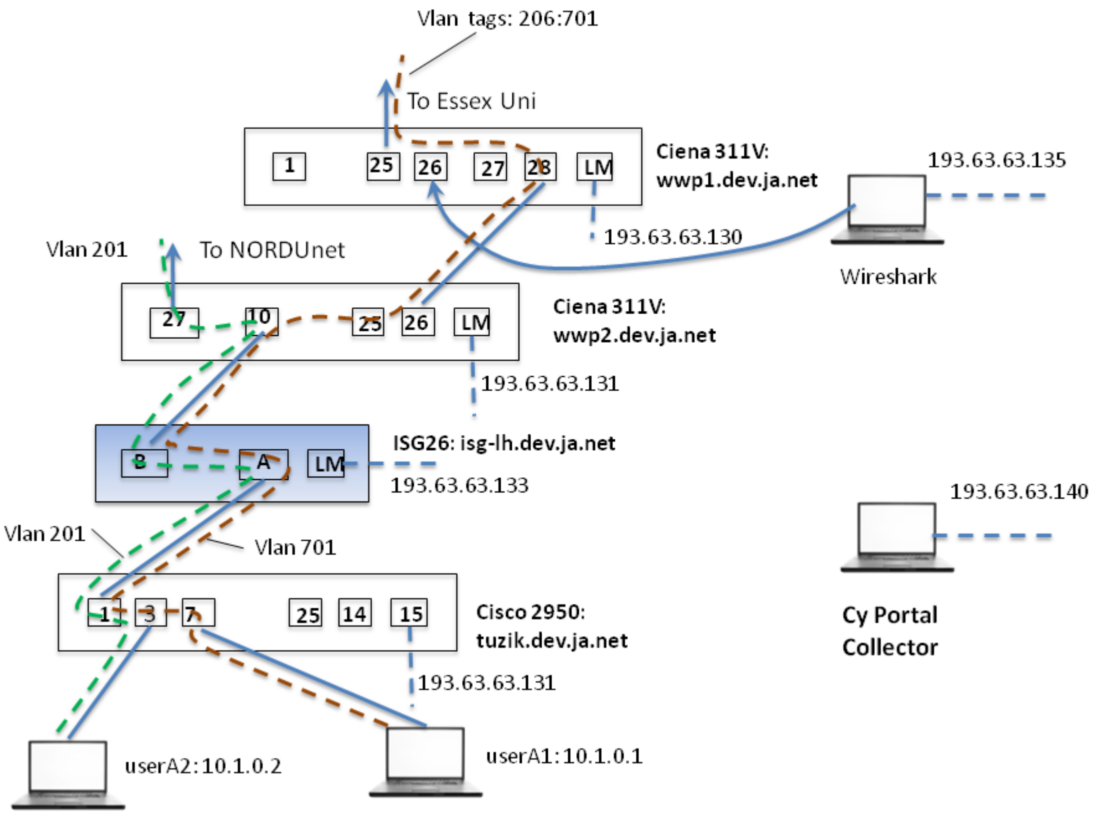


Figure 3.4: JANET LH testbed

Testbed Element Name	Testbed Element Type	Description
wwp1.dev.ja.net	Ciena switch	Connected to a link leading to the JANET Essex University testbed. Added a second (an outer) VLAN ID 206 to JANET LH–Essex University traffic as it required a JANET Lightpath connection used between JANET LH and Essex University. The outer VLAN ID 206 was used by the Lightpath infrastructure to differentiate Ethernet OAM traffic from traffic of other projects entering the Essex University campus network.
wwp2.dev.ja.net	Ciena switch	Connected to a link leading to the NORDUnet testbed in Copenhagen.
tuzik	Cisco 2950 switch	Used for tagging Ethernet traffic coming from two end hosts: userA1 and userA2. The tuzik switch tagged userA1 traffic with VLAN ID 701, and userA2 traffic with VLAN ID 201. These VLAN IDs represented the Ethernet services under monitoring.
userA1	End host	Communicated with the JANET Essex University testbed.
userA2	End host	Communicated with the NORDUnet testbed.
-	ISG26 agent box	Located between Cisco tuzik and Ciena wwp2 switches. This position allowed the agent to see traffic of both Ethernet services under monitoring and to inject its own Y.1731 frames into user data. (Only the JANET LH site used an ISG26 box. All other sites used ISG24 boxes. The OAM functionality of the ISG26 and ISG24 boxes was the same.)

Table 3.1: JANET LH testbed elements

3.3.2.2 *JANET Essex University Testbed*

The JANET Essex University testbed (Figure 3.5) was built on two Extreme BlackDiamond® 12K switches and several Brocade switches (not shown). The Extreme switches supported PB, PBB-TE and MPLS as transport technologies, and Y.1731 Ethernet OAM protocols.

VLAN 701:	*xtreme 4A*	*ports 2:3T, 2:4T, 2:9U*
	xtreme 4B	*ports 1:2 T*
	CFM	*port 2:3,md3,mep711*
SVLAN 800:	*xtreme 4B*	*ports 1:2U, 2:1 T*

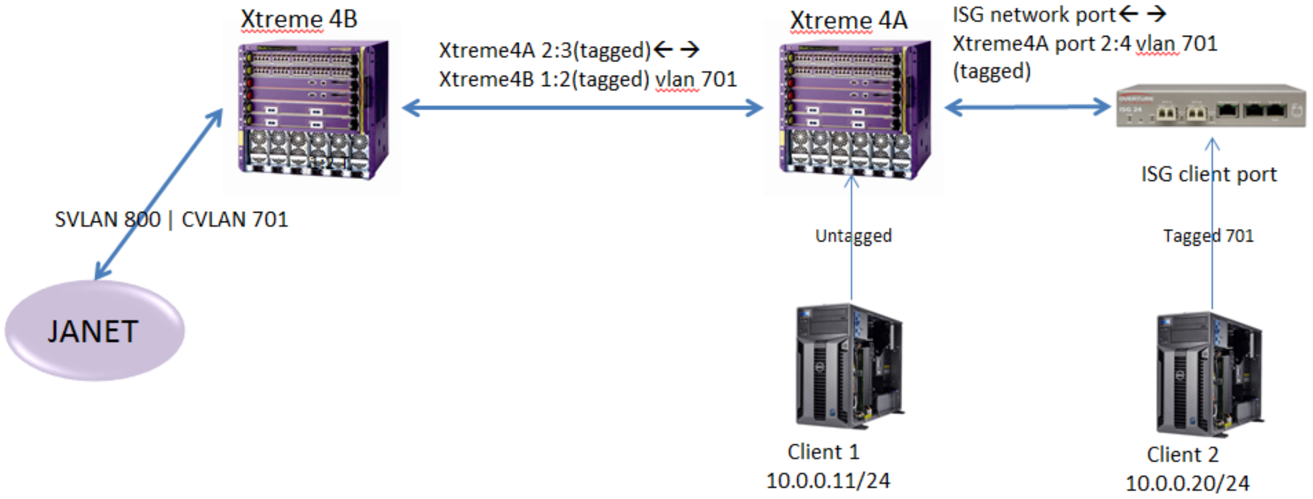


Figure 3.5: JANET Essex University testbed

The frames from JANET LH arrived at the JANET Essex University testbed double-tagged with VLAN ID 800 as an outer tag (it was translated by the Lightpath infrastructure from its original 206 value) and with VLAN ID 701 as an internal tag.

Two end hosts were used:

- Host 10.0.0.11 produced untagged frames which were then tagged by the Extreme 4A switch.
- Host 10.0.0.20 produced tagged VLAN ID 701 frames.

The ISG24 agent box sat between the Extreme 4A switch and host 10.0.0.20.

3.3.2.3 *NORDUnet Testbed*

Over the trial period, NORDUnet did not have equipment of their own that supported Y.1731 functions. The NORDUnet testbed shown in Figure 3.6 therefore consisted of an ISG24 agent box and host 10.2.0.1. The agent box was connected to a router, which terminated IP/MPLS tunnels from other participants established through the NORDUnet production IP/MPLS network.

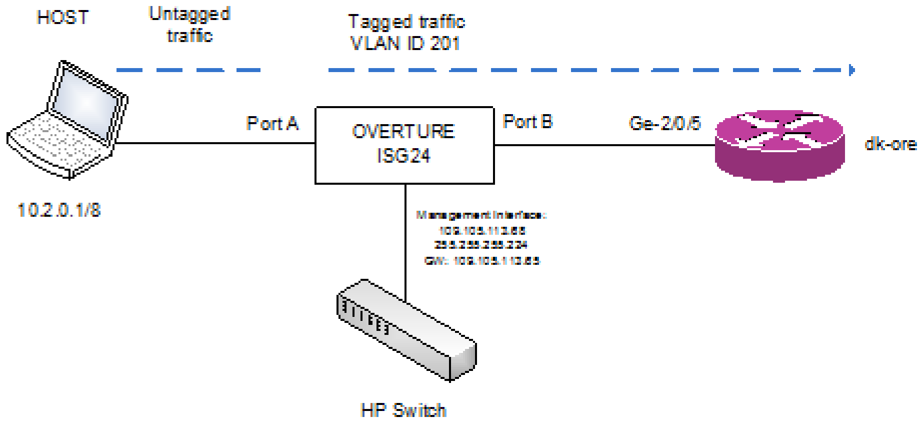


Figure 3.6: NORDUnet testebd

The ISG24 box tagged frames with the agreed VLAN IDs (e.g. VLAN ID 201 for traffic on the LH–NORDUnet connection).

3.3.2.4 *PSNC Testbed*

The PSNC testbed (Figure 3.7 below) was built on two Brocade MLX-8 switches with an ISG24 agent box in between. VLAN ID 339 was used for communication with CESNET, and VLAN ID 602 for communication with NORDUnet.

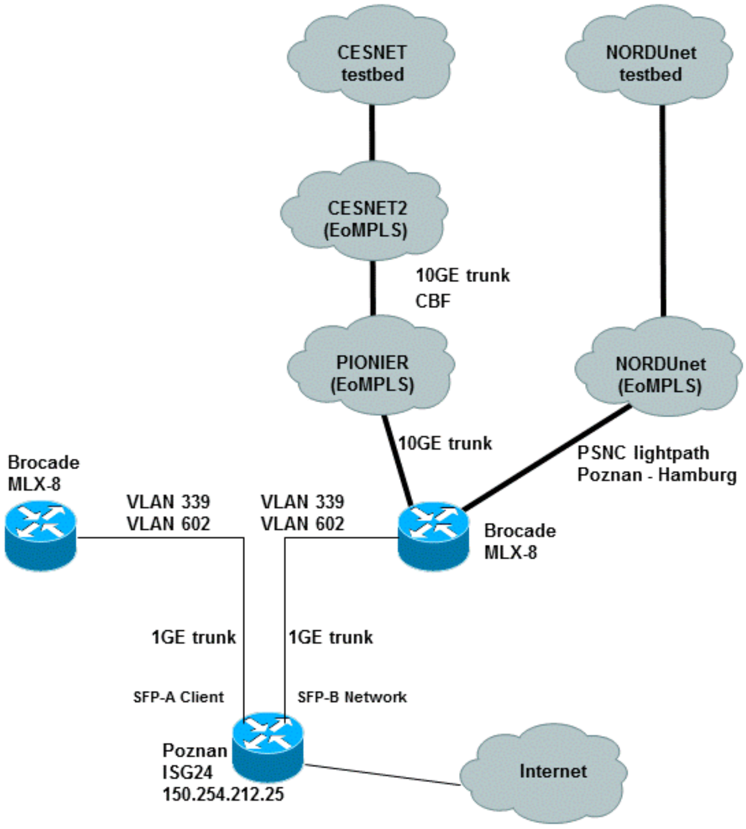


Figure 3.7: PSNC testbed

3.3.2.5 *CESNET Testbed*

The CESNET testbed (Figure 3.8) was built on two switches: a Cisco Catalyst 6503 and a Force10 E300. An ISG24 agent box sat between the two switches and was able to use VLAN ID 333 to communicate with the SURFnet testbed and VLAN ID 339 to communicate with the PSNC testbed. The particular models of the Cisco 6503 and Force10 E300 switches used in the CESNET testbed did not support Y.1731 functionality.

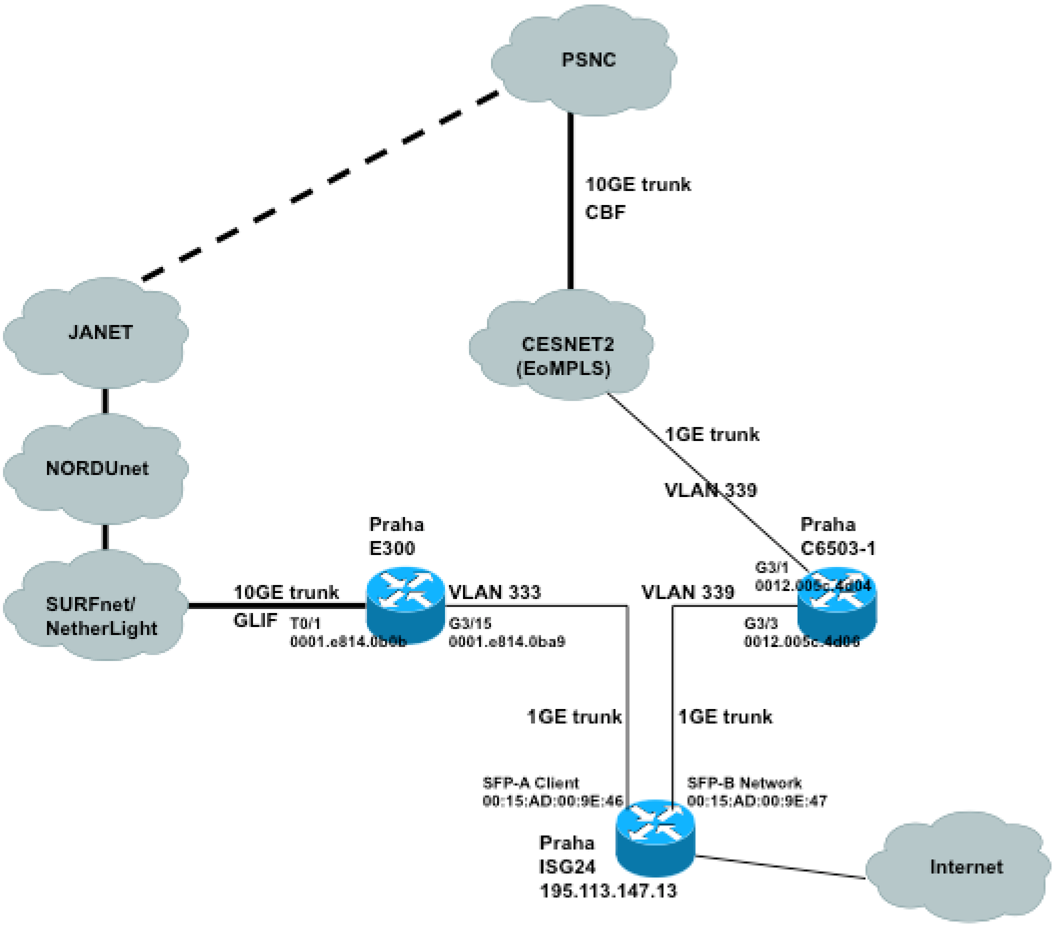


Figure 3.8: CESNET testbed

3.3.2.6 *SURFnet Testbed*

Two Avaya (Nortel) Metro Ethernet Routing Switch (MERS) 8600 switches were used to create the SURFnet testbed (Figure 3.9). An ISG24 agent box was connected to an untagged port of the E8K1T switch. The E8K02 switch was connected to a link going to CESNET and translated VLAN ID 333 frames assigned to SURFnet-CESNET connectivity into internal VLAN IDs used by the testbed switches (which were involved in other tests and hence already had some VLAN IDs in use).

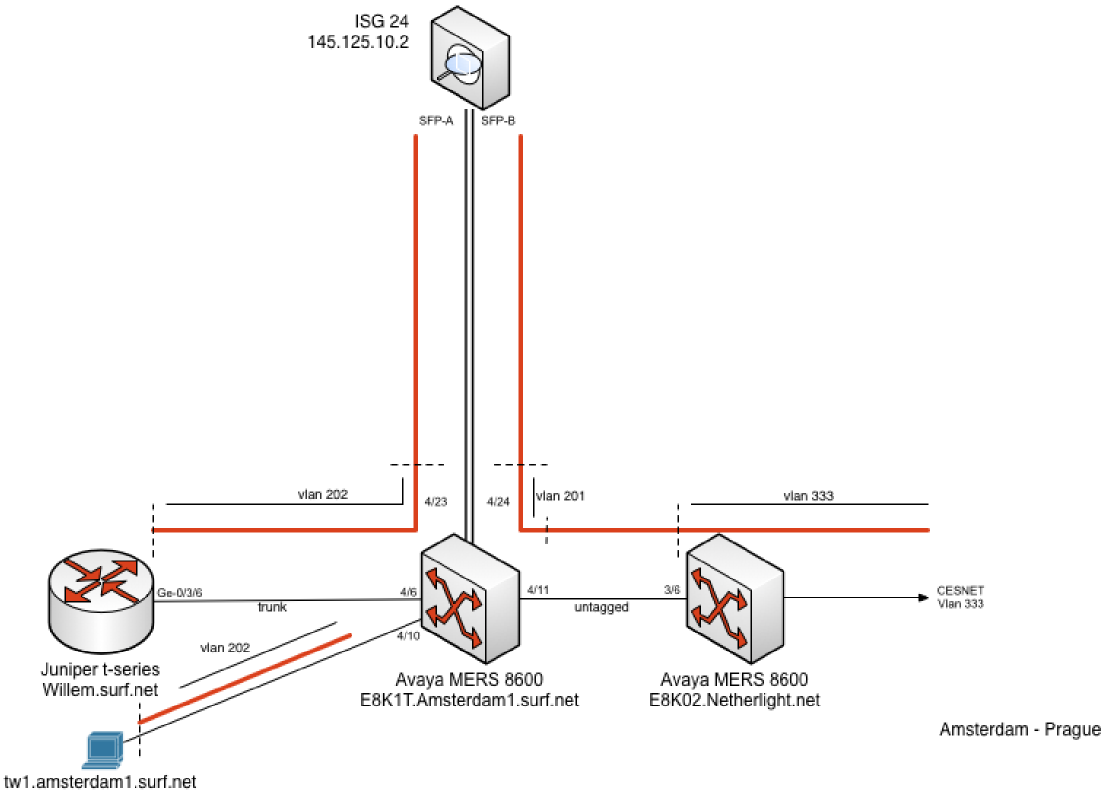


Figure 3.9: SURFnet testbed

3.3.3 Links between Testbeds

The links connecting the trial partner testbeds were of different types, as shown in Table 3.2 below; the Internal Name column gives the alternative link name sometimes used in the trial and in this document.

Link	Description	Internal Name
JANET LH–JANET Essex University	JANET Lightpath Ethernet over MPLS (EoMPLS) connection. This type of connection transfers Ethernet OAM frames transparently. This feature was tested during the JANET Carrier Ethernet project (described in Chapter 5 *Provider Backbone Bridge Traffic Engineering* on page 101). The connection is VLAN based, with double-tagged frames: 206/(701-799) at LH and 800/(701-799) at Essex University ingress points.	janet-essex
JANET LH–NORDUnet	JANET Lightpath EoMPLS connection, port based (1GE).	janet-copenhagen
JANET Essex University–PSNC	Based on a GEYSERS project connection between these organisations. Implemented as a VLAN on a lightpath provided by	–

Link	Description	Internal Name
	GÉANT.	
NORDUnet–PSNC	Implemented as tunnels through NORDUnet IP infrastructure and a VLAN over the PSNC lightpath towards NORDUnet.	copenhagen-poznan
PSNC–CESNET	Implemented as EoMPLS services in the PSNC and CESNET networks. The two domains were interconnected as native Ethernet with single-tagged frames.	–
CESNET–SURFnet	Implemented as a 1G Ethernet VLAN carried over the SDH 10G link.	prague-amsterdam

Table 3.2: Links between testbeds

3.4 Test 1: Monitoring of Ethernet Services State

3.4.1 Test Setup

The test goal was to establish Continuity Check Message (CCM) sessions between ISG24 agent boxes for Ethernet services between different participant sites and investigate how they allow the status (Up/Down) of these services to be monitored.

The following single-segment point-to-point Ethernet services were chosen for monitoring:

- JANET LH–JANET Essex University (inner VLAN ID 701; outer VLAN IDs 206 for LH end and 800 for Essex University end).
- JANET LH–NORDUnet (VLAN ID 201).
- PSNC–CESNET (VLAN ID 339).
- PSNC–NORDUnet (VLAN ID 602).
- CESNET–SURFnet (VLAN ID 333).

The ISG boxes in the JANET LH, CESNET and PSNC testbeds were involved in monitoring two services, while the ISG box in the JANET Essex University testbed monitored one.

The positions and connections of the ISG boxes were as shown in the testbeds' respective figures in Section 3.3.2 above. The management port of each ISG box had a public IP address connected to the Internet, so that every trial participant could manage every ISG box remotely. The staff of Overture (the ISG vendor) used the remote access facility to do the initial configuring of the ISG boxes.

3.4.2 Configuration

For each Ethernet service a CCM session was established between two ISG boxes sitting in the two testbeds that terminated the service.

Configuration of an ISG box is carried out by means of web access to a box and filling in web forms.

The steps to configure a CCM session between two ISG boxes are as follows:

1. Specify the names of Maintenance Domain (MD) and Maintenance Association (MA) and the MA levels.

   These three parameters have to be identical in both ISG boxes for a session to be established successfully.

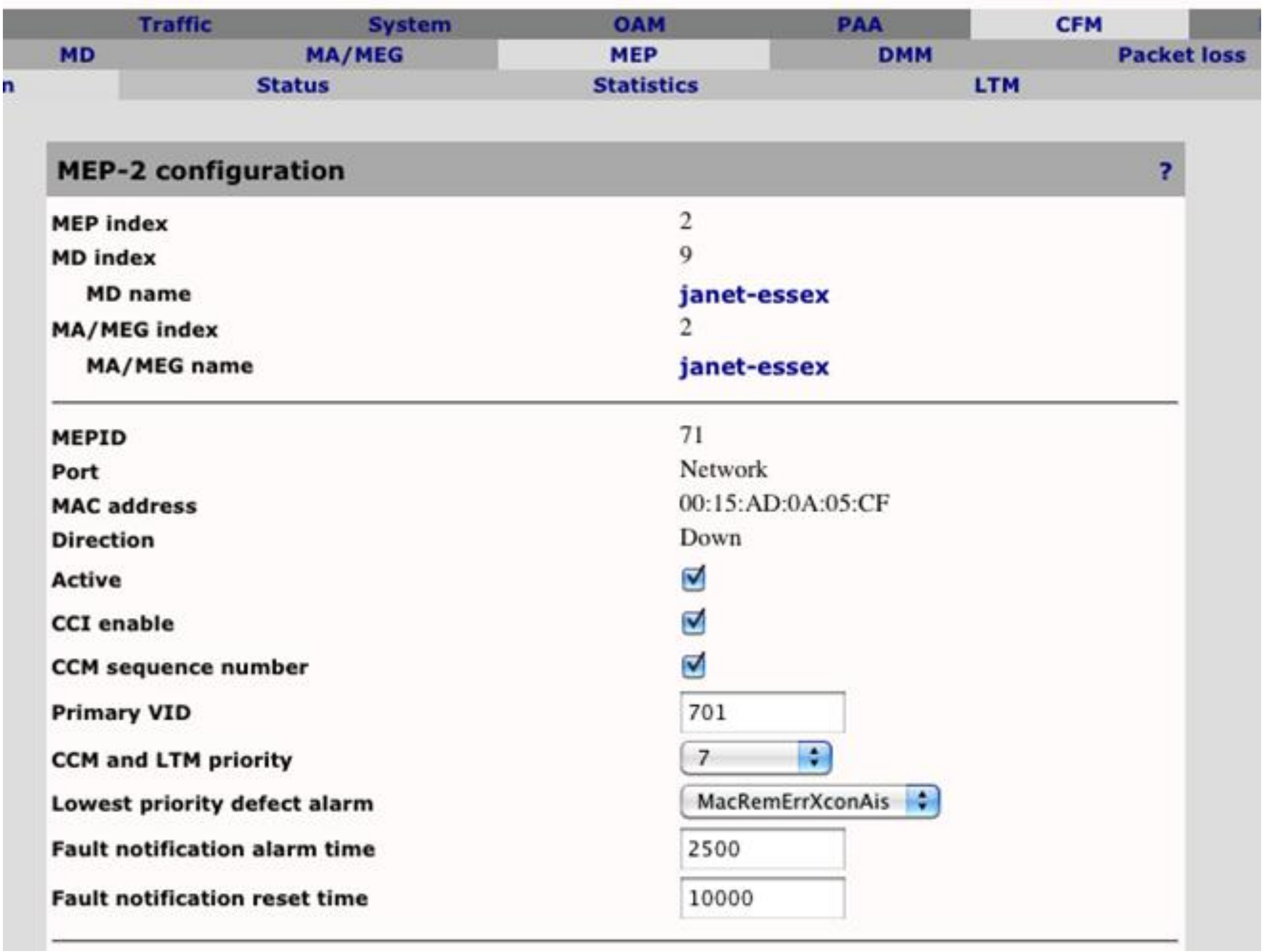
   The configurations of MDs and MAs with level 5 and appropriate names were entered in all six ISG boxes in the trial testbeds. For example, the JANET LH box was configured to support MD janet-essex and MA janet-essex with level 5 for monitoring the JANET LH–JANET Essex University service.

2. Set up a Maintenance Entity Group End Point (MEP) for each MA.

   An example of a MEP configuration form is shown in Figure 3.10 below. This example illustrates MEP parameters for the janet-essex MA for the JANET LH ISG. A MEP ID should be unique for a particular combination of MA/level number. In the case of the janet-essex MA, MEP ID 71 was chosen for the JANET LH ISG, and MEP ID 72 for the JANET Essex University ISG.

   The Primary VID parameter of a MEP is very important as it determines the encapsulation of CCM frames that a MEP generates towards its remote counterpart MEP. The encapsulation should be the same as for customer frames transported along the service being monitored, to guarantee that CCM frames go along the same route as service ones. In the case of the janet-essex MA, all CCM frames were single-tagged with VLAN ID 701, as service frames going through the ISG had this encapsulation (a second VLAN ID 206 is added further along the route to Essex University by the wwp1 switch, as stated in Table 3.1). As can be seen from the form in Figure 3.10, the JANET LH ISG box supported only single-tagged CCM frames (or untagged if the Primary VID field is left blank). This was a restriction of the ISG model used for the trial. The restriction meant that the ISG box in the JANET LH testbed could not be positioned after the wwp1 switch (see Figure 3.4) as it would need the ISG box to generate double-tagged frames (206:701).

   The parameters Port and Direction specify from which port a MEP should send CCM frames and in what direction. The direction Down means that frames should be sent down to the network; the direction Up directs frames inside a box (up to an internal protocol stack) towards another box port. In the example in Figure 3.10, MEP 71 is bound to the network port facing a connection to Essex University, so the direction has to be Down.



MEP-2 configuration

MEP index	2
MD index	9
MD name	janet-essex
MA/MEG index	2
MA/MEG name	janet-essex
MEPID	71
Port	Network
MAC address	00:15:AD:0A:05:CF
Direction	Down
Active	☑
CCI enable	☑
CCM sequence number	☑
Primary VID	701
CCM and LTM priority	7
Lowest priority defect alarm	MacRemErrXconAis
Fault notification alarm time	2500
Fault notification reset time	10000

Figure 3.10: A MEP configuration form

3.4.3 Test Description

The test consisted of the following steps:

1. For each Ethernet service under monitoring, set up MD, MA, and MEPs in the ISG boxes in the testbeds that terminate the service.
2. Launch a CCM session for one of the Ethernet services under monitoring.
3. Check whether the service is working by pinging a remote host at the far end of a service.
4. Check whether an ISG CCM session confirms the result of pinging, i.e. shows that the service is up.
5. Break a service (e.g. by shutting down a port along the service route) and check whether both pinging and CCM session show that the service is down.

3.4.4 Expected Results

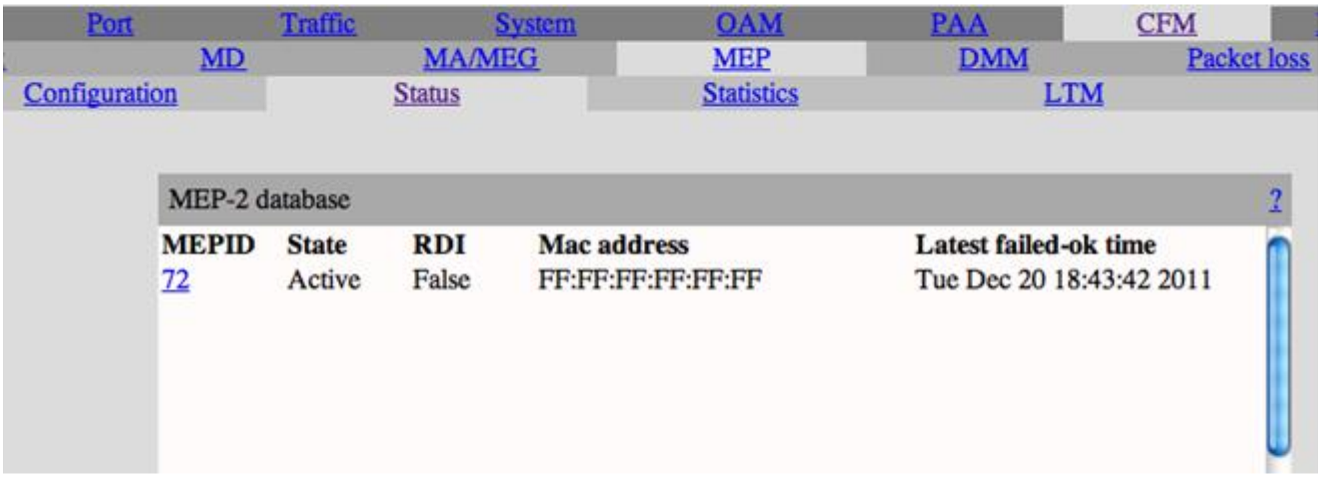
The data from the CCM sessions should correctly reflect the state of the service under monitoring.

3.4.5 Results

The CCM sessions between the ISG boxes correctly detected the state of the services that were established within the multi-NREN testbed.

3.4.5.1 *Monitoring the Up State of a Service*

This section considers the results of monitoring Ethernet services using an example where the actual state of the JANET LH–JANET Essex University service was Up (confirmed by pinging). The web form indicating the status of the remote MEP 72 (which sat in the JANET Essex University testbed) was as shown in Figure 3.11.

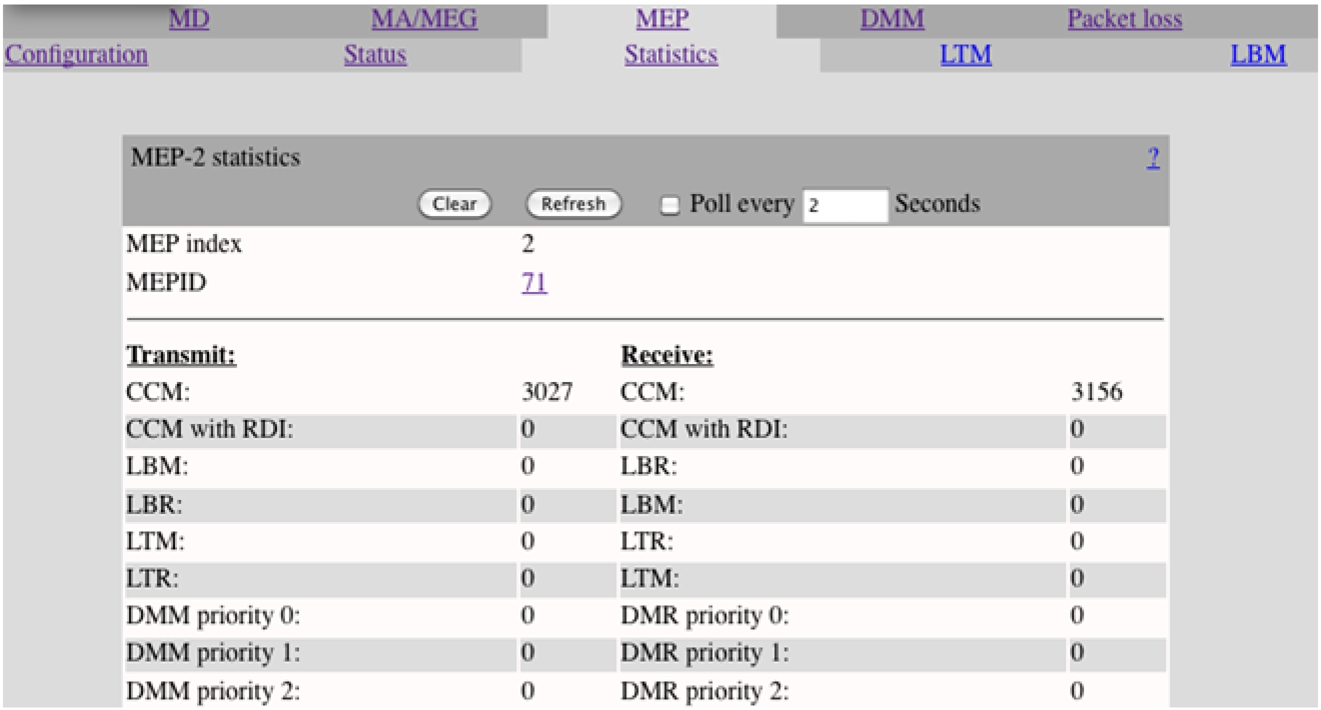


MEP-2 database

MEPID	State	RDI	Mac address	Latest failed-ok time
72	Active	False	FF:FF:FF:FF:FF:FF	Tue Dec 20 18:43:42 2011

Figure 3.11: Active state of remote MEP 72 indicating Up state of janet-essex service

MEP 71 statistics also confirmed that the service was Up (Figure 3.12):



MEP-2 statistics

□ Poll every 2 Seconds

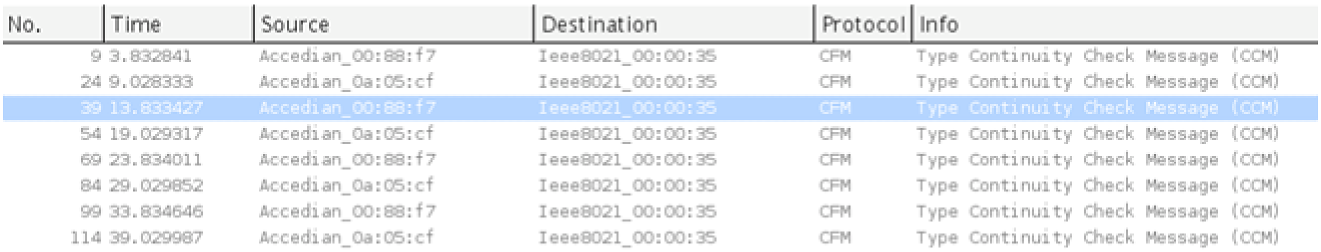
MEP index	2
MEPID	71

Transmit:		Receive:	
CCM:	3027	CCM:	3156
CCM with RDI:	0	CCM with RDI:	0
LBM:	0	LBR:	0
LBR:	0	LBM:	0
LTM:	0	LTR:	0
LTR:	0	LTM:	0
DMM priority 0:	0	DMR priority 0:	0
DMM priority 1:	0	DMR priority 1:	0
DMM priority 2:	0	DMR priority 2:	0

Figure 3.12: MEP 71 statistics when janet-essex service was Up

MEP 71 statistics in this example show that 3027 CCMs were sent towards the remote MEP, none of which had the Remote Defect Indicator (RDI) signal activated. The latter means that MEP 71 regularly received CCMs from the remote MEP and found all of them correct – otherwise the RDI would have been activated.

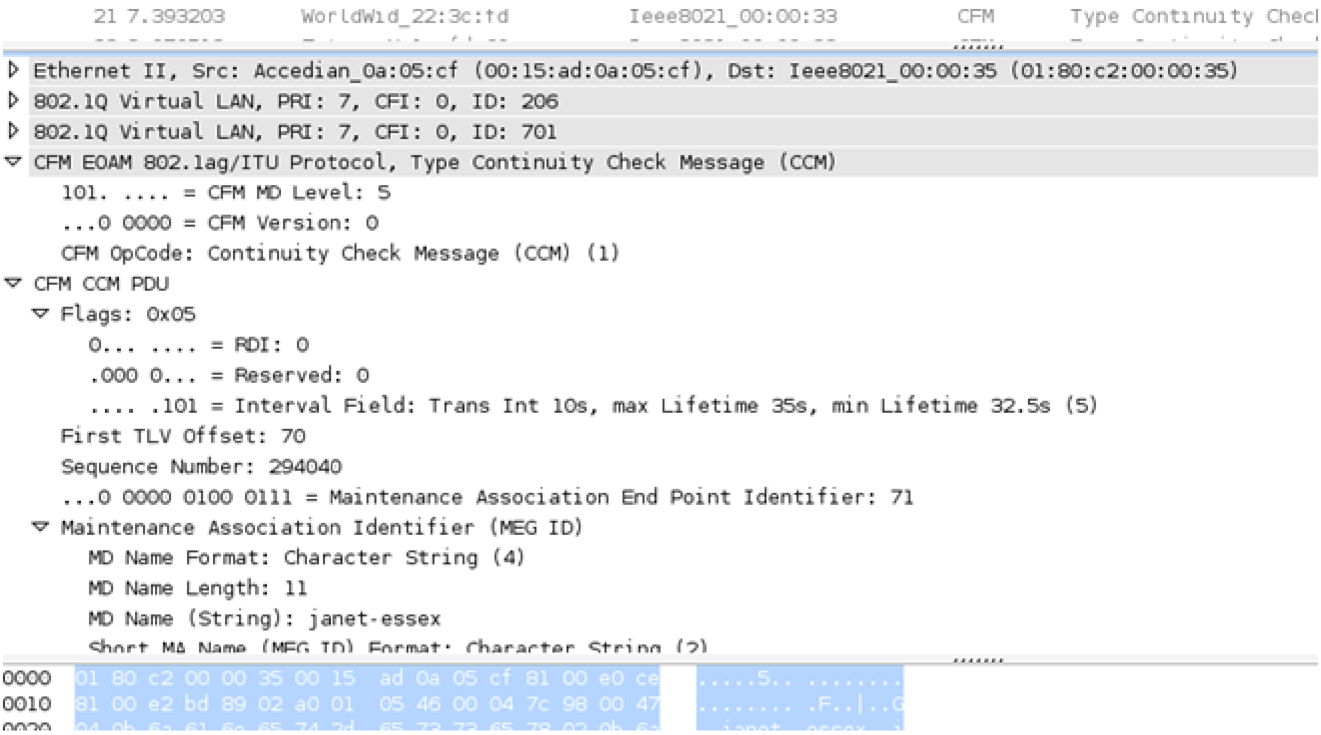
Finally, CCM frames going between the JANET LH and JANET Essex University sites also showed that the service was in an Up state. The screenshot in Figure 3.13 consists of the sequence of CCM frames going from MEP 71 (LH, the source address ending c7) and from MEP 72 (Essex University, the source address ending f7).



```
No.   Time        Source               Destination          Protocol  Info
    9 3.832841    Accedian_00:88:f7    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   24 9.028333    Accedian_0a:05:cf    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   39 13.833427   Accedian_00:88:f7    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   54 19.029317   Accedian_0a:05:cf    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   69 23.834011   Accedian_00:88:f7    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   84 29.029852   Accedian_0a:05:cf    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
   99 33.834646   Accedian_00:88:f7    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
  114 39.029987   Accedian_0a:05:cf    Ieee8021_00:00:35    CFM       Type Continuity Check Message (CCM)
```

Figure 3.13: CCM frames from the janet-essex MA session

As shown in Figure 3.14, each frame has RDI=0, which means that a MEP sees CCMs coming from a remote MEP regularly (the interval was 10 ms in that session) and finds them correct.

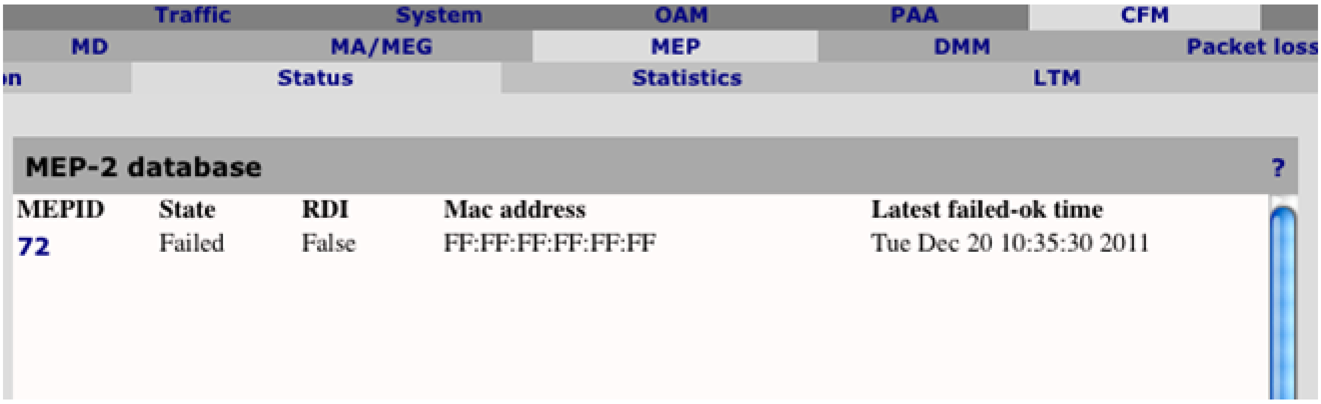


```
      21 7.393203        WorldWid_22:3c:fd          Ieee8021_00:00:33          CFM       Type Continuity Checl
▷ Ethernet II, Src: Accedian_0a:05:cf (00:15:ad:0a:05:cf), Dst: Ieee8021_00:00:35 (01:80:c2:00:00:35)
▷ 802.1Q Virtual LAN, PRI: 7, CFI: 0, ID: 206
▷ 802.1Q Virtual LAN, PRI: 7, CFI: 0, ID: 701
▽ CFM EOAM 802.1ag/ITU Protocol, Type Continuity Check Message (CCM)
    101. .... = CFM MD Level: 5
    ...0 0000 = CFM Version: 0
    CFM OpCode: Continuity Check Message (CCM) (1)
▽ CFM CCM PDU
  ▽ Flags: 0x05
      0... .... = RDI: 0
      .000 0... = Reserved: 0
      .... .101 = Interval Field: Trans Int 10s, max Lifetime 35s, min Lifetime 32.5s (5)
    First TLV Offset: 70
    Sequence Number: 294040
    ...0 0000 0100 0111 = Maintenance Association End Point Identifier: 71
  ▽ Maintenance Association Identifier (MEG ID)
      MD Name Format: Character String (4)
      MD Name Length: 11
      MD Name (String): janet-essex
      Short MA Name (MEG ID) Format: Character String (2)
0000  01 80 c2 00 00 35 00 15  ad 0a 05 cf 81 00 e0 ce   .....5.. ........
0010  81 00 e2 bd 89 02 a0 01  05 46 00 04 7c 98 00 47   ........ .F..|..G
0020  04 0b 6a 61 6e 65 74 2d  65 73 73 65 78 02 0b 6a   ..janet- essex..j
```

Figure 3.14: CCM frame from the janet-essex MA

3.4.5.2 *Monitoring of Down State of the Service*

When a service was in a Down state, the ISG boxes correctly indicated this fact in web forms relating to the service state. For example, the following screenshot (Figure 3.15) shows that MEP 71 in the JANET LH testbed saw the remote MEP 72 as Failed because no CCMs were received from it.

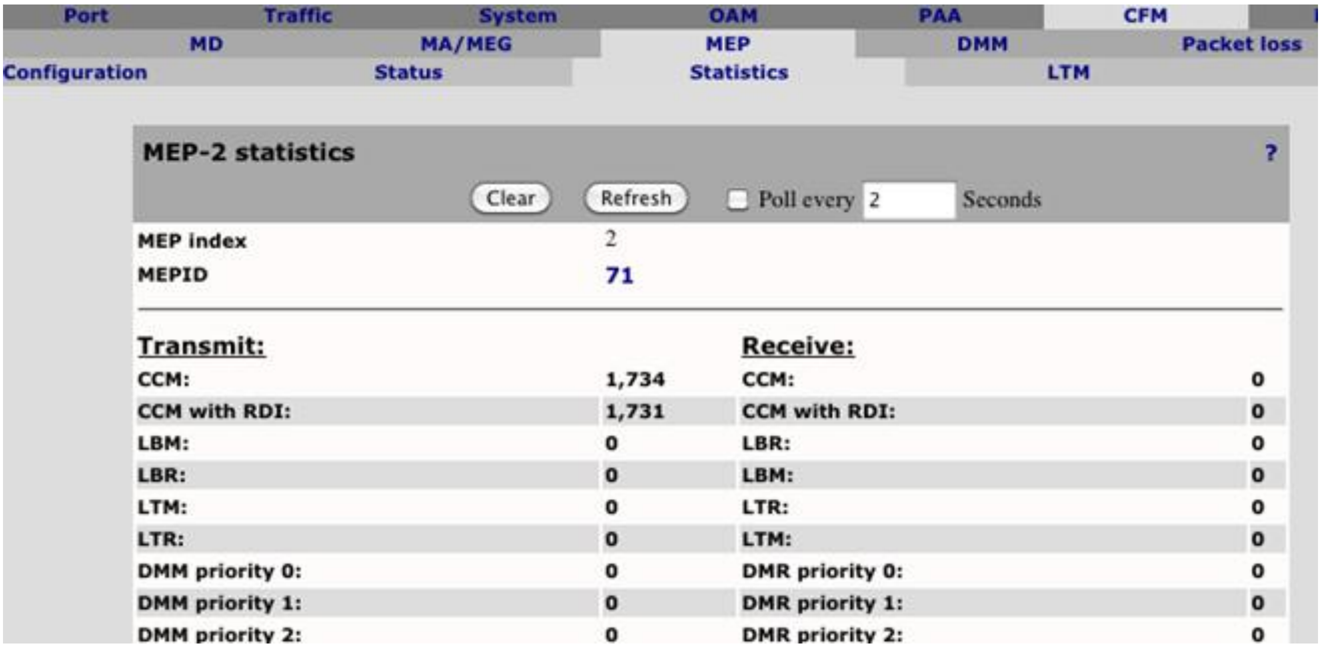


MEP-2 database

MEPID	State	RDI	Mac address	Latest failed-ok time
72	Failed	False	FF:FF:FF:FF:FF:FF	Tue Dec 20 10:35:30 2011

Figure 3.15: Failed state of remote MEP 72 indicating Down status of janet-essex service

The MEP 71 statistics also confirmed that fact, as MEP 71 was sending CCMs with RDI=1 and receiving no CCMs from MEP 72 (Figure 3.16).



MEP-2 statistics

MEP index 2

MEPID 71

Transmit:		Receive:	
CCM:	1,734	CCM:	0
CCM with RDI:	1,731	CCM with RDI:	0
LBM:	0	LBR:	0
LBR:	0	LBM:	0
LTM:	0	LTR:	0
LTR:	0	LTM:	0
DMM priority 0:	0	DMR priority 0:	0
DMM priority 1:	0	DMR priority 1:	0
DMM priority 2:	0	DMR priority 2:	0

Figure 3.16: MEP 71 statistics when janet-essex service was Down

3.4.6 Test Conclusions

The conclusions drawn from Test 1 were as follows:

- ISG agent boxes installed in the participants' testbeds were correctly able to detect the Up and Down state of point-to-point Ethernet services running within the multi-NREN OAM testbed.
- The information about a service state that is available in ISG24 web forms is not convenient for the routine daily operations of network administrators on monitoring service state as it requires from the administrator a lot of manual effort and some knowledge of the Y.1731 CCM protocol. In particular, there is no clear indication of a service status (Up/Down) in the forms.

3.5 Test 2: Monitoring of Ethernet Services Rate

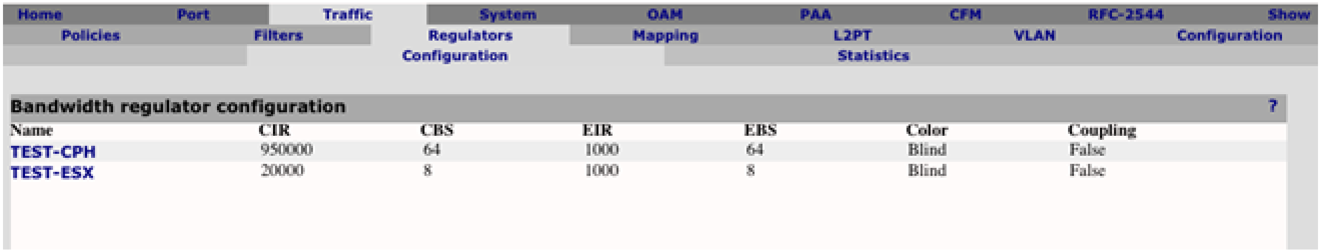
Monitoring the actual rate of traffic going over an Ethernet connection is useful as the results could be helpful in many situations, e.g. network planning, troubleshooting, etc.

3.5.1 Test Setup

The Test 2 setup was the same as for Test 1 (see Section 3.4.1). The only difference was that open source software such as Iperf [Iperf] and RUDE [RUDE] was used to generate traffic between sites.

3.5.2 Configuration

To measure the traffic rate, all ISG boxes used traffic policers, called regulators in ISG terminology. For example, the JANET LH ISG box had two regulators configured: TEST-CPH for the janet-copenhagen service, and TEST-ESX for the janet-essex service, as shown in Figure 3.17 below.



Bandwidth regulator configuration

Name	CIR	CBS	EIR	EBS	Color	Coupling
TEST-CPH	950000	64	1000	64	Blind	False
TEST-ESX	20000	8	1000	8	Blind	False

Figure 3.17: Traffic regulators

At the start of the trial, all of the traffic regulators on the ISG boxes were configured with Committed Information Rate (CIR) = 20000 Kbit/s. Later, some of them (such as TEST-CPH) were reconfigured to allow a higher traffic rate, to test the performance of the testbed with a load close to the links' capacity of 1 Gbit/s.

3.5.3 Test Description

The test consisted of the following steps:

1. Using Iperf or RUDE software, generate a traffic stream with a particular rate Rg Kbit/s for a particular Ethernet service between the trial sites.
2. Measure the actual traffic rate Ra of the service under monitoring using the rate counter of one of the ISG boxes terminating the service.
3. Compare Ra and Rg and evaluate the precision of the rate measurements.

3.5.4 Expected Results

The measured rate Ra should be close to the value Rg of traffic injected into the Ethernet connection. Different ways of calculating traffic rate should be taken into account, as RUDE and Iperf use User Datagram Protocol (UDP) or Transmission Control Protocol (TCP) payload size while ISG uses Ethernet frame size.

3.5.5 Results

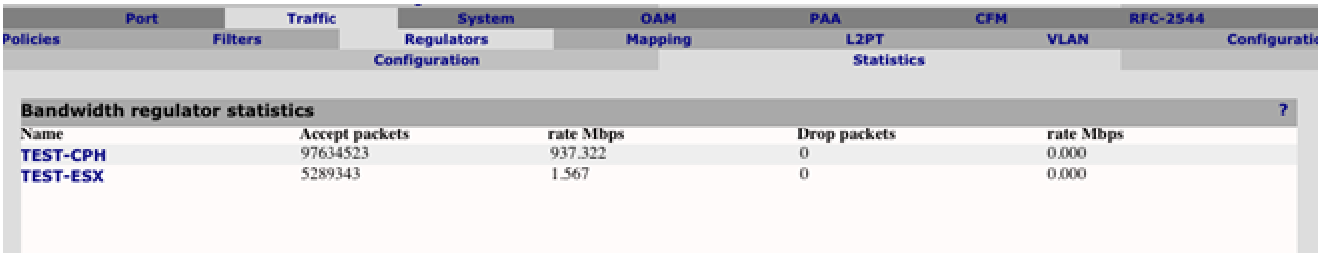
The test was carried out on the janet-essex, janet-copenhagen and copenhagen-poznan services.

All the results obtained during these tests showed quite good level of precision from the ISG boxes in measuring traffic rate. For example, during the test conducted between JANET LH and NORDUnet (the janet-copenhagen service) on 3 November 2011, the RUDE software had the following configuration:

- Packet size = 1400 bytes.
- Packet frequency = 80000 pps.

These traffic parameters give Rg = 80000 x 1400 x 8 = 896000 Kbit/s.

The observed traffic rate Ra in the JANET LH ISG was 937322 Kbit/s, as shown in Figure 3.18 below.



Bandwidth regulator statistics

Name	Accept packets	rate Mbps	Drop packets	rate Mbps
TEST-CPH	97634523	937.322	0	0.000
TEST-ESX	5289343	1.567	0	0.000

Figure 3.18: Results of the traffic rate measurement

At first glance, the difference between the generated and measured traffic rates seems too large, suggesting that something went wrong with either the traffic generator or the ISG box counter. However, the figure of Rg needs some amendment as it shows the traffic rate for the UDP payload, while the ISG box uses the full Ethernet frame size, including the preamble and inter-packet gap bytes. Thus, the bytes in Table 3.3 below need to be added to the size of the UDP payload to find the true Rg value for the generated Ethernet frames:

Frame Element	No. of Bytes
UDP Header	8
IP Header	20
VKAN tag	4
EtherType	2
MAC addresses	12
Preamble	8
Interpacket Gap	12
Total:	**66**

Table 3.3: Additional bytes for frame elements

The line rate of Rg was therefore: (1400 + 66) x 8 x 80000 = 938240 Kbit/s. With this correction, the difference between Rg and Ra was 818 Kbit/s or 0.08%, which can be viewed as good level of measurement precision.

3.5.6 Test Conclusions

The conclusions drawn from Test 2 were as follows:

- The measurements of Ethernet services traffic rate showed a good level of precision from the ISG counters – in the region of 0.1%.
- The traffic rate measurement results identified a source of confusion when the configured rate of streams from software traffic generators like RUDE or Iperf is compared against hardware rate counters, arising from the different rate-measurement criteria. Therefore, a network operator must be aware of this inconsistency and be able to translate between the two rate values.

3.6 Test 3: Monitoring of Frame Delay, Jitter and Loss

Frame delays, delay variations and loss are important performance parameters for Ethernet services [MEF 10.2.1, ITU-T Y.1563]. Some types of applications such as Voice over IP (VoIP) or videoconferencing are time sensitive and their performance usually deteriorates when frame delays, delay variations and loss exceed certain limits. Monitoring these parameters gives a network operator and their customers a quantitative background for assessing whether a service complies with its Service Level Description (SLD).

3.6.1 Test Setup

The Test 3 setup was the same as for Test 2 (see Section 3.5.1).

3.6.2 Configuration

ISG boxes support two different protocols for delay/loss measurement: Y.1731 and the proprietary Performance Assurance Agent (PAA). Y.1731 performance monitoring functions are called Connectivity Fault Management Delay Measurement Messages (CFM DMM) in ISG menus and result forms, which is the acronym used in this section for describing this particular ISG function.

Both protocols measure the following parameters:

- One-way Average Delay (OAD).
- One-way Average Delay Variation (OADV) (i.e. jitter).
- Two-way Average Delay (TAD).
- Two-way Average Delay Variation (TADV).
- Packet Loss (PL).

The Y.1731 protocol was used as the main tool for monitoring delay/loss and was activated in all the partners' ISG boxes. PAA was used as a secondary tool to check and verify the Y.1731 results of the PSNC–CESNET, JANET LH–JANET Essex University and JANET LH–NORDUnet services.

Configuration of both CFM DMM and PAA agents was straightforward and required only the specification of a service to monitor.

For measurement of one-way delay, synchronisation of an ISG box with an external Network Time Protocol (NTP) server was required. All ISG boxes were initially synchronised with the ntp.ubuntu.com server; during the trial period some ISG boxes used different NTP servers.

3.6.3 Test Description

The test consisted of the following steps:

1. Activate the CFM DMM function on all ISG boxes.
2. Activate the PAA function on selected ISG boxes.
3. Create a traffic load for a service under monitoring.
4. Periodically (e.g. once a day) write down the CFM DMM and PAA results.
5. Evaluate CFM DMM results with:
    a. Delay results obtained by other tools (e.g. by Iperf measurements).
    b. PAA results for services for which PAA was activated.

3.6.4 Expected Results

Delay and delay variation values should be within reasonable limits known from the NRENs' experience on network monitoring and correlate to results obtained by other tools, such as Network Management Systems and traffic monitors

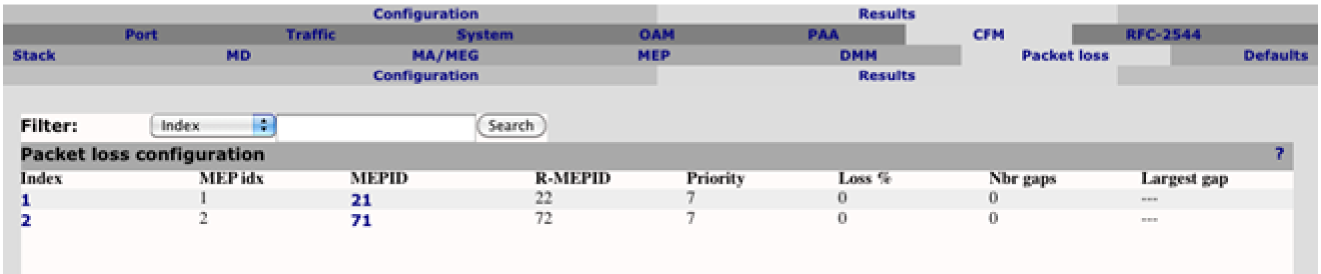
The quality of the links used for site interconnection was good, a priori Packet Loss should be close to 0% (all the links were built over production networks and services such as GÉANT Plus, JANET Lightpath, etc.).

3.6.5 Results

3.6.5.1 *Packet Loss Monitoring*

All results of Packet Loss monitoring were correct, as they were either 0% when the corresponding service was up, or 100% when the service was deliberately broken.

Typical results of monitoring Packet Loss for JANET-LH–JANET-Essex University and JANET LH–NORDUnet services are shown in Figure 3.19:



Packet loss configuration

Index	MEP idx	MEPID	R-MEPID	Priority	Loss %	Nbr gaps	Largest gap
1	1	21	22	7	0	0	---
2	2	71	72	7	0	0	---

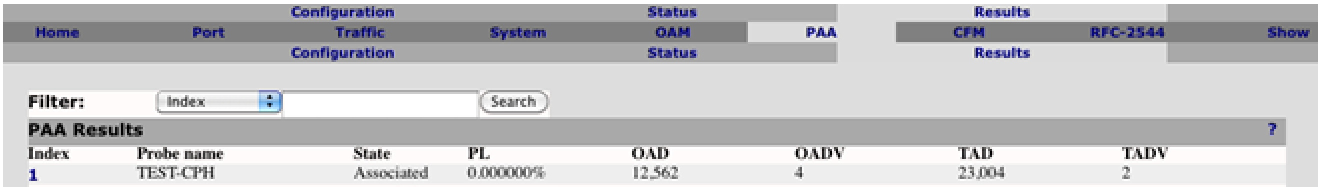
Figure 3.19: Packet Loss results

3.6.5.2 *Delay Monitoring*

PAA

PAA results were stable, satisfactory and corresponded to known values for NREN domains.

A typical PAA result for JANET LH–NORDUnet service monitoring is shown in Figure 3.20:



PAA Results

Index	Probe name	State	PL	OAD	OADV	TAD	TADV
1	TEST-CPH	Associated	0.000000%	12,562	4	23,004	2

Figure 3.20: PAA results for JANET LH – NORDUnet service

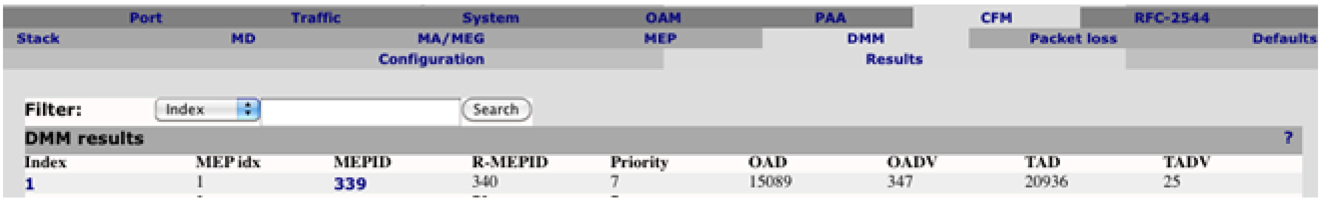
The screenshot in Figure 3.20 shows OAD from NORDUnet to JANET LH of about 12 ms, which corresponds to the known NORDUnet production network delay of 10 ms. The two-way delay of 23 ms also corresponds to a sum of OAD values for NORDUnet–JANET LH (12 ms) and JANET LH–NORDUnet (11 ms, not shown in Figure 3.20) directions.

CFM DMM

CFM DMM results were not so straightforward.

For PSNC–CESNET, PSNC–NORDUnet and CESNET–SURFnet, CFM DMM results were stable, satisfactory and corresponded to known values for NREN domains.

PSNC–NORDUnet CFM DMM results were also checked by the corresponding PAA results for this service. A screenshot of these results is shown in Figure 3.21.



DMM results

Index	MEP idx	MEPID	R-MEPID	Priority	OAD	OADV	TAD	TADV
1	1	339	340	7	15089	347	20936	25

Figure 3.21: CFM DMM results for the PSNC–CESNET service

However, CFM DMM measurement for the JANET LH–JANET Essex University and JANET LH–NORDUnet services showed some peculiar behaviour:

- No OAD values were measured for either service; the CFM DMM result forms showed dashes in their OAD field instead of figures.
- The TAD graph of the JANET LH–JANET Essex University service had a regular saw-tooth shape with about 1 second peaks, with no reasons for such regular changes.

The strange behaviour of delay measurements for the JANET LH–NORDUnet and JANET LH–JANET Essex University services was investigated separately and described in Test 6 (Section 3.9 on page 83) and Test 7 (Section 3.10 on page 87) respectively.

3.6.6 Test Conclusions

The conclusions drawn from Test 3 were as follows:

- One-way delay measurements required tight synchronisation of ISG agents; synchronisation with public NTP servers turned out to be enough for this purpose.
- The proprietary PAA protocol showed better robustness than the standard Y.1731. The reasons for this required a special investigation, which is described in Test 6 (Section 3.9 on page 83) and Test 7 (Section 3.10 on page 87).
- All the PAA results were robust and satisfactory.
- The CFM DMM protocol showed robust and satisfactory results for three out of five services; it could not measure OAD on the JANET LH–JANET Essex University and JANET LH–NORDUnet connections and measured incorrectly TAD on the JANET LH–JANET Essex University connection.

3.7 Test 4: Investigation of CyPortal Functionality

All monitoring results are difficult to use in day-to-day network operations without a software system that stores and visualise those results in a convenient form. CyPortal is a cloud-based software system which was developed for storing and visualising Y.1731 data. Such a service is still not very widespread. For example, the main GÉANT NREN monitoring system, perfSONAR, currently lacks such capabilities and a new GN3 Year 4 sub-activity Services Assurance and Monitoring has been established to bridge this gap. In addition to being one of the few systems on the market, CyPortal was selected as the company (Cyan, Inc.) was keen to test their new service in an academic environment and suggested a free trial.

3.7.1 Test Setup

The test setup corresponded to the diagram shown in Figure 3.3 and described in Section 3.3.1.

3.7.2 Configuration

All the ISG boxes were configured to ftp fresh monitoring data to the controller located in the JANET LH testbed. The minimum value of 15 minutes was chosen as the data-sending interval.

3.7.3 Test Description

The test consisted of the following steps:

1. Give all the trial participants access to the CyPortal instance set up by Cyan specialists for the purpose of the trial.
2. Add the five trial services monitored by ISG agents to CyPortal.
3. Configure CyPortal to use thresholds for some monitored service parameters to indicate the health of the service.
4. Store CyPortal-monitored data over the trial period and visualise it on trial-participant request.
5. Produce CyPortal reports and screenshots for major trial events such as artificial line breaks or creating extra load for a service.

3.7.4 Expected Results

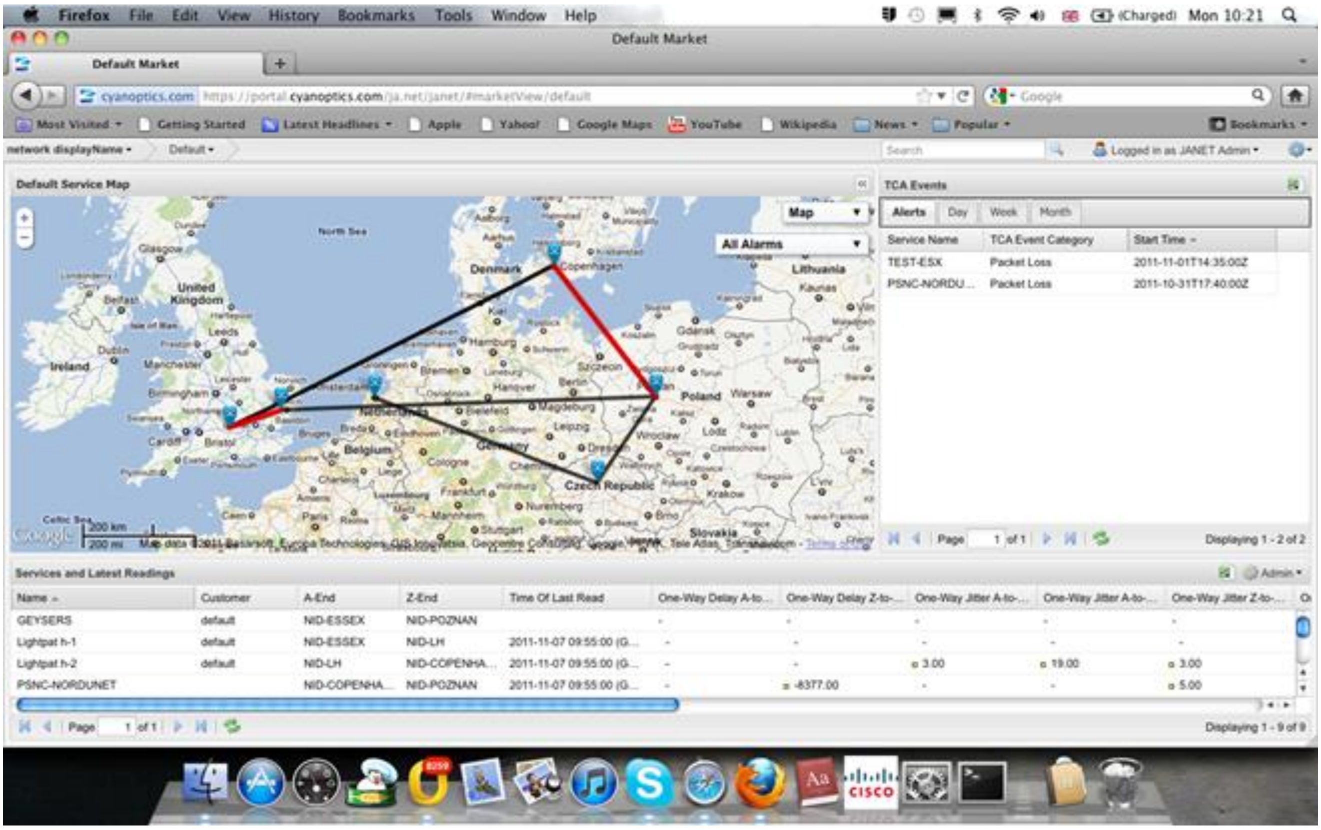
CyPortal was expected to show the following functionality:

- Store all monitoring data received from ISG agents for each service over the trial period of six months.
- Visualise the overall health of the services on a topology map. An overall “green” health status means that all parameters of a service are within threshold limits; a “red” health status means that at least one parameter is outside the threshold limits.
- Visualise data for a selected service with a different level of granularity for a period specified by the user.
- Refresh monitored data in real time according to the interval of data push from ISG agents.

3.7.5 Results

3.7.5.1 *Visualisation of Overall Health of Services*

CyPortal was able to visualise the overall health of services in a convenient form on a Google map (Figure 3.22).


Firefox File Edit View History Bookmarks Tools Window Help
Default Market
cyanoptics.com https://portal.cyanoptics.com/ja.net/janet/#marketView/default
Most Visited Getting Started Latest Headlines Apple Yahoo! Google Maps YouTube Wikipedia News Popular Bookmarks
network displayName Default
Logged in as JANET Admin
Default Service Map
Map
All Alarms
North Sea
United Kingdom
Ireland
Denmark
Netherlands
Belgium
Germany
Poland
Czech Republic
Lithuania
Slovakia
200 km
200 mi
TCA Events
Alerts Day Week Month
Page 1 of 1
Displaying 1 - 2 of 2
Services and Latest Readings
Admin
Page 1 of 1
Displaying 1 - 9 of 9

Service Name	TCA Event Category	Start Time
TEST-ESX	Packet Loss	2011-11-01T14:35:00Z
PSNC-NORDU...	Packet Loss	2011-10-31T17:40:00Z

Name	Customer	A-End	Z-End	Time Of Last Read	One-Way Delay A-to-...	One-Way Delay Z-to-...	One-Way Jitter A-to-...	One-Way Jitter A-to-...	One-Way Jitter Z-to-...
GEYSERS	default	NID-ESSEX	NID-POZNAN		-	-	-	-	-
Lightpat h-1	default	NID-ESSEX	NID-LH	2011-11-07 09:55:00 (G...	-	-	-	-	-
Lightpat h-2	default	NID-LH	NID-COPENHA...	2011-11-07 09:55:00 (G...	-	-	3.00	19.00	3.00
PSNC-NORDUNET		NID-COPENHA...	NID-POZNAN	2011-11-07 09:55:00 (G...	-	-8377.00	-	-	5.00

Figure 3.22: CyPortal services health map

Two services in the Figure 3.22 example of the CyPortal map are shown in red (JANET LH–JANET Essex University and PSNC–NORDUnet) which means that some of the parameters of these services exceeded threshold limits. The pane in the top-right corner of the web page shows which parameters did so, namely:

- Packet Loss of the TEST-ESX service (an internal name for the JANET LH–JANET Essex University service monitored by the PAA protocol).
- Packet Loss of the PSNC–NORDUnet service.

The bottom pane in Figure 3.22 shows the latest monitored data of all the services added to CyPortal.

The rest of the services are shown in black (this colour was chosen to indicate a “green” service status; the colour is tuneable), which means that all parameters of these services were within threshold limits (or that limits have not been set up for these services).

The mechanism for setting up a parameter threshold by a CyPortal administrator is described in Section 3.7.5.2 below. The only threshold that could not be changed on the fly (as it was set up by CyPortal developers) was the percentage of packet loss, which indicates that the service is down; it was 100% during the trial (but it can be changed by CyPortal developers on request).

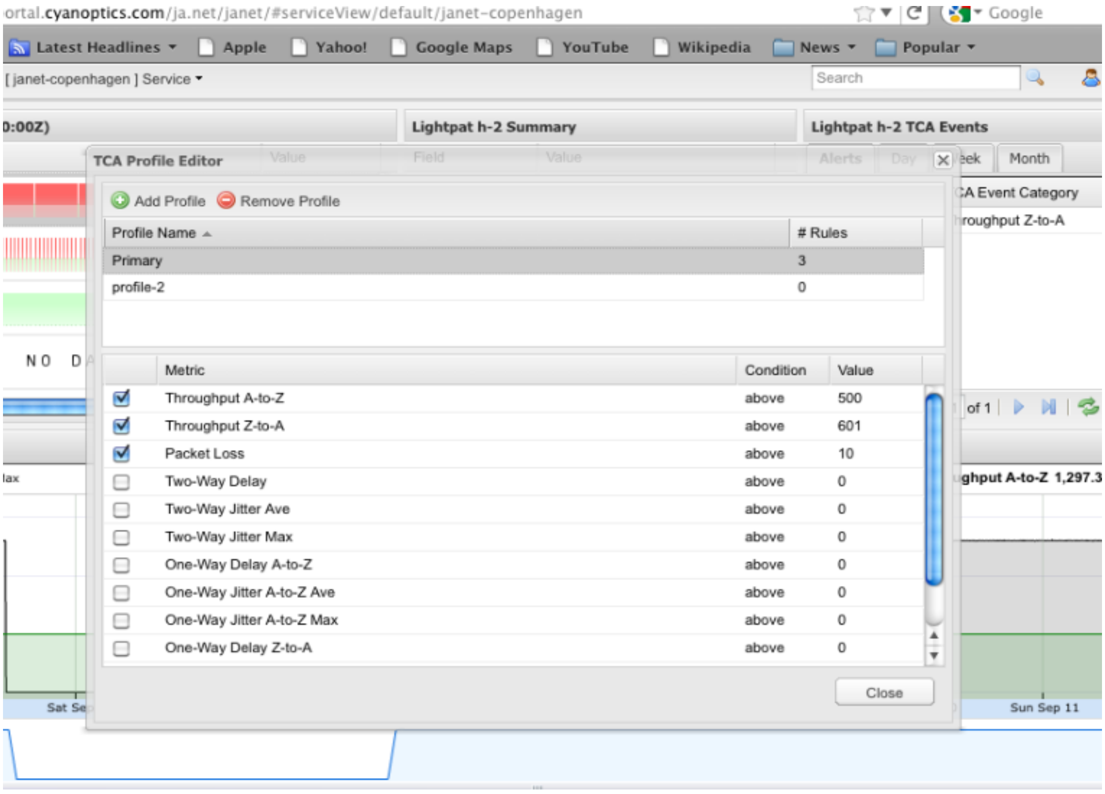
The CyPortal feature to paint services-violators on the map in red was not available at the beginning of the trial. CyPortal developers added the feature at the request of the trial participants.

Generally, the CyPortal map features turned out to be flexible enough to monitor services status and identify problematic services for more detailed investigation.

The only inconvenience noticed was the long interval for data refreshing (15 minutes). However, this was limited by the ISG agents, not by CyPortal functionality. The possibility to decrease this interval was discussed with CyPortal specialists. The most convenient way seems to be for the controller to poll Simple Network Management Protocol (SNMP) agents of ISG and network equipment of other vendors. This ability would not only use a shorter interval of data refreshing but would collect and visualise Y.1731 data from production network equipment without (or in parallel with) polling extra Y.1731 agents like ISG boxes.

3.7.5.2 *Setting Alarm Thresholds on Service Parameters*

CyPortal allows users with administrator rights to set up thresholds for any Y.1731 parameter monitored by ISG agents: throughput, packet loss, one-way and two-way delay and jitter; the thresholds can be different for A-Z and Z-A service directions. CyPortal supports several profiles for each service; one of them can be activated. The screenshot in Figure 3.23 below shows an example pane of the Traffic Conformance Agreement (TCA) Profile Editor that should be used for setting thresholds.



Profile Name	# Rules
Primary	3
profile-2	0

	Metric	Condition	Value
☑	Throughput A-to-Z	above	500
☑	Throughput Z-to-A	above	601
☑	Packet Loss	above	10
□	Two-Way Delay	above	0
□	Two-Way Jitter Ave	above	0
□	Two-Way Jitter Max	above	0
□	One-Way Delay A-to-Z	above	0
□	One-Way Jitter A-to-Z Ave	above	0
□	One-Way Jitter A-to-Z Max	above	0
□	One-Way Delay Z-to-A	above	0

Figure 3.23: Setting up service parameter thresholds

The example in Figure 3.23 shows that three thresholds were set up for the janet-copenhagen service:

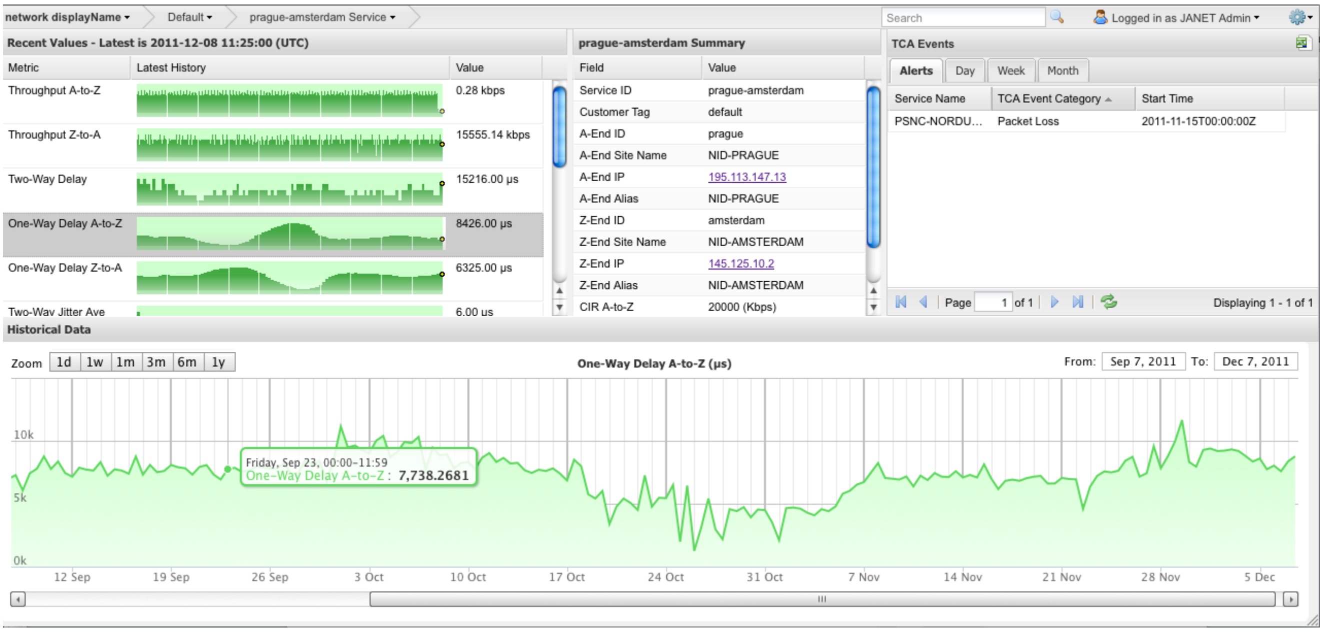
- Throughput A-to-Z: 500 (Kbit/s).
- Throughput Z-to-A: 601 (Kbit/s).
- Packet Loss: 10%.

If these thresholds are exceeded, CyPortal generates an alarm event and paints the service on the map in red. There is a possibility of examining alarm events in a drop-down window. Posting alarm emails to an administrator would be a useful feature, but it was not implemented at the time of the trial. (It is, however, on the roadmap for CyPortal.)

The mechanism of CyPortal thresholds and alarms worked correctly in all cases over the trial.

3.7.5.3 *Detailed Visualisation of Service Parameters*

CyPortal provides the user with detailed information about the parameters of a service when that service is selected (by clicking on it on the map or on the services list in the bottom pane of the map view). A typical detailed view of a service is shown in the screenshot of the prague-amsterdam service (the internal name for the CESNET–SURFnet service) shown in Figure 3.24.



Recent Values - Latest is 2011-12-08 11:25:00 (UTC)

Metric	Latest History	Value
Throughput A-to-Z		0.28 kbps
Throughput Z-to-A		15555.14 kbps
Two-Way Delay		15216.00 µs
One-Way Delay A-to-Z		8426.00 µs
One-Way Delay Z-to-A		6325.00 µs
Two-Way Jitter Ave		6.00 us

prague-amsterdam Summary

Field	Value
Service ID	prague-amsterdam
Customer Tag	default
A-End ID	prague
A-End Site Name	NID-PRAGUE
A-End IP	195.113.147.13
A-End Alias	NID-PRAGUE
Z-End ID	amsterdam
Z-End Site Name	NID-AMSTERDAM
Z-End IP	145.125.10.2
Z-End Alias	NID-AMSTERDAM
CIR A-to-Z	20000 (Kbps)

TCA Events

Service Name	TCA Event Category	Start Time
PSNC-NORDU…	Packet Loss	2011-11-15T00:00:00Z

Historical Data

Figure 3.24: Detailed view of prague-amsterdam service

The pane in the top-left corner shows the latest values of all the service parameters, with compact graphs showing how the parameters changed during the previous day. The bottom pane shows a graph of the selected parameter. The user can choose an interval of monitoring: 1 day, 1 week. 1 month, 3 months or 1 year. In the example in Figure 3.24, the interval of 3 months from 7 September 2011 to 7 December 2011 was chosen.

It is also possible to drag a graph left or right to see a different time interval of monitoring. By positioning the mouse cursor over a particular time point, the user can see the precise time stamp and value of a parameter.

CyPortal correctly visualised such events as line breaks. The trial participants artificially created such breaks by switching off a port on a network box along a service route. Figure 3.25 below illustrates how CyPortal visualised a line break between JANET LH and NORDUnet, which happened on 18 October 2011 and lasted until 31 October 2011. The dotted red line shows the 10% threshold for Packet Loss set up for this service.

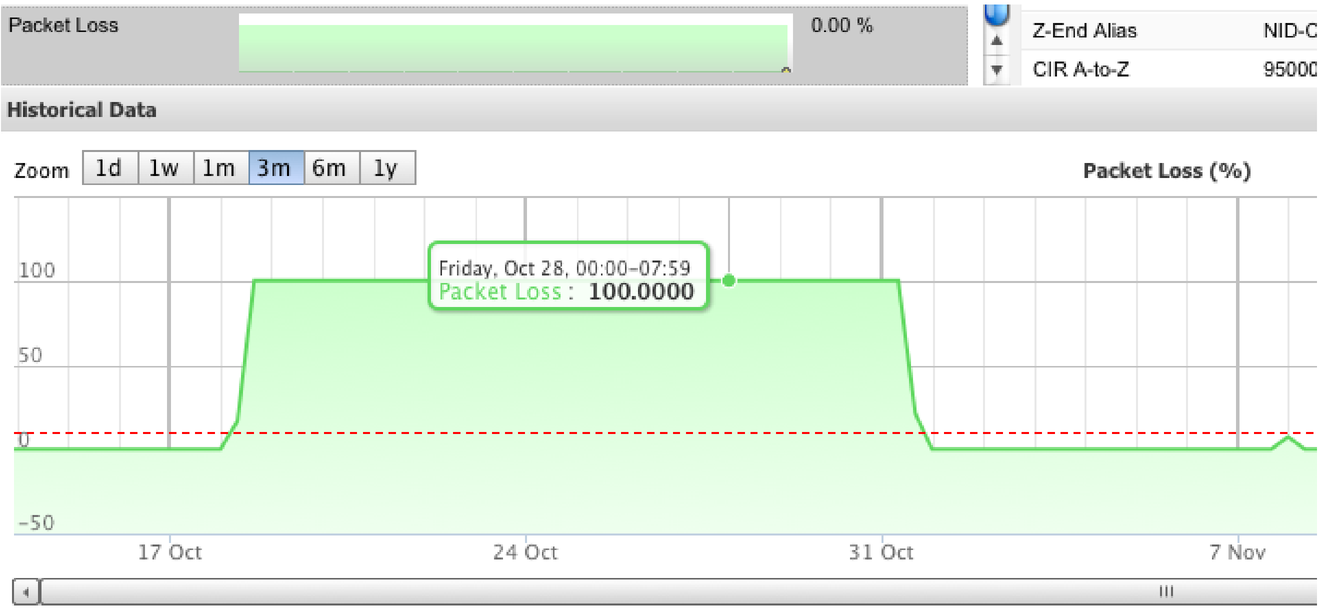


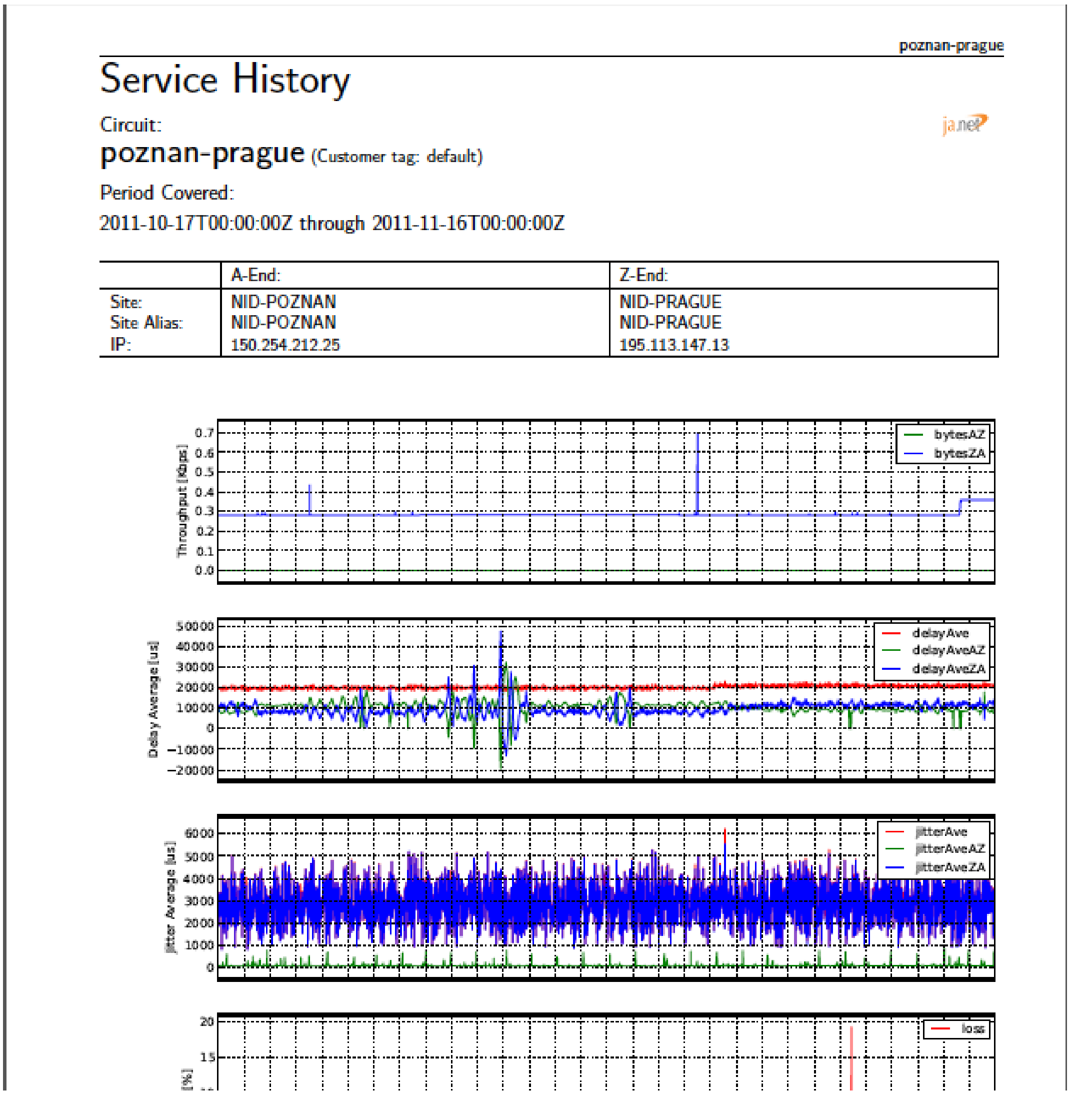
Figure 3.25: Visualising of line break

3.7.5.4 *CyPortal Reports*

CyPortal can generate different kinds of reports in the form of pdf documents on request. These include:

- Summary report, which includes data about the aggregate bandwidth usage of all services. The time period is specified by the user.
- Full monthly report which, along with aggregated bandwidth usage, covers:
  - Top 10 Worst Offenders by Number of Violations and Top 10 Worst Offenders by Time Outage Percentage.
  - Service statistics (per service) with maximum, minimum and average values of availability, loss, bandwidth usage, delay and jitter.
- Service report (daily, weekly, monthly or for a user-selected period).

An example of a service report is shown in Figure 3.26 below.



poznan-prague

Service History

Circuit:

poznan-prague (Customer tag: default)

Period Covered:

2011-10-17T00:00:00Z through 2011-11-16T00:00:00Z

	A-End:	Z-End:
Site: Site Alias: IP:	NID-POZNAN NID-POZNAN 150.254.212.25	NID-PRAGUE NID-PRAGUE 195.113.147.13

Figure 3.26: Example of CyPortal daily report

3.7.6 Test Conclusions

The conclusions drawn from Test 4 were as follows:

- CyPortal proved to be capable of storing and visualising Ethernet services parameters according to Y.1563 and Y.1731 recommendations in quite a flexible way.
- The CyPortal map function effectively visualises the geographic topology of services and their status with regard to violation of traffic parameter thresholds.
- Traffic parameter thresholds can be set up and changed by a user who has administrative rights.
- Monitoring data can be refreshed at a specific interval equal to or greater than 15 minutes.
- The version of CyPortal used in the trial had some limitations:

- ○ It worked only with ISG probe boxes from Overture (or Accedian Metro Ethernet as an Overture original equipment manufacturer (OEM) box). This limitation is going to be lifted according to Cyan, Inc.’s roadmap for CyPortal, as it assumes support of SNMP pools of equipment from a wider range of vendors.
- ○ It supported only a flat model of Y.1731 monitoring, i.e. it monitored a single segment of service without the possibility of monitoring several nested segments in a hierarchical manner as described by the Y.1731 and IEEE 802.1ag standards. It is a good solution for end users but a partial one for network providers and operators as it doesn’t give a visibility of multi-domain problems. It is unclear whether hierarchical monitoring is on CyPortal’s roadmap.
- ○ The feasibility of integrating CyPortal with the main GÉANT NREN monitoring system, perfSONAR, will be investigated in a new GN3 Year 4 joint sub-activity Service Assurance and Monitoring, which has been established by JRA1 Task 1 and JRA2 Task 3 participants,

3.8 Test 5: Two-Level Monitoring of Ethernet Services State

3.8.1 Test Setup

The equipment used in the testbeds of most of the trial participants supported Y.1731 functionality on their own (i.e. independent of the ISG agent boxes); only the NORDUnet and CESNET testbeds did not have such equipment. In principle, therefore, it was possible to establish separate Y.1731 CCM sessions between testbed network boxes along with sessions between ISG boxes, and to investigate two-level hierarchical monitoring of Ethernet services over the trial multi-domain testbed. In practice, this was done only for the JANET LH–JANET Essex University service because the other trial participants lacked time and resources.

A diagram of the combined JANET LH and JANET Essex University testbeds is shown in Figure 3.27 below.

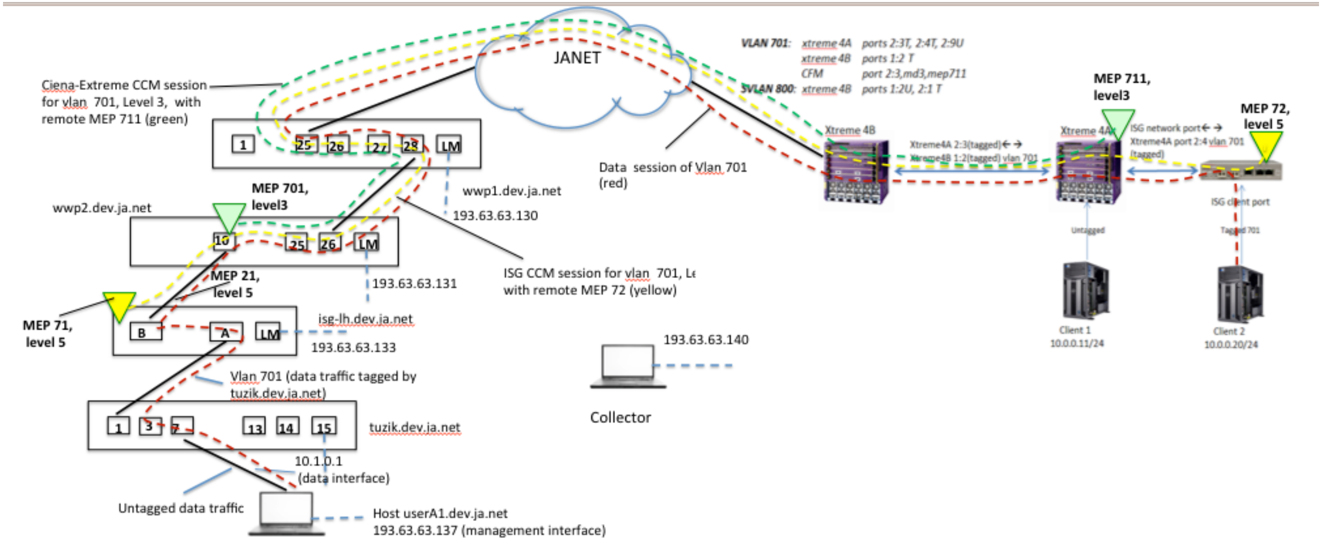


Figure 3.27: Two-level hierarchical monitoring of JANET LH–JANET Essex University service

The Level 5 CCM session janet-essex between the ISG boxes represented a monitoring session between customer end points, while the Level 3 session e-oam-lh-essex represented a monitoring session between service provider end points.

3.8.2 Configuration

The configuration of the Ciena wwp2.dev.ja.net switch to support a CFM CCM session with a remote Extreme switch was as follows:

```
cfm service create vs vs_206 name e-oam-lh-essex next-mepid 702
cfm service enable service e-oam-lh-essex
cfm mep create service e-oam-lh-essex port 10 vlan 701 type up mepid 701
```

This configuration created the CCM session e-oam-lh-essex with MEP 701 and bound that MEP to port 10 and VLAN 701 as port 10 was an egress port of the monitored service with VLAN ID 701. MEP 701 had the Direction parameter set to Up as it had to send its messages up to the bridge entity of switch wwp2.dev.aj.net (i.e. inside the switch towards the exit port 26). The CCM session e-oam-lh-essex had to be a Level 3 session, which is the default level in Ciena's implementation of CFM.

A virtual switch e-oam was created on wwp2.dev.ja.net to add an outer VLAN tag with VLAN ID 206:

```
virtual-circuit ethernet create vc e-oam vlan 206
virtual-switch ethernet create vs vs_206 vc e-oam
virtual-switch ethernet add vs vs_206 port 10 vlan 701
```

Switch Extreme 4A had the following configuration to support the e-oam-lh-essex session:

```
create cfm domain string "md3" md-level 3
create cfm domain string "md4" md-level 4
configure cfm domain "md3" add association string "e-oam-lh-essex" vlan
      "JRA_TEST"
configure cfm domain "md3" association "e-oam-lh-essex" destination-mac-type
      unicast
configure cfm domain "md3" association "e-oam-lh-essex" add remote-mep 701 mac-
      address 00:02:a1:22:3a:6b
configure cfm domain "md3" association "e-oam-lh-essex" ports 2:3 add end-point
      down 701
```

3.8.3 Test Description

The test consisted of the following steps:

1. Establish CFM CCM session e-oam-lh-essex on Level 3 between the Ciena wwp2.dev.ja.net switch in the JANET LH testbed and the Extreme 4B switch in the JANET Essex University testbed.
2. Check whether the e-oam-lh-essex session correctly shows the status of the VLAN 701 service when the service is up and down by invoking the service status through an appropriate CLI command.
3. Check whether both the e-oam-lh-session and the ISG CFM session show the same status of the VLAN 701 service.

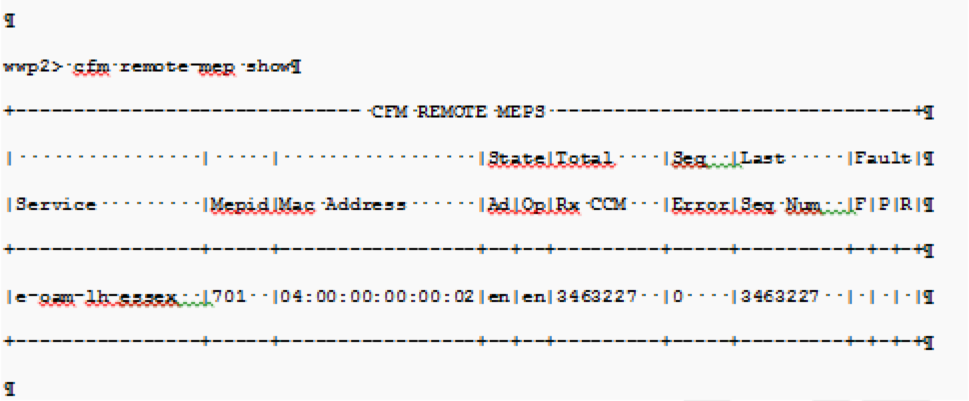
3.8.4 Expected Results

The expected result was the correct detection of the VLAN 701 service's status in its up and down state by the e-oam-lh-essex session.

3.8.5 Results

3.8.5.1 *Monitoring of a Healthy Service*

When the VLAN 701 service was up (i.e. the end hosts 10.1.0.1 and 10.0.0.20 could ping each other), the CFM CCM session e-oam-lh-essex correctly showed the service status. For example, the command `cfm remote-mep show` on wwp2.dev.ja.net showed the following:



```
¶
wwp2> cfm remote-mep show¶
+------------------------------ CFM REMOTE MEPS -------------------------------+¶
|                |     |                 |State|Total    |Seq  |Last     |Fault|¶
|Service         |Mepid|Mac Address      |Ad|Op|Rx CCM   |Error|Seq Num  |F|P|R|¶
+----------------+-----+-----------------+--+--+---------+-----+---------+-+-+-+¶
|e-oam-lh-essex  |701  |04:00:00:00:00:02|en|en|3463227  |0    |3463227  | | | |¶
+----------------+-----+-----------------+--+--+---------+-----+---------+-+-+-+¶
¶
```

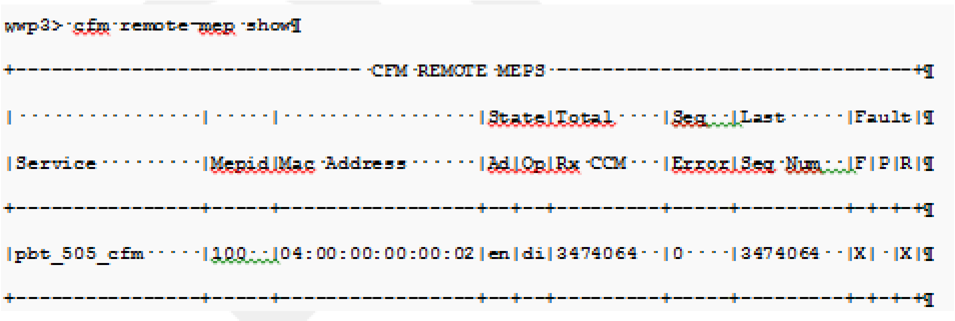
Figure 3.28: Result of cfm remote–mep show command for a healthy service

The screenshot in Figure 3.28 confirms that the administrative and operational states of the session have the value en (i.e. enabled) and no faults were indicated.

This result coincided with the result of monitoring the service VLAN 701 by the ISG boxes, which was visualised by CyPortal.

3.8.5.2 *Monitoring of a Broken Service*

When the service VLAN 701 was down because port 25 of wwp1.dev.ja.net was switched off, the CLI confirmed the Fault state of the e-oam-lh-essex session, as shown in Figure 3.29 below:



```
wwp3> cfm remote-mep show

+------------------------------ CFM REMOTE MEPS --------------------------------+
|                 |     |                 |State|Total    |Seq  |Last       |Fault|
|Service          |Mepid|Mac Address      |Ad|Op|Rx CCM   |Error|Seq Num    |F|P|R|
+-----------------+-----+-----------------+--+--+---------+-----+-----------+-+-+-+
|pbt_505_cfm      |100  |04:00:00:00:00:02|en|di|3474064  |0    |3474064    |X| |X|
+-----------------+-----+-----------------+--+--+---------+-----+-----------+-+-+-+
```

Figure 3.29: Result of cfm remote–mep show command for a broken service

This result coincided with the result of monitoring the service VLAN 701 by the ISG boxes, which was visualised by CyPortal.

3.8.5.3 *Cross-Connection of Levels*

Actually, the diagram in Figure 3.27 shows the setup of the combined testbeds after fixing a problem with cross-connecting the e-oam-lh-essex session (Level 3) and the janet-essex session (Level 5) of the ISG boxes. Initially the setup was different as the ISG box in the JANET Essex University testbed was mistakenly positioned between the Extreme 4A and 4B switches.

As a result, the Level 3 and Level 5 sessions became cross-connected (i.e. overlapped, as shown in Figure 3.30 below), while according to the standards IEEE 802.1ag and ITU-T Y.1731, sessions of different levels that monitor the same service must be nested but not overlapped. In the present case, this meant that the Level 3 session should be nested into the Level 5 session.

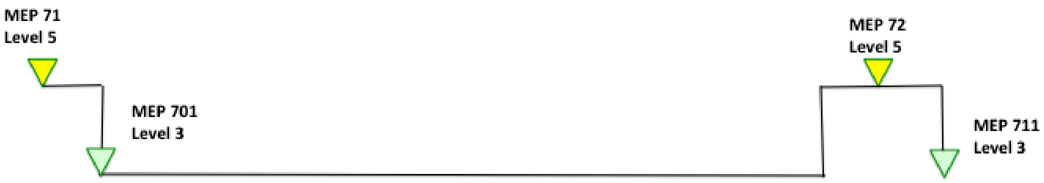


Figure 3.30: Overlapping of janet-essex and e-aom-lh-essex sessions

The incorrect setup was in place for several weeks before the misplacement was noticed and fixed. The setup was changed to the correct arrangement with nested sessions, as shown in Figure 3.27.

In theory, the janet-essex session should not have shown correct results for the status of the VLAN 701 service while overlapping was taking place, because MEP 72 at Essex University should have blocked CCM messages from and to MEP 711. This had to happen according to the principle of a multi-level MEP organisation described in Section 5.7 of Y.1731 [ITU-T Y.1731]:

> "The MEP allows OAM frames from outside administrative domains belonging to higher level MEs to pass transparently; while it blocks OAM frames from outside administrative domains belonging to same or lower level MEs."

In practice, however, MEP 72 did not block messages going between MEP 701 and MEP 711, so both sessions, janet-essex and e-oam-lh-essex, were able to receive CCMs for their level and to determine correctly the status of the VLAN 701 service. At the same time, MEP 711 and MEP 72 sent CCM with RDI=1, which indicated the problem seen from the Essex end of the sessions. The ISG boxes in the JANET LH testbed also indicated cross-connect condition as X-CCM=A (i.e. Active).

The reason for the non-blocking of lower level CCMs by MEP 72 is unknown.

3.8.6 Test Conclusions

The conclusions drawn from Test 5 were as follows:

- Hierarchical two-level monitoring of VLAN 701 between the JANET LH and JANET Essex University testbeds was successful and provided independent and correct results for the service status.
- The incorrect overlapping setup of two CCM sessions resulted in the correct indication of this situation by Y.1731 agents in ISG, Ciena and Extreme switches by means of the RDI. However, this did not distort the normal operation of both sessions, which showed the service status correctly, despite the Y.1731 standard's recommendations that leaking CCMs of a lower level session should be blocked.

3.9 Test 6: Investigation of the One-Way Delay Measurement Problem

This test was carried out to identify the reason for the ISG boxes' incorrect measurement of delay values between the JANET LH and NORDUnet testbeds in Test 3 (described in Section 3.6.5.2 on page 70).

3.9.1 Test Setup

As reported in Section 3.6.5.2, two monitoring sessions, janet-essex and janet-copenhagen, by ISG boxes gave no results for One-way Average Delay (OAD) parameters while other delay and jitter results such as OADV, TAD and TADV were correct.

The most likely causes were the Ciena 311v switches in the JANET LH testbed, as DMM (Delay Measurement Message) and DMR (Delay Measurement Response) frames passed those switches and hence the switches had an opportunity to tamper somehow with the ISG messages. (The CFM functionality of the Ciena switches was enabled as they supported a separate Level 3 CCM session with the Extreme switches in the JANET Essex University testbed.)

In principle, the Extreme switches could also have interfered with the janet-essex session to prevent the correct measurement of OAD, but it seemed less likely as the janet-copenhagen session showed the same problem with no transit switches on the NORDUnet side.

Neither the Lightpath equipment nor the NORDUnet routers had been configured to support Y.1731 functionality, so the probability of their interfering was low.

To check whether the Ciena switches were interfering with the ISG delay measurement sessions, two test setups for the JANET LH–NORDUnet service were used:

- Setup 1: with a transit Ciena 311v switch, which was the normal setup used in all other tests (Figure 3.31).
- Setup 2: without a transit Ciena switch, where the ISG boxes in the JANET LH and NORDUnet testbeds were connected directly through JANET Lightpath and a NORDUnet IP/MPLS tunnel (Figure 3.32).

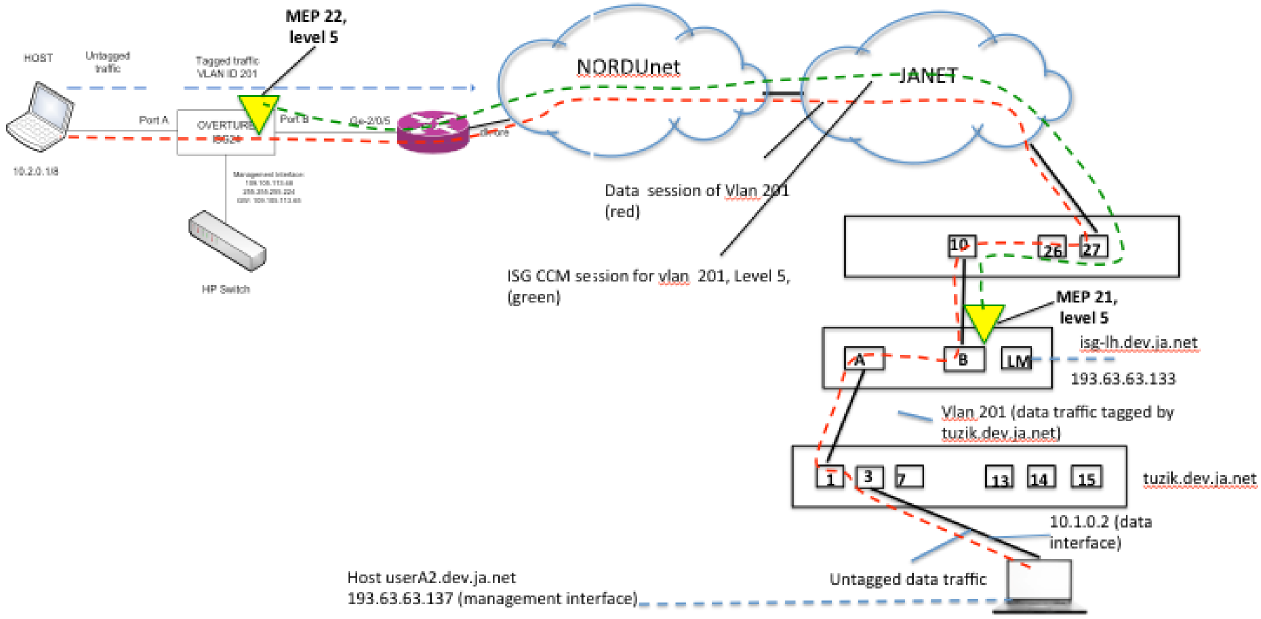


Figure 3.31: Setup 1: combined JANET LH and NORDUnet testbeds with a transit Ciena 311v switch

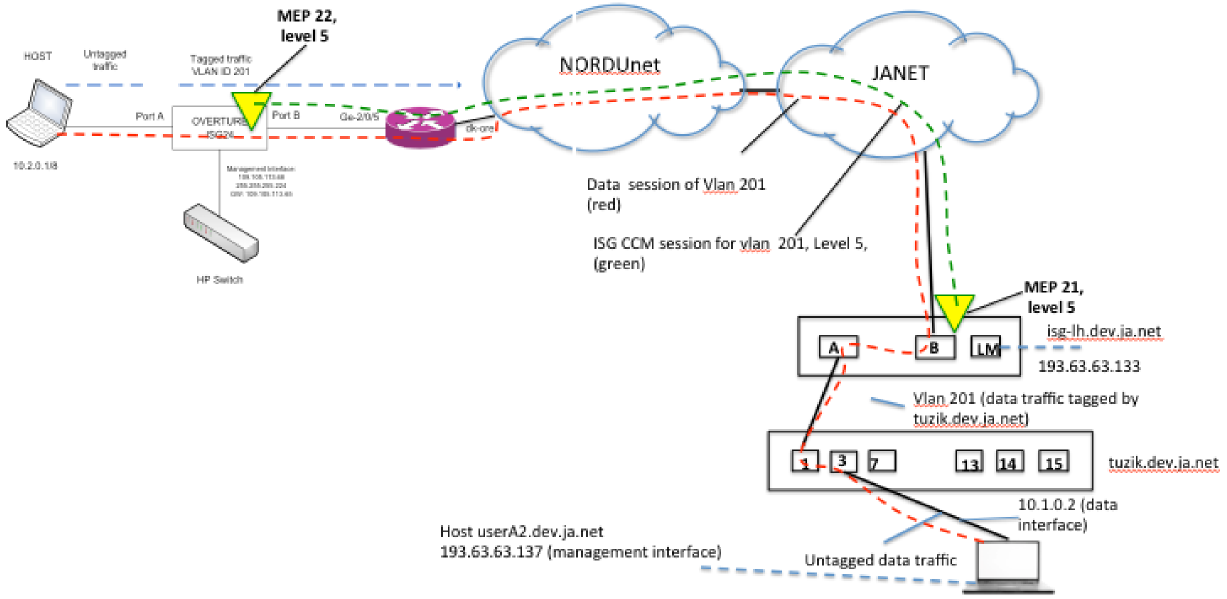


Figure 3.32: Setup 2: Combined JANET LH and NORDUnet testbeds without a transit Ciena 311v switch

Similar setups for the JANET LH–JANET Essex University service turned out to be impossible as direct ISG connection to the Lightpath would need the ISG box to generate double-tagged frames (206:701). The version of ISG box used could not do this, therefore only the JANET LH–NORDUnet service was investigated.

3.9.2 Configuration

The configuration of the CFM DMM sessions in both the JANET LH and NORDUnet ISG boxes was unchanged from that used for Test 3 (described in Section 3.6.2) when the strange behaviour of the delay measurements was noticed.

The results of PAA session TEST-CPH were used as a baseline for comparison with the CFM DMM results.

3.9.3 Test Description

The test consisted of the following steps:

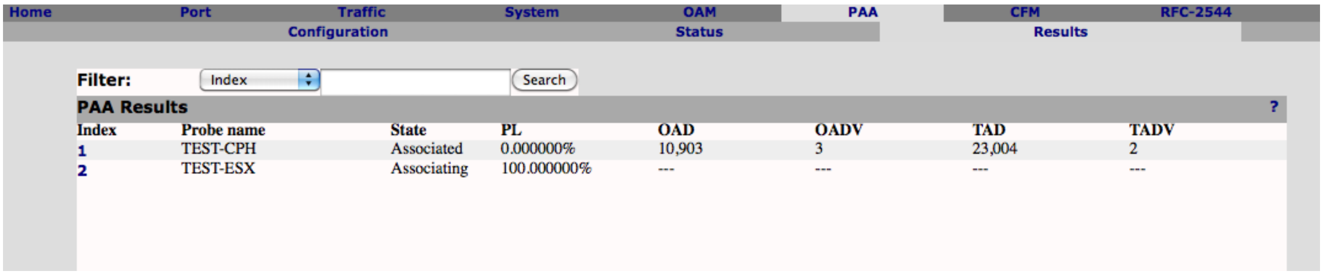
1. Deploy Setup 1, as shown in Figure 3.31.
2. Generate traffic between the JANET LH and NORDUnet testbeds using Iperf software.
3. Record the results of CFM DMM and PAA sessions for OAD, OADV, TAD and TADV parameters.
4. Deploy Setup 2, as shown in Figure 3.32.
5. Repeat steps 2 and 3.
6. Compare the results obtained for Setup 1 and Setup 2.

3.9.4 Expected Results

It was expected that with Setup 2 deployed the ISG boxes would be able to measure OAD values.

3.9.5 Results

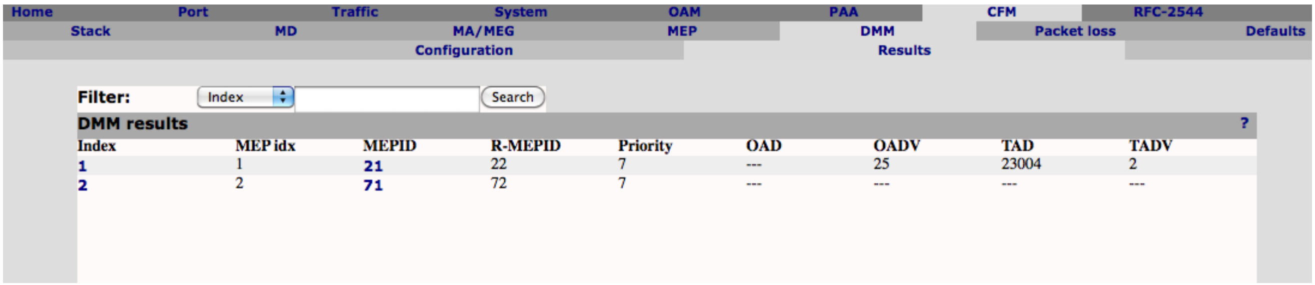
When Setup 1 was deployed, the results of the PAA and CFM DMM sessions on the JANET LH ISG box were as shown in Figure 3.33 and Figure 3.34 below respectively.



PAA Results

Index	Probe name	State	PL	OAD	OADV	TAD	TADV
1	TEST-CPH	Associated	0.000000%	10,903	3	23,004	2
2	TEST-ESX	Associating	100.000000%	---	---	---	---

Figure 3.33: TEST-CPH PAA session (JANET LH box)



DMM results

Index	MEP idx	MEPID	R-MEPID	Priority	OAD	OADV	TAD	TADV
1	1	21	22	7	---	25	23004	2
2	2	71	72	7	---	---	---	---

Figure 3.34: CFM DMM session (JANET LH box)

When Setup 2 was deployed, the results turned out to be the same, i.e. the ISG boxes couldn’t measure OAD in either setup.

Different NTP servers had been used to synchronise the JANET LH and NORDUnet boxes, as it was assumed that poor synchronisation was the reason for the void OAD measurements. However, changing the NTP server did not impact on the OAD measurement results.

3.9.6 Test Conclusions

The conclusions drawn from Test 6 were as follows:

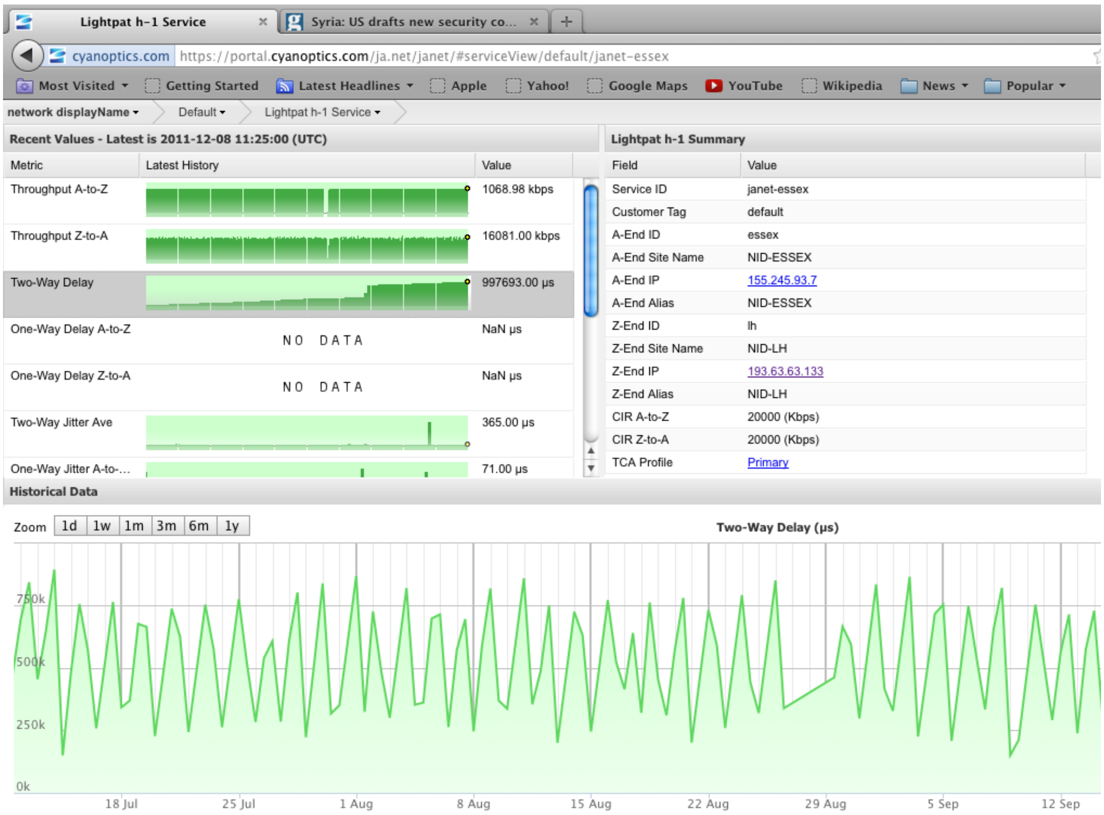
- The reason for the inability of the ISG boxes to measure OAD values for the janet-copenhagen session was not interference from a transit Ciena 311v box, as direct connection of the ISG boxes resulted in the same behaviour of the CFM DMM protocol.
- The proprietary PAA protocol of the ISG boxes turned out to be more robust than the standard CFM DMM one as the PAA protocol was able to measure OAD values correctly while CFM DMM was not.
- Changing the NTP server as a source of synchronisation didn’t impact on CFM DMM measurements.
- The reason for the inability of the CFM DMM protocol to measure OAD values for the janet-copenhagen session remains unclear, despite discussions with the vendor and a site visit from a vendor engineer. The matter will be investigated further in the planned Service Assurance and Monitoring activity.

3.10 Test 7: Investigation of Saw-Tooth Shape of janet-essex Jitter Measurements

This test was carried out to identify the reason for the saw-tooth shape of the ISG boxes’ measurement of jitter values between the JANET LH and NORDUnet testbeds in Test 3 (described in Section 3.6.5.2 on page 70).

3.10.1 Test Setup

As reported in Section 3.6.5.2, the Two-way Average Delay (TAD) parameter of the janet-essex CFM DMM sessions measured by the ISG boxes in the JANET LH and JANET Essex University testbeds showed a strange regular saw-tooth shape, with peaks at about 800 ms and drops to 50 ms – 200 ms over periods of about 2 days. This TDA behaviour is shown in Figure 3.35 below. At the same time, the PAA session TEST-ESX showed stable results for TAD of 15.5 ms.



Recent Values - Latest is 2011-12-08 11:25:00 (UTC)

Metric	Latest History	Value
Throughput A-to-Z		1068.98 kbps
Throughput Z-to-A		16081.00 kbps
Two-Way Delay		997693.00 µs
One-Way Delay A-to-Z	NO DATA	NaN µs
One-Way Delay Z-to-A	NO DATA	NaN µs
Two-Way Jitter Ave		365.00 µs
One-Way Jitter A-to-...		71.00 µs

Lightpat h-1 Summary

Field	Value
Service ID	janet-essex
Customer Tag	default
A-End ID	essex
A-End Site Name	NID-ESSEX
A-End IP	155.245.93.7
A-End Alias	NID-ESSEX
Z-End ID	lh
Z-End Site Name	NID-LH
Z-End IP	193.63.63.133
Z-End Alias	NID-LH
CIR A-to-Z	20000 (Kbps)
CIR Z-to-A	20000 (Kbps)
TCA Profile	Primary

Historical Data

Figure 3.35: The saw-tooth shape of TAD for janet-essex session

To investigate the reason for this phenomenon, two changes were made to the setup of the JANET LH testbed. MEP 701 was moved from port 10 of the wwp2.dev.ja.net switch to port 28 of the wwp1.dev.ja.net switch, as shown in Figure 3.36 below. The move was made to allow the capture of janet-essex DMM traffic on the Ciena switches' ports before and after MEP 701 to see whether the Ciena MEP was interfering in the “foreign” (from the Ciena box's point of view) janet-essex session of the ISG boxes.

Another change to the setup was to replace the Overture ISG26 box with an Accedian [Accedian] MetroNID box. This was done because the original Overture ISG26 box had an experimental version of software used for demo sessions, which could have been the reason for the non-standard measurement of TAD. To exclude this possibility, a box with standard software from Accedian was used (MetroNID was the original model which was used by Overture for OEMing its ISG model line).

3.10.2 Configuration

The Accedian MetroNID was configured exactly as the Overture ISG26 had been.

The session janet-essex and MEP 701 at port 10 of wwp2.dev.ja.net were disabled and then activated with the same parameters at port 28 of wwp1.dev.ja.net.

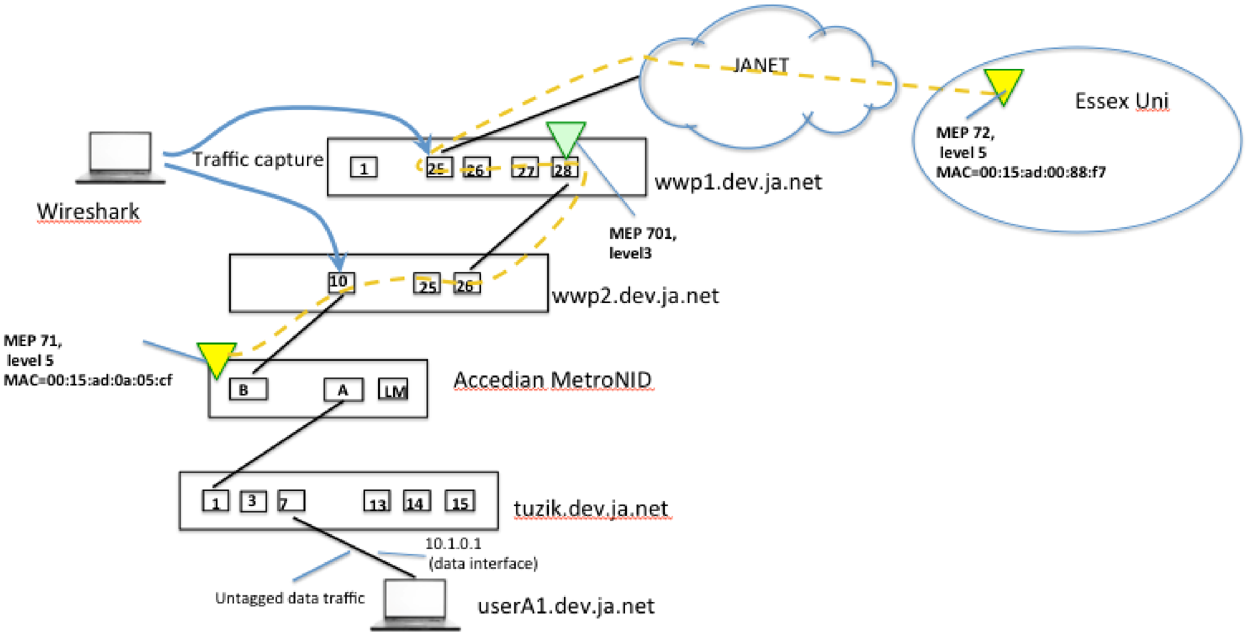


Figure 3.36: JANET LH testbed setup to investigate the reason for the TAD saw-tooth shape

3.10.3 Test Description

The test consisted of the following steps:

1. Monitor the TAD parameter of the janet-essex session with an Accedian MetroNID box installed instead of the Overture ISG26 in the JANET LH testbed.
2. Capture DMM and DMR frames of the janet-essex session at port 10 of wwp2.dev.ja.net and port 28 of wwp1.dev.ja.net and compare their fields.
3. Disable the CFM protocol on wwp1.dev.ja.net and wwp2.dev.ja.net boxes and check whether TAD behaviour returns to normal.

3.10.4 Expected Results

The expected results of the test were as follows:

- Replacing the Overture ISG26 with an Accedian MetroNID should result in stable measurement of TAD at a 15 ms level (the value from PAA TAD monitoring).
- If TAD keeps showing the saw-tooth shape, then try to identify what is wrong using the DMM and DMR frames captured at different ports of the Ciena switches.

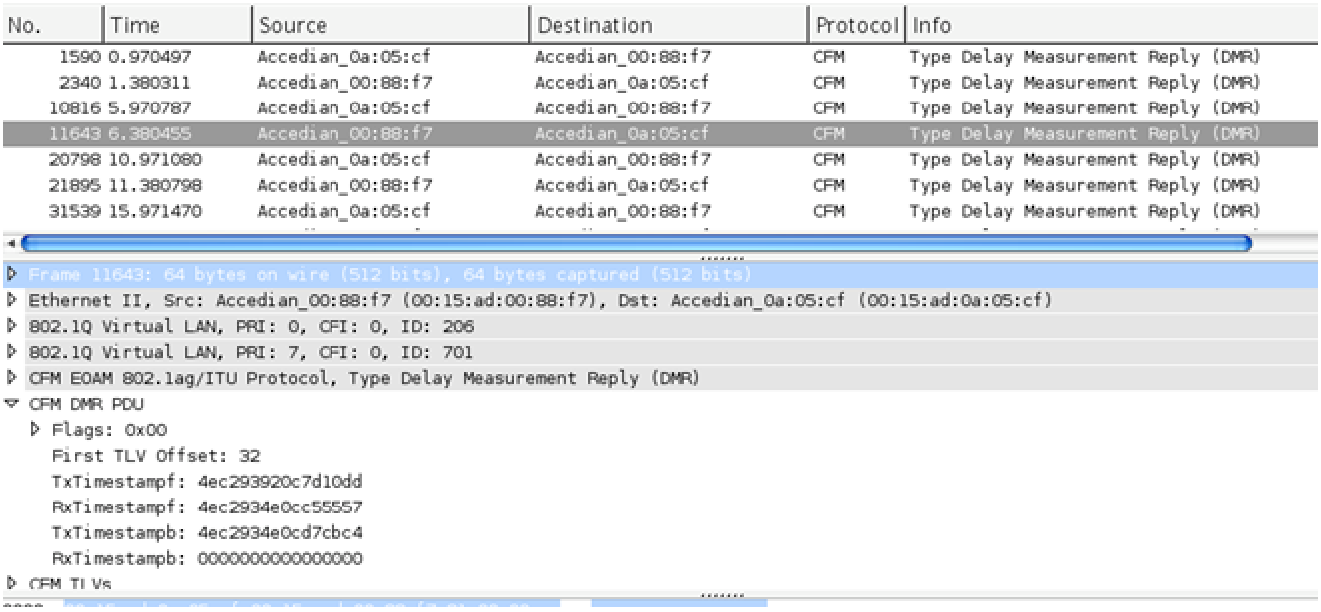
3.10.5 Results

3.10.5.1 *Replacing Overture ISG26 with Accedian MetroNID*

Replacing the ISG box did not change the behaviour of TAD for the janet-essex session – the graph repeated the saw-tooth shape. The demo model of ISG26 was therefore not the reason for such TAD measurements.

3.10.5.2 *Analysis of Captured Frames*

An analysis of captured frames from the janet-essex session allowed the real reason for the incorrect TAD measurements to be identified: the wwp1.dev.ja.net switch was inserting its own timestamps into the DMR messages of the janet-essex session. A DMR message sent by the Essex University ISG box (MAC 00:15:ad:00:88:f7) to the JANET LH Accedian MetroNID box (MAC 00:15:ad:0a:05:cf), captured at port 28 of wwp1.dev.ja.net (Figure 3.37), helps to explain how this was happening.



```
No.    Time       Source              Destination         Protocol Info
  1590 0.970497   Accedian_0a:05:cf   Accedian_00:88:f7   CFM      Type Delay Measurement Reply (DMR)
  2340 1.380311   Accedian_00:88:f7   Accedian_0a:05:cf   CFM      Type Delay Measurement Reply (DMR)
 10816 5.970787   Accedian_0a:05:cf   Accedian_00:88:f7   CFM      Type Delay Measurement Reply (DMR)
 11643 6.380455   Accedian_00:88:f7   Accedian_0a:05:cf   CFM      Type Delay Measurement Reply (DMR)
 20798 10.971080  Accedian_0a:05:cf   Accedian_00:88:f7   CFM      Type Delay Measurement Reply (DMR)
 21895 11.380798  Accedian_00:88:f7   Accedian_0a:05:cf   CFM      Type Delay Measurement Reply (DMR)
 31539 15.971470  Accedian_0a:05:cf   Accedian_00:88:f7   CFM      Type Delay Measurement Reply (DMR)

▷ Frame 11643: 64 bytes on wire (512 bits), 64 bytes captured (512 bits)
▷ Ethernet II, Src: Accedian_00:88:f7 (00:15:ad:00:88:f7), Dst: Accedian_0a:05:cf (00:15:ad:0a:05:cf)
▷ 802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 206
▷ 802.1Q Virtual LAN, PRI: 7, CFI: 0, ID: 701
▷ CFM EOAM 802.1ag/ITU Protocol, Type Delay Measurement Reply (DMR)
▽ CFM DMR PDU
  ▷ Flags: 0x00
    First TLV Offset: 32
    TxTimestampf: 4ec293920c7d10dd
    RxTimestampf: 4ec2934e0cc55557
    TxTimestampb: 4ec2934e0cd7cbc4
    RxTimestampb: 0000000000000000
▷ CFM TLVs
```

Figure 3.37: DMR message captured at port 28 of wwp1.dev.ja.net

The DMR message is a reply from MEP 72 (on the JANET Essex University ISG24 box) to a DMM message sent by MEP71 (on the JANET LH MetroNID box). It has 4 timestamp fields:

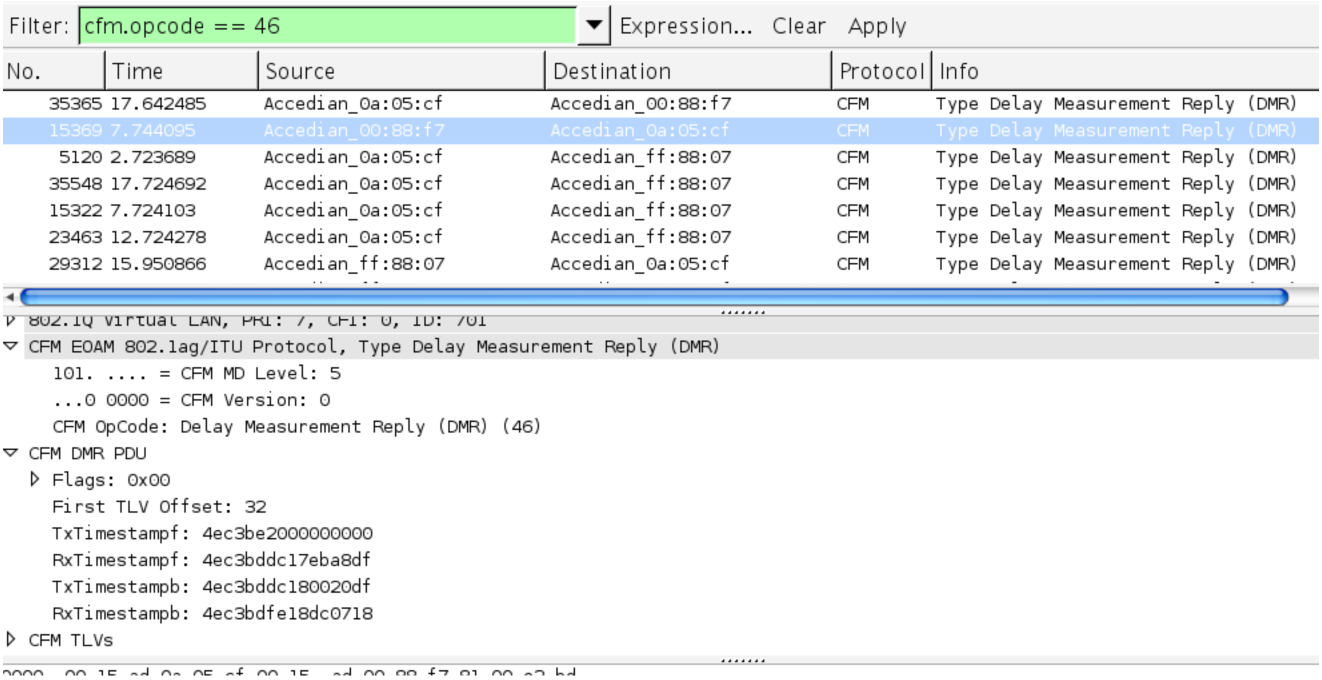
- TxTimestampf: the timestamp at the transmission time of the DMM frame (made by MEP 71 in its DMM frame and repeated by MEP 72 in its DMR reply).
- RxTimestampf: the time of receiving the DMM frame by MEP 72.
- TxTimestampb: the time of transmitting the DMR frame by MEP 72.
- RxTimestampb: the time of receiving the DMR frame by MEP 71.

When a MEP receives a DMR message from its remote counterpart, it uses the timestamps to calculate two-way frame delay using the following equation [ITU-T Y.1731]:

Frame Delay = (RxTimestampb – TxTimestampf) – (TxTimestampb – RxTimestampf)

In the DMR message in Figure 3.37, RxTimestampb = 0; this is correct as the frame was travelling to its destination and this field had to have 0 value as the destination hadn't been reached yet.

What is surprising is that the same frame captured at port 10 of wwp2.dev.ja.net had RxTimestampb filled with a non-zero value! This is clearly seen in Figure 3.38:



Filter: cfm.opcode == 46

```
No.   Time        Source               Destination          Protocol Info
35365 17.642485   Accedian_0a:05:cf    Accedian_00:88:f7    CFM      Type Delay Measurement Reply (DMR)
15369 7.744095    Accedian_00:88:f7    Accedian_0a:05:cf    CFM      Type Delay Measurement Reply (DMR)
 5120 2.723689    Accedian_0a:05:cf    Accedian_ff:88:07    CFM      Type Delay Measurement Reply (DMR)
35548 17.724692   Accedian_0a:05:cf    Accedian_ff:88:07    CFM      Type Delay Measurement Reply (DMR)
15322 7.724103    Accedian_0a:05:cf    Accedian_ff:88:07    CFM      Type Delay Measurement Reply (DMR)
23463 12.724278   Accedian_0a:05:cf    Accedian_ff:88:07    CFM      Type Delay Measurement Reply (DMR)
29312 15.950866   Accedian_ff:88:07    Accedian_0a:05:cf    CFM      Type Delay Measurement Reply (DMR)
```

```
▷ 802.1Q Virtual LAN, PRI: 7, CFI: 0, ID: 701
▽ CFM EOAM 802.1ag/ITU Protocol, Type Delay Measurement Reply (DMR)
    101. .... = CFM MD Level: 5
    ...0 0000 = CFM Version: 0
    CFM OpCode: Delay Measurement Reply (DMR) (46)
▽ CFM DMR PDU
  ▷ Flags: 0x00
    First TLV Offset: 32
    TxTimestampf: 4ec3be2000000000
    RxTimestampf: 4ec3bddc17eba8df
    TxTimestampb: 4ec3bddc180020df
    RxTimestampb: 4ec3bdfe18dc0718
▷ CFM TLVs

0000  00 15 ad 0a 05 cf 00 15  ad 00 88 f7 81 00 e2 bd
```

Figure 3.38: DMR message captured at port 10 of wwp2.dev.ja.net

The non-zero value of RxTimestampb was definitely a violation of the Y.1731 protocol as the DMR message at port 10 was still travelling to its destination of MEP 71 at port 28 of the MetroNID and hence had to have a zero value.

It was therefore the CFM stack of the wwp1.dev.ja.net switch that put the non-zero value into the RxTimestampb field. According to the Y.1731 standards, this should not have happened as the CFM stack of wwp1.dev.ja.net was configured to support the Level 3 session e-oam-lh-essex and not the Level 5 session janet-essex, whose DMR messages were the ones being interfered with.

It seems that receiving DMR messages with the RxTimestampb values already set disoriented the normal algorithm for calculating the TAD of MEP 71 and caused the saw-tooth shape of the TAD graph. Why it resulted in this particular shape of graph is not clear, as it depends on the details of the implementation of TAD calculations in the Overture and Accedian boxes. (The matter was reported to Accedian and Cyan, Inc. engineers. However, since the vendors were participating in the trial free of charge, it was difficult for them to commit the resources necessary to identify the cause, particularly when the equipment of a third party (Ciena), who did not actively participate in the trial (and so was not contacted about the problem), was involved.) What is clear is that such a situation is non-standard and it should not have happened.

3.10.5.3 *Disabling CFM Functionality on Ciena Switches*

When the CFM protocol was disabled on the Ciena switches, the results of the TAD calculation by the MetroNID box became stable at close to 15.5 ms, which coincided with the PAA TAD results.

3.10.6 Test Conclusions

The conclusions drawn from Test 7 were as follows:

- The reason for the incorrect calculation of the TAD parameter by the Overture and Accedian boxes in the JANET LH testbed was the incorrect timestamping of their DMR messages by the transit Ciena switch.
- This timestamping was in violation of the Y.1731 standard, as the Ciena switch supported a CFM session on Level 3 while the Accedian box supported a CFM session on level 5.

3.11 Conclusions

The main finding from the trial is that the Ethernet OAM functions embedded in the carrier-grade Ethernet equipment can be used for effective monitoring and visualising of the health and performance of wide-area Ethernet services in a range of scenarios: within NRENs, between NRENs and on the GÉANT backbone. These functions are standardised by a number of IEEE, ITU-T and MEF specifications, and vendor implementations are close enough to those specifications to provide interoperability between OAM agents. In some cases, Ethernet OAM functions are already available in the network routers and switches currently used in NRENs. However, NREN staff's experience in the use of these functions is still limited and, as a result, the available functions remain largely unused. In cases when existing network equipment does not support Ethernet OAM functions, NRENs and GÉANT should consider making such support a mandatory requirement in all future procurements.

The use of dedicated Ethernet demarcation boxes with a rich set of OAM functions (Overture ISG and Accedian MetroNID) proved to be an effective way of monitoring Ethernet services end to end. It was especially useful in the current situation, where not all the trial participants (notably the NORDUnet and CESNET testbeds) had equipment that supported the required monitoring OAM functions. For NRENs in the same situation, the use of Ethernet OAM demarcation boxes would also seem to be an effective option.

The CyPortal software, which stored and visualised monitoring data in diverse forms, was an important and useful element of the trial. It showed that this type of software is a crucial element in an NREN's Operations Support System (OSS) environment when the NREN wants to offer managed, monitored and visible Ethernet services to their customers. For example, such software could be used for end-to-end monitoring of the GÉANT Plus or JANET Lightpath services. Integrating such systems with monitoring tools developed by the R&E community – with perfSONAR, for example – looks promising and will be explored in GN3 Y4 by the Service Assurance and Monitoring activity.

The version of CyPortal that was used in the trial had some limitations, such as support of only a few types of equipment (namely, only Accedian and Overture demarcation boxes); a long data refreshing interval (no less that 15 minutes); and monitoring services only end to end (i.e. no support for multi-level hierarchical monitoring). For large-scale deployments of managed Ethernet services, such limitations should be lifted.

In most tests the results of monitoring the services' health (status) and such performance parameters as frame delay, jitter, and loss were stable and corresponded to expected results. Only the measurements of delays between JANET LH and two other testbeds (JANET Essex University and NORDUnet) by means of the standard Y.1731 protocol turned out to be incorrect. The investigation of the inability of the Y.1731 protocol to measure One-way Average Delay did not reveal the reason for such behaviour; one of the possible reasons might be a problem with the synchronisation of the boxes. The investigation of incorrect measurements of Two-way Average Delay established the cause of the saw-tooth shape of the data graph: one of the Ciena boxes in the JANET LH testbed incorrectly inserted its timestamps in a "foreign" monitoring session. The exact reason for such incorrect behaviour of the Ciena box is unknown; possibly it was just a bug.

The trial participants managed to carry out tests only on an end-to-end basis. Monitoring and troubleshooting multi-segment services with intermediate monitoring points, which are needed for a large-scale deployment of monitoring tools and systems in the multi-domain NREN environment, are a logical extension of the trial and will be conducted in GN3 Y4 by the Service Assurance and Monitoring activity.

4 Synchronous Ethernet

4.1 Overview

Synchronous Ethernet (SE) was tested on Cisco Systems equipment. SE is a new standard to provide synchronisation between Ethernet interfaces. A detailed description of SE features and functionalities can be found in "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1]. Additional testing was performed with Precision Time Protocol (PTP), since all the available equipment tested was PTP capable.

4.2 Introduction

This chapter contains the results of testing Synchronous Ethernet (SE) implementation on Cisco Systems equipment. SE is a new Ethernet development which has been identified as a Carrier Class Transport Network Technology (CCTNT) by the GN3 project. The main goal of the tests was to gain practical experience with the new technology, because it is rather different from traditional Ethernet. Another goal was to assess the current status of implementation of SE on different series of routers. Additional testing was performed with Precision Time Protocol (PTP), since all the available equipment tested was PTP capable.

4.3 Technology Briefing

SE is a new standard defined by ITU-T for the distribution of accurate frequency over Ethernet ports and links. This frequency synchronisation is usually referred to as "timing" and must be distinguished from time synchronisation, which is the distribution of so-called "time of day" or "wall-clock time". SE cannot distribute wall-clock time; it distributes only accurate frequency by means of the very precise control of the bit rate of an Ethernet link. Other methods and protocols have been proposed and developed to accomplish the task of accurate time synchronisation and distribution, for example IEEE 1588 Precision Time Protocol (PTP), IEEE 802.1 AS (newer, based on IEEE 1588) or IETF's Network Time Protocol (NTP). More information is available in "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1] and is not repeated here.

4.4 Test Objective

The objective of the tests was to gain practical experience with the new technology, and to assess the status of implementation of SE and PTP on Cisco Systems routers and other equipment. The new Ethernet features could be interesting for new applications in research and education networks.

4.5 Test Infrastructure

4.5.1 Description

The tests can be divided into two categories. One, testing in the CESNET lab with Cisco MWR 2941 boxes (Mobile Wireless Routers supporting both SE and PTP) together with Meinberg LANTIME M600. M600 is a Stratum 1 NTP Time Server capable of PTP, but SE is not supported. Detailed descriptions of these devices are out of scope of this document; interested readers can find more information online [CiscoMWR2941], [MbergLANTIMEM600]. Two, tests performed remotely with the help of Cisco labs and student courses. In these tests, more equipment was available, together with demonstration of frequency/time measurements.

4.5.2 Configurations

The configurations were relatively simple and information about them is given in each test description.

4.6 Test 1: Network Clocking with PTP and SE

4.6.1 Test Objective

The objective of the test was to verify the operation of PTP and SE between devices in a simple topology.

4.6.2 Test Setup

There is no radio access network (RAN) equipment (e.g. base transceiver stations (BTSs) or radio network controllers (RNCs)) available in the CESNET lab, so two MWR 2941 routers were configured to provide trivial SE connectivity between them. Network clocking, which is the most important part, can be configured to receive clocks from different sources. Three options are available on MWR 2941: PTP, pseudowire-based or SE. Configuration options are rich; for example, for PTP the MWR can act as slave or master, can use unicast or multicast modes. New models of MWR 2941 can provide a GPS antenna option, but this feature was not available; the Meinberg LANTIME M600 was therefore used to connect to GPS antennas and obtain GPS signals. PTP was enabled between the Meinberg and Cisco equipment.

In the first test, the purpose was to enable PTP on every device and configure network clocking to be distributed via PTP. The Meinberg M600 was configured as the master clock (i.e. very accurate source of synchronisation because of a very stable PTP clock) for PTP clients or “slaves” (see Figure 4.1). PTP can run over any IP-based infrastructure; in this test, standard Ethernet ports were used. It should be noted that MWR 2941 is SE capable but M600 is not. However, this is not a problem because PTP has been designed to work in IP networks and transfer accurate timing signals without any special requirements. Synchronous Ethernet with accurate frequency transfer can help PTP, but SE can be disabled on MWR 2941 and PTP works without any problem.

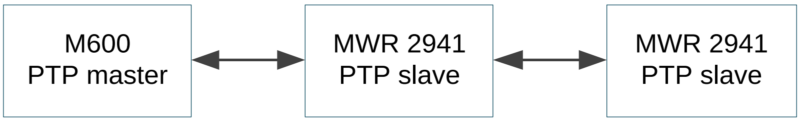


Figure 4.1: PTP master and slave clocks

In the second test (see Figure 4.2), MWR 2941 routers were connected and SE was used as the source of accurate timing. An E1 tester was used to check the quality of transmission. Because of the very simple setup, no errors were observed. Ethernet Synchronisation Messaging Channel (ESMC) and Synchronisation Status Message (SSM) options are available for additional configuration; troubleshooting these is less straightforward.

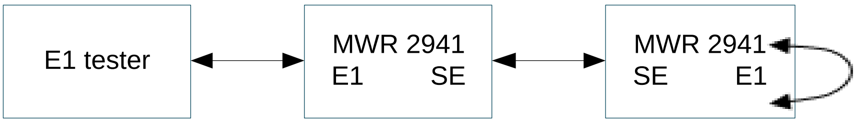


Figure 4.2: Synchronous Ethernet and E1 testing

4.6.3 Configuration

Internal commands can be used to verify clocking distribution and proper PTP and SE configurations. There are many options for configuring PTP, but few commands available for SE and ESMC.

4.6.4 Expected Results

PTP or SE clocking is enabled and PTP/SE is running between nodes.

4.6.5 Results

PTP was configured in unicast mode and operation was verified with Cisco show commands. PTP multicast options were not appropriate given the simple experimental setup. SE configuration options are simpler than for PTP, and ESMC/SSM configurations are relatively uncomplicated.

4.6.6 Test Conclusions

PTP and SE worked as expected, mainly because of the rather simple setup with two instances of SE/PTP equipment and one PTP Stratum 1 server only. More configuration and debugging challenges can be expected in more complex environments. Testing equipment to verify synchronisation of a network was available during the test, but the hardware failed to operate properly. (There are no plans to re-do the test, or to enhance the test setup, since the main test objective was met and network synchronisation was achieved. However, CESNET is exploring PTP (and SE to some extent) deployment, and any new findings could be shared with the GÉANT community.)

4.7 Test 2: Remote Labs Testing

4.7.1 Test Objective

The objective of the SE lab was to provide synchronisation over SE interfaces and observe the operation of synchronisation messages and protocols. In addition, the aim was to demonstrate time measurements, like Maximum Time Interval Error (MTIE), which are very important for verifying the status of network synchronisation.

The objective of the PTP lab was to verify the effectiveness of PTP functionality. MTIE measurements were performed for both SE and PTP labs.

4.7.2 Test Setup

The remote Cisco labs can provide complex environments with RAN equipment, for both SE and PTP deployments. More complex topologies, such as rings, require clocking resiliency with access ring protection on different equipment. Procedures and tests include enabling synchronisation and choosing different clocking sources. Different quality-based selection algorithms can then be configured and verified.

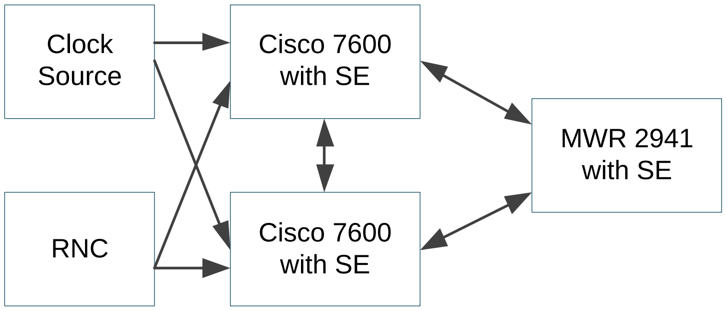


Figure 4.3: SE and PTP in a simple ring topology

For the PTP tests, different configurations were implemented on the routers (e.g. enabling PTP master/slave functionality); the PTP setup is shown in Figure 4.3.

4.7.3 Configuration

Two 7600 routers were shared and configured as the primary clock sources. MWR 2941 routers were then configured to use SE ports to recover the clock across the network. MWR routers must be in the lock state, i.e. clocking was properly recovered from physical Gigabit Ethernet lines.

PTP was enabled on 7600 routers with master functionality and MWR routers acting as PTP slaves with sessions established to the master 7600, and configured to distribute clock information between network elements.

The ESMC protocol was configured to enable quality-based selection algorithms to choose clocking from different sources.

4.7.4 Test Description

The test consisted of the following steps:

1. Configure the routers as outlined in Section 4.7.3.
2. Observe clock state changes.
3. Configure ESMC protocol as outlined in Section 4.7.3.
4. Observe clock state changes and debug as necessary.
5. Configure PTP as outlined in Section 4.7.3.
6. Observe clock state changes and debug as necessary.
7. Perform MTIE measurements.

4.7.5 Expected Results

SE is enabled on all routers and after additional configuration the ESMC protocol is working properly. MWR routers should obtain clocking from the best clock available. When PTP is configured, clocking information is recovered without SE enabled.

4.7.6 Results

SE was configured on all network elements and MWR routers recovered clocking from physical links connected to 7600 routers. 7600 routers used an external clock source connected to the timing input on the special ES+ line card. External input is preferred over an internal clock because it gives better quality. ESMC protocol was configured and MWR routers were able to select the best clocking source in the network.

PTP was configured with master functionality on one 7600 router and using unicast negotiation transport; MWR routers were configured as slaves. The clock was recovered from the established session with the PTP master.

MTIE measurements were performed for both SE and PTP labs but problems were encountered and the results were not predicative.

4.7.7 Test Conclusions

It can be noted that configuration commands are relatively straightforward but debugging and troubleshooting may be cumbersome, especially for routing and switching professionals not very familiar with clocking subtleties. On the other hand, people familiar with legacy Synchronous Digital Hierarchy (SDH) equipment should have no problems when working with SE equipment. The last challenge was to measure clock quality. However, testing equipment for MTIE was not working properly in either the CESNET lab or the Cisco remote labs. The problem was most likely hardware related, but the precise reason for the problems could not be identified.

4.8 Conclusions

SE and PTP were configured and tested in both the CESNET lab and Cisco remote labs. It seems that R&E networks and the R&E community may be focused more on PTP deployment because PTP can provide both frequency and time distribution and no additional hardware is required (though of course the software must be PTP capable). SE, on the other hand, can distribute accurate frequency only and requires new hardware since standard Ethernet ports cannot work with SE ports and provide the required functionality. Future requests from the R&E community are needed to confirm this expectation. Some NRENs may benefit from distribution of accurate frequency for certain new applications like high-definition videoconferencing or metrology. Beyond this, however, there is still uncertainty about the use of SE in the NREN community. While the technology has clearly defined applications in the commercial world – mobile networks, for example – the authors are not aware of any other applications in the NREN environment that currently require the functionality it offers.

Nevertheless, they expect that, in the future, new applications will be identified by the NREN community that will benefit from its use.

Note: SE has been included in the report due to its relevance to NRENs and carriers. However, it should be understood more as a function of the Ethernet technology than as a CCTNT as defined by [DJ1.1.1].

5 Provider Backbone Bridge Traffic Engineering

5.1 Overview

This chapter describes the results of the trial of Provider Backbone Bridge Traffic Engineering (PBB-TE) technology, which belongs to the Carrier Ethernet Transport (CET) family. The trial was carried out in the context of two projects: the JANET Carrier Ethernet project [JANET CE] and the GN3 project (as part of JRA1 Task 1). The JANET Carrier Ethernet project had a wider scope than just investigating and trialling PBB-TE, as some of its participants also investigated Ethernet over Multi-Protocol Label Switching (EoMPLS). This chapter highlights only those results that deal with PBB-TE functionality and its interoperability with EoMPLS, and that were obtained by the participants who worked on both the GN3 JRA1 Task 1 and JANET Carrier Ethernet project trials.

To place PBB-TE in context, the chapter begins by providing an overview of the technology areas to which it belongs: Carrier Ethernet and Carrier Ethernet Transport.

The trial objectives were to:

- Gain practical experience with PBB-TE by trialling its core functionality in a multi-domain and multi-technology (PBB-TE and Ethernet over MPLS) environment.
- Investigate the manageability of PBB-TE within a multi-domain and multi-technology environment (path provisioning, troubleshooting).
- Evaluate the deployment of PBB-TE on JANET (for the core, for Regional Networks (RNs), for IP and Lightpath traffic) and to identify areas for further investigation.
- Evaluate the state of standardisation of PBB-TE.

To achieve these objectives, a series of tests were conducted over a multi-domain, multi-technology testbed consisting of the trial participants' local testbeds plus a core testbed. The testbed included equipment from Ciena, Extreme Networks, Juniper, Foundry Networks / Brocade and Dell. (Ciena was actively involved in the tests, configuring examples of PBB-TE connections over the JANET Lumen House (LH) testbed. Extreme was also involved, giving advice to the Task participants.) The chapter describes the testbed, and the setup, configuration, results and problems encountered for each of the three sets of tests carried out: PBB-TE core tests, and local Essex University and JANET LH tests. All conclusions are presented at the end of the chapter.

The PBB-TE core tests investigated network-to-network connectivity, resilience, OAM and Ethernet Service Manager functionality. The test results were positive. Two problems were encountered during the network-to-network connectivity test: frame dropping, which was solved by increasing the Maximum Transmission Unit (MTU) value of the core switches, and traffic selection from Essex University, which was caused by a bug in the switch software and solved by translating the EtherType value used by the switches in the Essex University testbed.

The local Essex University tests investigated PBB-TE connectivity, resilience, and CFM and PBB-TE protection and resilience. The results were positive; no problems were encountered.

The local JANET Lumen House tests investigated unprotected tunnels, CFM, performance monitoring, path protection and traffic policing. The results were positive. Two constraints were encountered in the unprotected tunnels test: PBT tunnels can be created only on gigabit ports (25-27), and it is not possible to change a B-VID value for a tunnel that was specified during tunnel creation. One problem was encountered in the CFM test: CFM for virtual switches stops working after rebooting, probably the result of an unknown bug in LEOS v4.6. A fix was found and the problem has been reported to the vendor.

Generally, the tests conducted showed that the PBB-TE functionality of the equipment used corresponded to expectations, and demonstrated proper behaviour in the following key areas:

- Manual establishment of TE point-to-point tunnels.
- Separation of customer and provider address spaces.
- CCM monitoring of state of tunnels.
- Fast protection switching.
- Per-VLAN traffic policing.

The participants concluded that PBB-TE as a transport technology is better suited to single- than multi-domain applications; that PBB-TE deployed in one domain can smoothly interoperate with EoMPLS deployed in other domains; and that EoMPLS should be used in the core networks while PBB-TE could be used in access networks and large campus networks.

5.2 Technology Briefing

5.2.1 Carrier Ethernet Area

For a better understanding of PBB-TE, it is useful to start with an overview of the Carrier Ethernet area to which PBB-TE belongs, including its structure.

Carrier Ethernet (or, more precisely, Carrier-Grade Ethernet) is an attempt to expand Ethernet beyond the borders of Local Area Networks (LANs) and into the territory of Wide Area Networks (WANs). The reasons for this expansion include the fact that Ethernet has become a de facto lower-layer network technology on LANs. Traditionally, Ethernet frames were repackaged into some other format for wide area transmission, only to be converted back to Ethernet at the destination. The idea therefore emerged of having Ethernet everywhere, and not only within LANs. The global availability of Ethernet-like services from service providers gives customers the ability to connect their geographically distributed networks on Layer 2 without the use of routing and routers. Such a way of connecting sites is, in some cases, very desirable because of its simplicity and independence from a provider's IP infrastructure.

There is another, very important reason for the Ethernet expansion: Ethernet interfaces are normally less expensive than other technologies, as their popularity has allowed mass production and thus cheaper unit prices. Previously, this benefit did not come without hidden costs: other technologies have far more sophisticated failure- and error-detection systems built in, which enable more accurate, and thus faster, tracing and rectifying of faults. The development of equivalent mechanisms for Ethernet is part of what prompted the Carrier Ethernet variant of the protocol.

The difference in cost of equipment is significant if Ethernet is compared with Synchronous Digital Hierarchy (SDH). SDH is the technology of choice, which has been used by communications carriers for building robust and high-speed WAN circuits for many years. It is feature-rich in detecting, signalling and recovering from failures. However, this comes at a high financial cost, and thus SDH could be the main target for replacement by Carrier Ethernet.

It is very useful to understand that there are actually two different areas under the Carrier Ethernet umbrella: one is Ethernet as an internal carrier transport technology, also known as Carrier Ethernet Transport (CET), and the other is Ethernet as a carrier service, also known as Ethernet over X, where X is a technology other than Ethernet carrier grade, which a provider uses internally but "wraps" into the Ethernet cover so that for end users it looks like Ethernet. While these areas obviously overlap – for example, when Carrier Ethernet Transport is used for providing Ethernet services for customers – understanding this difference helps to sort out many Carrier Ethernet-related issues.

Ethernet as a carrier service takes the customer's point of view, in that a carrier network has an Ethernet user network interface, which customers can use to connect their sites on Layer 2 as a Virtual Local Area Network (VLAN). Service providers do not take into account any IP or other Layer 3 protocol information about the customer's networks – a Carrier Ethernet link simply transports frames between the two end points, giving the illusion that they are directly connected.

However, different transport technologies can operate within the service provider network to underpin a global Ethernet service. Currently three options are available: Ethernet itself, Internet Protocol / Multi-Protocol Label Switching (IP/MPLS) and optical transport (SDH, Optical Transport Network (OTN), Dense Wavelength-Division Multiplexing (DWDM)). Respectively, three versions of Carrier Ethernet service are available: Carrier Ethernet Transport (CET), Ethernet over MPLS (EoMPLS) and Ethernet over Transport (EoT).

This trial investigated only packet-based versions of Carrier Ethernet, i.e. Carrier Ethernet Transport and Ethernet over MPLS.

5.2.2 Carrier Ethernet Transport (CET) and PBB-TE

Several strands of improvements aim to transform Ethernet into a carrier-grade transport technology, i.e. into Carrier Ethernet Transport. It is worth stressing that these target both the external (a service for customers) and internal (a reliable transport for connecting the provider's equipment) aspects of a provider's business. The main standards bodies working in the Ethernet improvement area are the IEEE, the ITU-T and, to a lesser extent, the IETF. The primary improvement strands are as follows:

- *De-coupling of provider and user networks.* If a provider network and all customer sites worked as a single LAN, the result would be hardly manageable. The IEEE has developed two standards: Provider Bridges (PB) [IEEE 802.1ad] and Provider Backbone Bridges (PBB) [IEEE 802.1ah]. 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. Both PB and PBB support point-to-point (Virtual Private Wire Service (VPWS)) and multipoint-to-multipoint (Virtual Private LAN Service (VPLS)-style) types of connectivity.
- *Traffic engineering, bandwidth guarantees and Quality of Service (QoS).* Several standards have been or are being developed to support these very desirable and interrelated features. One of the key standards here is PBB-TE (based on the Nortel proprietary Provider Backbone Transport (PBT) technology but standardised as IEEE 802.Qay), which adds support to PBB for deterministic paths and thus provides the feature set needed for providing bandwidth guarantees, resilience and robust QoS. PBB TE supports both point-to-point and point-to-multipoint connections [IEEE 802.1Qay].
- *Resilience.* The main tool for supporting resilience in LANs – the Spanning Tree Protocol (STP) – does not satisfy the requirements of the latest networks for fast switching to an alternative route when a failure occurs. Neither do the recent improvements in STP such as Rapid STP or Multiple STP. The IEEE has specified two standards in this area, which are based on different approaches: PBB-TE (also known as 802.1Qay), which uses a protection switching mechanism based on a deterministic backup path (in SDH style) and Shortest Path Bridging (SPB) (also known as 802.1aq), which uses link-state routing protocols for finding the proper topology (in a similar fashion to link-state IP routing protocols such as Open Shortest Path First (OSPF) and Intermediate System to Intermediate System (IS-IS)).
- *Operation, Administration and Maintenance (OAM)*. OAM has probably been (and still is) the main concern of providers who are looking at Ethernet as a carrier technology. Traditional Ethernet supports no OAM functionality at all; the recent standards 802.1ag and 802.3ah from the IEEE and Y.1731 from the ITU-T [ITU-T Y.1731] bridge this gap.

The above overview of the key features of different CET versions shows that of the three existing versions of CET (PB, PBB and PBB-TE), PBB-TE meets most requirements of a carrier-grade transport technology as described in [DJ1.1.1].

5.3 Trial Objective

The objectives of the PBB-TE trial were as follows:

- To gain practical experience with PBB-TE by trialling its core functionality in a multi-domain and multi-technology (PBB-TE and Ethernet over MPLS) environment.
- To investigate the manageability of PBB-TE within a multi-domain and multi-technology environment (path provisioning, troubleshooting).
- To evaluate the deployment of PBB-TE on JANET (for the core, for Regional Networks (RNs), for IP and Lightpath traffic) and to identify areas for further investigation.
- To evaluate the state of standardisation of PBB-TE.

5.4 Multi-Domain Testbed

To achieve the objectives of the trial, a multi-domain, multi-technology testbed was created. The testbed consisted of the participants' local testbeds, plus the core testbed, as shown in Figure 5.1. (The figure and the description that follows include participants/testbeds from the JANET CE project, which were shared with the GN3 project PBB-TE trial, namely Lancaster University, Manchester University/Net North West (NNW) and Oxford University/Oxford University Computing Services (OUCS).)

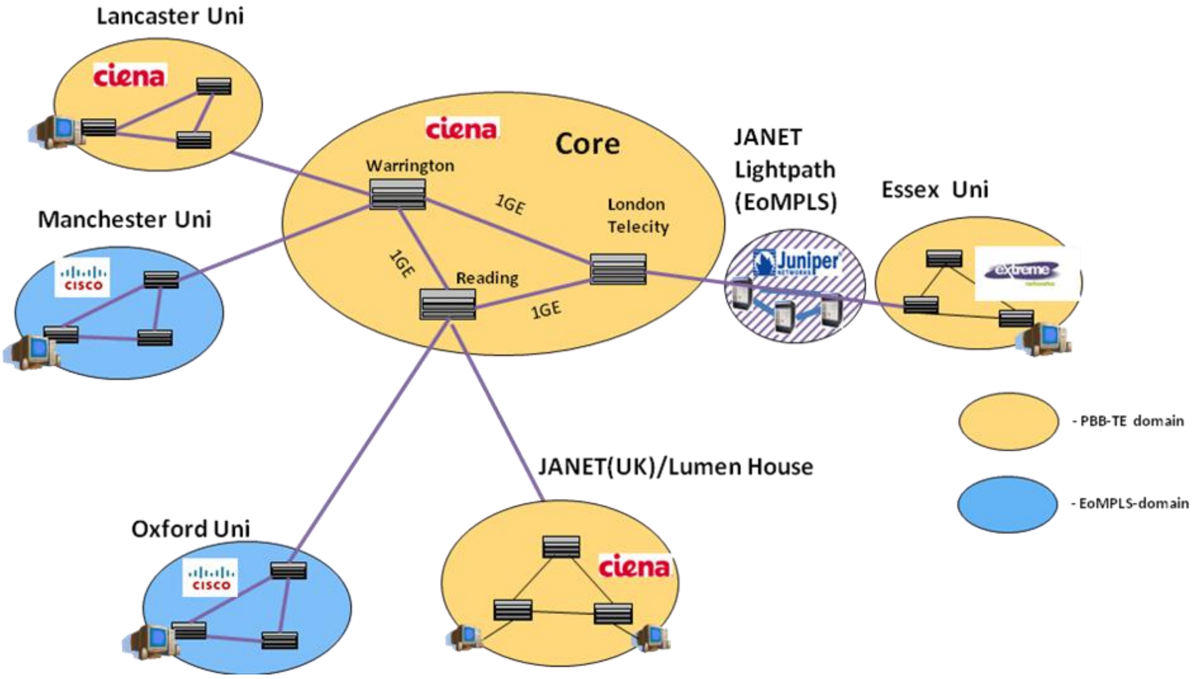


Figure 5.1: The multi-domain, multi-technology testbed

The core testbed was built using Ciena 5305 switches configured to support PBB-TE technology. The choice of technology for the core testbed resulted from the fact that the JANET Lightpath infrastructure has been using EoMPLS (along with wavelength/OTN) as a transport since March 2009 and hence some experience with EoMPLS behaviour in the core role had already been obtained by the start of the trial.

Two partners' testbeds, NNW and OUCS, were built using Cisco switches and configured to support EoMPLS.

Three partners' testbeds, Lumen House (LH), Lancaster University, and Essex University were built to support PB and PBB-TE, the former two on Ciena switches and the latter on Extreme switches.

1 Gbit/s Time-Division Multiplexing (TDM) links were used to connect the core testbed with the NNW, OUCS and LH testbeds. The JANET Lightpath connection was used to connect the core testbed with the Essex University testbed. This option was chosen because Essex University had already used a number of Lightpath connections before the start of the trial and it was relatively straightforward to extend the existing range of VLANs allocated to the university to serve the needs of the project.

5.5 PBB-TE Core Tests

This section covers the PBB-TE core tests:

- Network-to-network connectivity.
- Resilience.
- OAM investigation.
- Investigation of Ethernet Service Manager functionality.

It begins by describing the core testbed setup and configuration, explaining the naming conventions used and how the PBB-TE tunnels were configured.

5.5.1 Core Testbed Setup

The core of the testbed consisted of three Ciena 5305 switches connected by 1 Gbit/s links (established through OTN infrastructure, managed by Verizon for JANET).

The core diagram is shown in Figure 5.2.

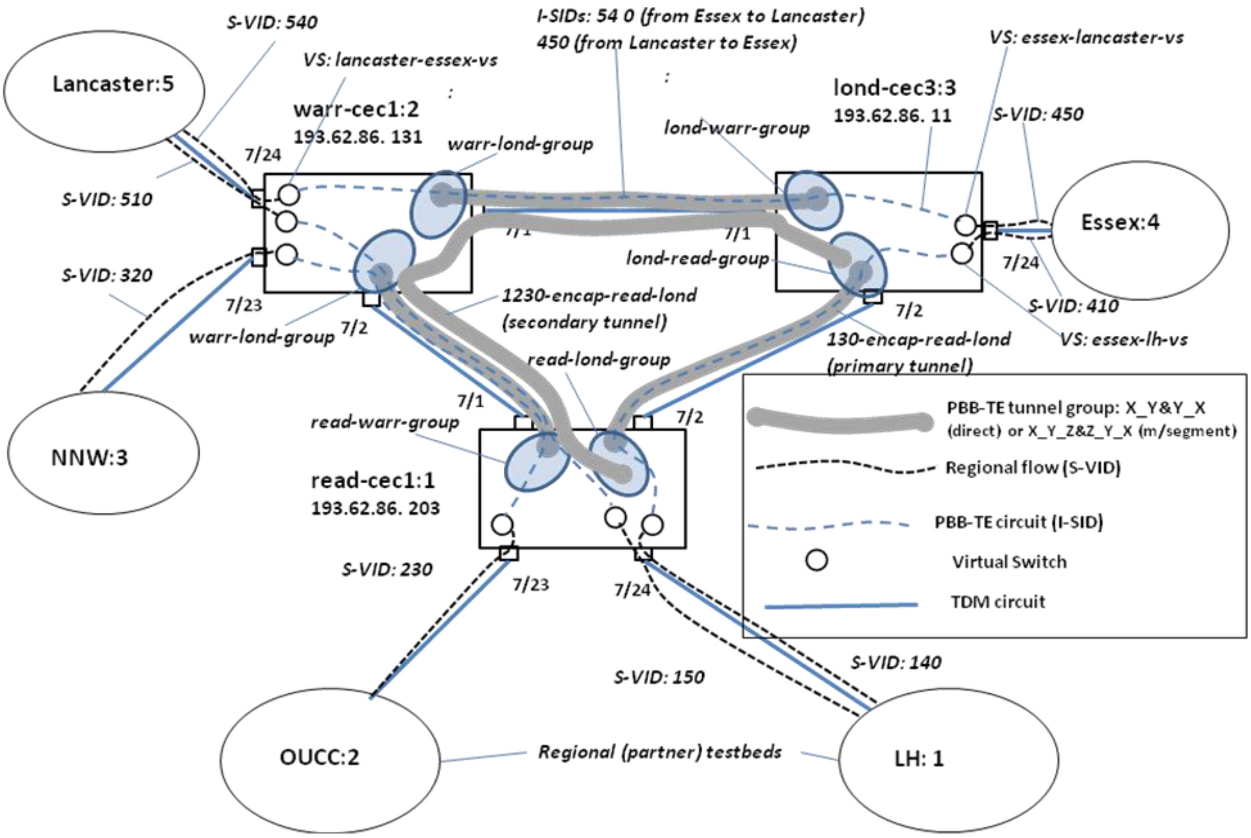


Figure 5.2: Core testbed

The core switches were located at JANET Core Points of Presence (C-PoPs) in London Telecity, Reading and Warrington.

The core switches were connected by 1 Gbit/s links to 5 project partner testbeds in Manchester (NNW), Oxford (OUCS), Essex (Essex University), Lancaster (Lancaster University) and Oxfordshire (Lumen House).

Links to NNW and Lancaster were similar to the core links as they were established by Verizon through OTN infrastructure.

Links to LH and OUCS were established through the Thames Valley Network (TVN) OTN and Layer 2 (MPLS) infrastructure.

The link to Essex University was established through the JANET Lightpath infrastructure. The Lightpath Juniper MX960 switch at Telecity used the VLAN range 401-499 to pass traffic from the core testbed towards the Essex University testbed.

5.5.2 Core Testbed Configuration

5.5.2.1 *Naming Convention*

The configuration elements of the core switches (PBB-TE tunnels, tunnel groups, sub-ports, etc.) used the following naming convention for switches and partner testbeds:

Element Type	Element	Name	Alternative
Switch	Reading	read-cec1	1
	Warrington	warr-cec1 (from warr-cec1.ja.net)	2
	London	lond (from lond-cec3.ja.net DNS name)	3 (when numbers were needed for creation of an element name)
Partner testbed	LH	1	-
	OUCS	2	-
	NNW	3	-
	Essex	4	-
	Lancaster	5	-

Table 5.1: Naming convention for switches and partner testbeds

5.5.2.2 *PBB-TE Tunnels*

Three pairs of PBB-TE tunnels performed the task of transferring traffic between partner testbeds through the core, and three pairs of PBB-TE tunnels were established to provide protected tunnelling. Their configuration can be explained by taking the read-cec1 to lond-cec3 tunnels as an example.

On the read-cec1 switch, read-lond group tunnels were established to provide protected connectivity to the lond-cec3 switch. For this tunnel group, two PBB-TE tunnels were established, as shown in Table 5.2. The names are the result of concatenating the switches' numeric and symbolic names, as explained further in the table.

Tunnel Name	Tunnel Type	Explanation of Name
130-encap-read-lond	A primary tunnel	A concatenation of the switch numerics 1 and 3, plus 0 as it is the first tunnel between switches 1 and 3, and a symbolic name.
1230-encap-read-lond	A secondary tunnel	The result of connecting switches 1 and 3 through an intermediate switch 2, and a symbolic name.

Table 5.2: PPE-TE tunnels for read-lond-group tunnels

As discussed below, the further Backbone Virtual LAN Identifier (B-VID) values of the tunnels share the same numeric part of the tunnel name; e.g. the 130-encap-read-lond tunnel has B-VID 130.

The role of tunnels within a tunnel group was determined by their weights: the heavier the tunnel, the higher the priority for it to be used for group connectivity. For this purpose, the 130-encap-read-lond tunnel received weight 8, while the 1230-read-lond tunnel received weight 1; the 130-encap-read-lond tunnel was therefore always used if it was up, while the 1230-read-lond tunnel was only used if 130-encap-read-lond was down.

The full configuration of the read-lond tunnel group was as follows:

```
pbt tunnel-group create group read-lond group tunnel-sync on logical-id 1
pbt encap-tunnel create static-encap 130-encap-read-lond dest-bridge-name lond-
      cec3 port 7/1 bvid 130 logical-id 1 tunnel-group read-lond-group pair-
      index 1 weight 8
pbt encap-tunnel create static-encap 1230-encap-read-lond dest-bridge-name
      lond-cec3 port 7/2 bvid 1230 logical-id 4 tunnel-group read-lond-group
      pair-index 4 weight 1
pbt decap-tunnel create static-decap 130-decap-read-read port 7/1 bvid 130
      logical-id 1 tunnel-group read-lond-group pair-index 1
pbt decap-tunnel create static-decap 1230-decap-read-lond port 7/2 bvid 1230
      logical-id 4 src-bridge-name read-cec1 tunnel-group read-lond-group pair-
      index 4
```

The tunnels use B-VID values.

The counterpart configuration of the primary tunnel on the read-cec1 switch looks similar and is therefore not presented here (the full configuration files of all the core switches can be found at [Configs]). However, the configuration of the secondary tunnel is specific, because it is not a direct tunnel between two adjacent switches. Instead, the secondary tunnel goes through an intermediate switch, warr-cec1 (which provides resilience in case of a direct link or interface 7/1 fault).

The configuration of warr-cec1 for supporting a secondary tunnel for the lond-warr-group uses a Ciena technique for transit PBB-TE tunnels:

```
pbt transit create pbt-transit TRANSIT-1230-702 parent-port 7/2 logical-id 3
pbt transit add pbt-transit TRANSIT-1230-702 class-element 1 bvid 1230
pbt transit create pbt-transit TRANSIT-1230-701 parent-port 7/1 logical-id 4
pbt transit add pbt-transit TRANSIT-1230-701 class-element 1 bvid 1230
```

Transit tunnels connect respective ports with respective B-VID values.

In Figure 5.2, only one secondary tunnel is shown – between lond-cec3 and read-cec1 – to ensure the readability of the diagram. The other two secondary tunnels were established in the same way.

5.5.3 Test 1: Network-to-Network Connectivity

5.5.3.1 *Test Objective*

The main function of the core testbed was to support connections between the partner testbeds. Thus, the objective of this test was to check whether the core was able to connect testbeds that use different technologies (e.g. EoMPLS and PBB-TE) and transfer their traffic transparently.

5.5.3.2 *Test Setup*

As a service delimiter common for both technologies, the outer VLAN ID value was chosen. In Figure 5.2 these VLAN IDs are shown as Service VLAN IDs (S-VIDs), in accordance with Ethernet terminology. It was the responsibility of the partners to configure their testbed so that the outer VLAN ID had a value assigned for specific connectivity through the core. The scheme chosen for the core testbed organisation complied with the Metro Ethernet Forum (MEF) External Network to Network Interface (ENNI) specification [MEF 26.1].

The VLAN IDs for the testbeds' connectivity were assigned according to the testbeds' numeric names, e.g. for Lancaster-Essex connectivity S-VID = 540, while for LH-Essex connectivity S-VID = 140.

It was important to test the capability of the core testbed to translate customer S-VIDs; this capability gives customers the freedom to use their own S-VID values and hence simplifies their interoperability. Thus, different S-VID values were assigned to each direction of a connection between the partner testbeds. For example, the Essex-LH connection had S-VID =410.

The core testbed translated the customer S-VID at egress, e.g. it accepted traffic from LH with S-VID = 140 and passed it towards Essex with S-VID = 410.

5.5.3.3 *Expected Result*

The expected result – and indicator of a successful outcome – was that all the partner testbeds pairs would be able to communicate through the core.

5.5.3.4 *Configuration*

The configuration of the core switches on selecting customer traffic and translating S-VIDs was based on the Ciena sub-ports technique. This can be explained by taking the lond-cec3 switch as an example.

To support a connection between Essex and Lancaster, the following configuration was used:

```
sub-port create sub-port essex-lancaster-subport parent-port 7/24 classifier-
      precedence 2 logical-id 2 egress-l2-transform stamp-*.450.*
sub-port add sub-port essex-lancaster-subport class-element 1 vtag-stack 450
```

The first statement creates a sub-port essex-lancaster-subport for a physical port 7/24, which is connected to the Essex link. The statement instructs the switch to put S-VID 450 into the outer VLAN tag (the first asterisk leaves the EtherType of a frame intact while the second one does the same with the inner VLAN ID – if it exists). This means that the frames going from Lancaster to Essex will have S-VID = 450 despite their original value.

The second statement adds a selector to the sub-port essex-lancaster-subport, which tells the sub-port that traffic with S-VID = 450 should be selected at the physical interface 7/24 for this logical sub-port.

To transfer traffic selected by a sub-port through a PBB-TE tunnel it was necessary to create a customer Information Service Identifier (I-SID) connection (analogous to an MPLS pseudowire) for this PBB-TE tunnel and a virtual switch to connect the sub-port and the I-SID connection. This was achieved by the following statements:

```
pbt service create service essex-lancaster-conn-1 ingress-isid 450 egress-isid
      540 logical-id 2 tunnel-group lond-warr-group

virtual-switch create vs essex-lancaster-vs logical-id 2

virtual-switch interface attach sub-port essex-lancaster-subport vs essex-
      lancaster-vs

virtual-switch interface attach pbt-service essex-lancaster-conn-1 vs essex-
      lancaster-vs
```

5.5.3.5 *Results*

The test results were positive, that is, the core testbed successfully managed to connect the partners' testbeds for all the tests conducted by the project partners over the lifetime of the PBB-TE trial (the results of these successful tests are described in further sections).

The following partner testbed pairs were connected:

- LH-Lancaster.
- LH-Essex.
- Lancaster-Essex.
- OUCS-NNW.

The core testbed passed partners' traffic in a transparent way, including service frames; no problems with filtering of any kind of traffic were detected.

The first three connections were a PB/PBB-TE/PB type, which means that partner frames arrived at the core ingress interfaces (i.e. ENNIs) as frames in PB (QinQ) encapsulation with a proper S-VID in the outer VLAN tag. Then they were encapsulated into PBB-TE frames and transferred through the core in that format. At the core egress interfaces, the PBB-TE headers were stripped off and frames were sent towards the destination testbed in the original PB format with the S-VID forcibly changed into the value assigned for the destination testbed.

The last (OUCS-NNW) connection was an EoMPLS/PBB-TE/EoMPLS type. The job of the core switches on transferring EoMPLS frames through the core was similar to the PB/PBB-TE/PB case. As long as both OUCS and NNW testbeds produced EoMPLS frames with proper S-VID values in the outer VLAN tag, the core switches managed to select those frames and direct them into their respective I-SID core connections. The existence of two MPLS shim headers (for an MPLS tunnel and a pseudowire) was transparent to the core switches as they paid attention only to the outer VLAN tags at the edge and the PBB-TE headers on the transit interfaces.

5.5.3.6 *Problems Encountered*

The following problems were encountered during the test:

- Frame drops were detected during a LH-Essex Ultra High Definition (UHD) video test. An investigation showed the reason to be Ethernet frame sizes exceeding 1518 bytes. This happened because of the extra header fields of PBB-TE framing in the core. The effect became apparent when the UHD codec sent frames with 1400 bytes of Real-Time Protocol (RTP) payload, which resulted in larger than standard Ethernet frame sizes. To fix this, the Maximum Transmission Unit (MTU) value of the core switches was increased to accommodate up to 3000 byte frames. This MTU value turned out to be large enough for all connections and tests, as no frame drops due to exceeding the MTU size were detected afterwards.
- The core Ciena 5305 switches could not select traffic from Essex University properly. At the same time, the core switches *could* properly select incoming traffic from *other* partners’ testbeds. The reason was an EtherType value used by the Extreme switches in the Essex University testbed: it used EtherType 0x88a8 in the outer VLAN tag (S-VID tag), while all other equipment (Cisco 6500 and Ciena LE-311v) generating traffic towards the core used EtherType 0x8100 in the outer S-VID tag. Actually, the value 0x88a8 is a standard for S-VID tagging according to the IEEE 802.1ad specification and the Ciena 5305 switches should have recognised it – but failed to do so. A Ciena specialist said that this was a bug in the switch software, and it was fixed in the subsequent software release. The problem was solved by translating the EtherType value from 0x88a8 to 0x8100 using an intermediate Juniper switch (not shown in Figure 5.2), which provided a transparent connection between the lond-cec3 switch and Essex University’s Extreme switch. After that, traffic injected into the core by the Essex University testbed was selected and transferred through the core without any problems.

5.5.4 Test 2: Resilience

5.5.4.1 *Test Objective*

The objective was to test the capability of the core switches to protect tunnel groups by switching from a primary group tunnel to a secondary one when the primary tunnel went down (e.g. because of port or link failure).

5.5.4.2 *Test Description*

The test consisted of the following steps:

1. Establish a ping session (in both directions) between LH and Lancaster testbeds with a ping interval of 100 ms.
2. After some time, manually deactivate port 7/1 of the read-cec1 switch (using the Command Line Interface (CLI)).
3. Several seconds later, stop the ping sessions to check the number of lost replies.
4. Record the states of the read-warr group tunnels.
5. Resume the ping sessions and reactivate port 7/1.
6. Repeat this procedure, again recording the number of lost replies and the state of the tunnels.

5.5.4.3 *Configuration*

The configuration of protected tunnel groups and Continuity Check Message (CCM) sessions that monitored the tunnel states was as described in Section 5.5.4.2.

5.5.4.4 *Results*

The results were recorded from both sides – at LH and Lancaster. The results showed that the protection switching happened correctly in all cases (the test was repeated three times) and within the expected interval of 300 ms (three times the CCM interval, as the loss of three consecutive CCMs triggers protection switching). The number of lost pings during switching from a primary to a secondary tunnel was between 0 and 2.

When the state of port 7/1 was restored, the reverse protection switching happened – with a delay of about 10 seconds after restoring the port state. The number of lost pings was the same as for the direct switching – between 0 and 2. After consulting with Ciena, the reason for the delay in reverse switching was identified to be a design feature intended to prevent the network oscillating between tunnels when a port is experiencing erratic connectivity problems.

5.5.4.5 *Problems Encountered*

No problems were encountered.

5.5.5 Test 3: OAM Investigation

5.5.5.1 *Test Objective*

The objective was to test Ethernet OAM’s hierarchical capability to monitor the health of the core tunnels without interfering with customer traffic and customer OAM sessions. The core tunnel monitoring was carried out by means of the 802.1ag/Y.1731 [IEEE 802.1ag, ITU-T Y.1731] CCM mechanism supported by Ciena 5305 boxes.

5.5.5.2 *Test Description*

The test consisted of the following steps:

1. Establish CCM sessions for all primary and secondary tunnels of all tunnel groups connecting the core switches.
2. For each tunnel, create two Maintenance Entity Group End Points (MEPs), one at each end of the tunnel.
3. Each MEP continuously sent CCM frames towards its far-end counterpart with a 100 ms interval. To prevent potential interference with customer CCMs, the core CCM sessions worked at Level 3, leaving Levels 4 and 5 for partners’ customer traffic.
4. Use a number of Ciena 5305 CLI Show commands to manually check the health of the core tunnels.
5. The CCM mechanism was also configured and used in the core resilience tests (described in Section 5.5.4) as the main mechanism for triggering protection switching (i.e. switching between the primary and secondary tunnels of a tunnel group).

As Ciena 5305 switches do not support MEG Intermediate Points (MIPs) for PBT tunnels, this element of the Connectivity Fault Management (CFM) standard could not be tested.

5.5.5.3 *Configuration*

The following configuration example is taken from the lond-cec3 switch configuration file; it illustrates the technique used for monitoring the lond-read tunnel group.

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! CFM GLOBAL CONFIG:
!
cfm enable
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! CFM SERVICE CONFIG:
!
```

```
cfm service create static-encap 130-encap-lond-read name cfm-130-read-lond
      next-mepid 311 ccm-interval 100msecCCM logical-id 1
cfm service enable service cfm-130-read-lond
cfm service create static-encap 1230-encap-lond-read name CFM-1230 next-mepid
      2101 ccm-interval 100msecCCM logical-id 4
cfm service enable service CFM-1230
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

CCM Level 3 is the default level for Ciena 5305 CCM sessions, which is why it is omitted in the example configuration statements.

5.5.5.4 *Results*

All the results of the Ethernet CCM tests were positive. The sessions between MEPs were established without problems and showed the state of the core tunnels correctly. For example, the `cfm rem show` command at the lond-cec3 switch produced the following output:

```
lond-cec3> cfm rem show

+--------------------------- CFM REMOTE MEPS ----------------------------------+
|                        |Mep |                 |State|             |Total
|Service                 |ID  |Mac Address      |Ad|Op| Fault(s)    | RX
+------------------------+----+-----------------+--+--+-------------+----------
|cfm-130-read-lond       |131 |02:00:00:10:00:01|En|En|             |179381712|
|cfm-230-lond-warr       |2300|02:00:00:10:00:02|En|En|             |167539145|
|PBT-1212                |1212|02:00:00:10:00:02|En|En|             |203524943|
|CFM-1230                |135 |02:00:00:10:00:01|En|En|             |166267693|
+------------------------+----+-----------------+--+--+-------------+---------+
```

The `pbt encap-tunnel show` command at the lond-cec3 switch showed the following output when all the tunnels were in an operational and administrative Up state:

```
lond-cec3> pbt encap-tunnel show

+------------------------ ENCAP TUNNEL TABLE ----------------------------------+
|                               |   State    |     |  Port   |Group|Pair |W|CFM/
|Name                           |Op |Adm|Fwd|B-VID|  Name  |Index|Index|t|y1731|
+-------------------------------+---+---+---+-----+--------+-----+-----+-+-----+
|130-encap-lond-read            |En |En |En |130  |7/1     |1    |1    |8| Y/N|
|230-encap-lond-warr            |En |En |En |230  |7/2     |2    |1    |8| Y/N|
|U_warr-cec1                    |En |En |En |1212 |7/2     |3    |1    |6| Y/N|
|1230-encap-lond-read           |En |En |Dis|1230 |7/2     |1    |4    |1| Y/N|
+-------------------------------+---+---+---+-----+--------+-----+-----+-+-----+
```

When there was a problem with a tunnel, the output was as follows:

```
lond-cec3> cfm rem show

+--------------------------- CFM REMOTE MEPS ---------------------------------+
|                        |Mep |                 |State|             |Total    |
|Service                 |ID  |Mac Address      |Ad|Op| Fault(s)    | RX      |
+------------------------+----+-----------------+--+--+-------------+---------+
|cfm-130-read-lond       |131 |02:00:00:10:00:01|En|En|             |179386189|
|cfm-230-lond-warr       |2300|02:00:00:10:00:02|En|Di|rMep         |167543564|
|PBT-1212                |1212|02:00:00:10:00:02|En|En|             |203529420|
|CFM-1230                |135 |02:00:00:10:00:01|En|En|             |166272170|
+------------------------+----+-----------------+--+--+-------------+---------+

lond-cec3> pbt encap-tunnel show

+------------------------ ENCAP TUNNEL TABLE ---------------------------------+
|                              |   State    |     |  Port   |Group|Pair |W|CFM/|
|Name                          |Op |Adm|Fwd|B-VID|  Name   Index|Index|t|y1731|
+------------------------------+---+---+---+-----+--------+-----+-----+-+-----+
|130-encap-lond-read           |En |En |En |130  |7/1     |1    |1    |8| Y/N|
|230-encap-lond-warr           |Dis|En |Dis|230  |7/2     |2    |1    |8| Y/N|
|U_warr-cec1                   |En |En |En |1212 |7/2     |3    |1    |6| Y/N|
|1230-encap-lond-read          |En |En |Dis|1230 |7/2     |1    |4    |1| Y/N|
+------------------------------+---+---+---+-----+--------+-----+-----+-+-----+
```

The output above corresponds to a remote MEP at warr-cec1 being down. The session cfm-230-lond-warr shows an rMep fault, which changes the 230-encap-lond-warr status to read Dis for disabled (this is specific to Ciena 5305 switches – a fault in the CCM session caused the respective tunnels to be reported as down).

Further information about Ethernet OAM, using the IEEE 802.1ag / Y.1731 functionality of network equipment, is provided in Chapter 3.

5.5.5.5 *Problems Encountered*

No problems were encountered.

5.5.6 Test 4: Investigation of Ethernet Service Manager Functionality

5.5.6.1 *Background*

The ability to manage and provision PBB-TE switches with a management system is very important since, in reality, PBB-TE is a zero control plane technology (some Internet drafts describing GMPLS extensions for PBB-

TE have been produced but none has been implemented in real PBB-TE equipment). Consequently, network managers have two possibilities for managing PBB-TE networks:

1. Manual provisioning and management by accessing each device with telnet or SSH.
2. Automated provisioning and management using some kind of Network Management System (NMS).

Ethernet Service Manager (ESM) version 5.4 from Ciena, an NMS with provisioning functionality, was selected for the trial.

5.5.6.2 *Test Objective*

The objective was to test whether ESM version 5.4 could manage the core and LH testbeds.

5.5.6.3 *Results*

The main results of this trial were:

- ESM can discover 5305 and LE-311v switches in automated mode and visualise them on the network hierarchical map (see Figure 5.3).

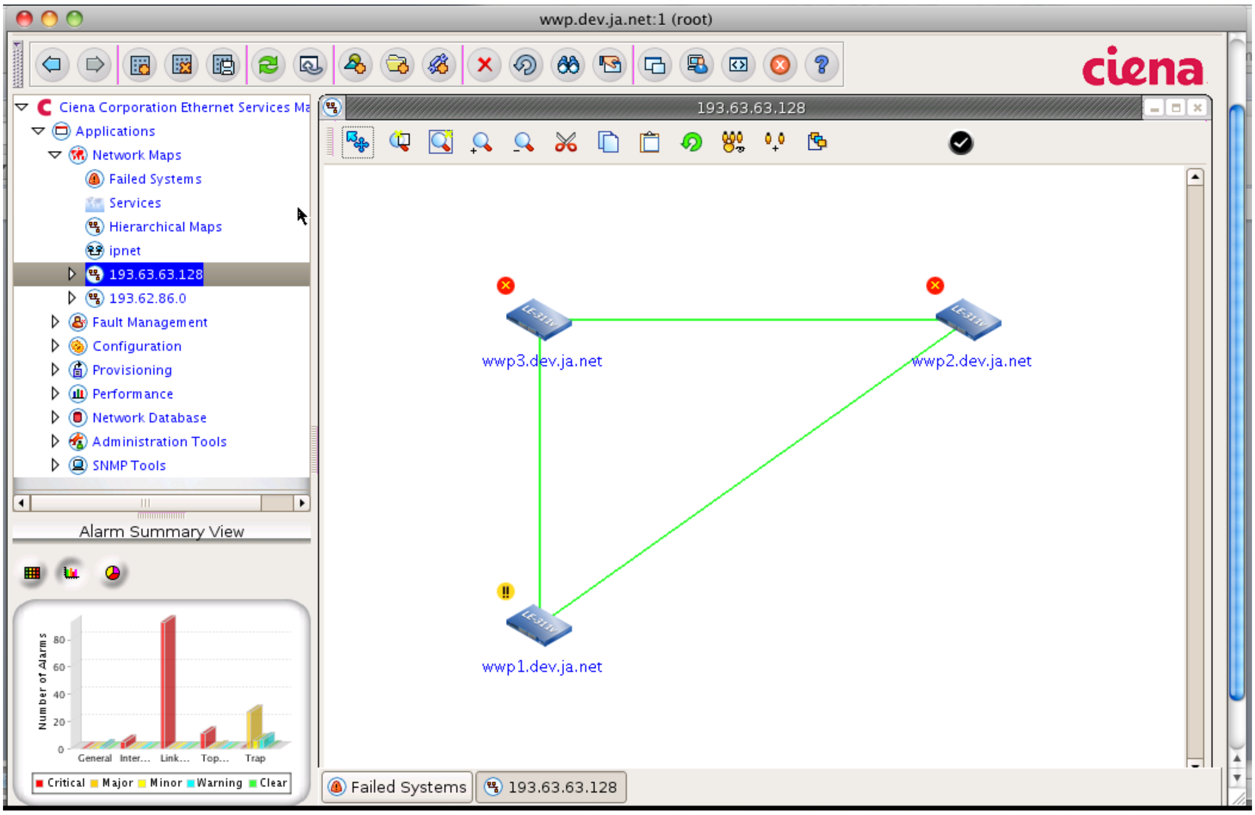
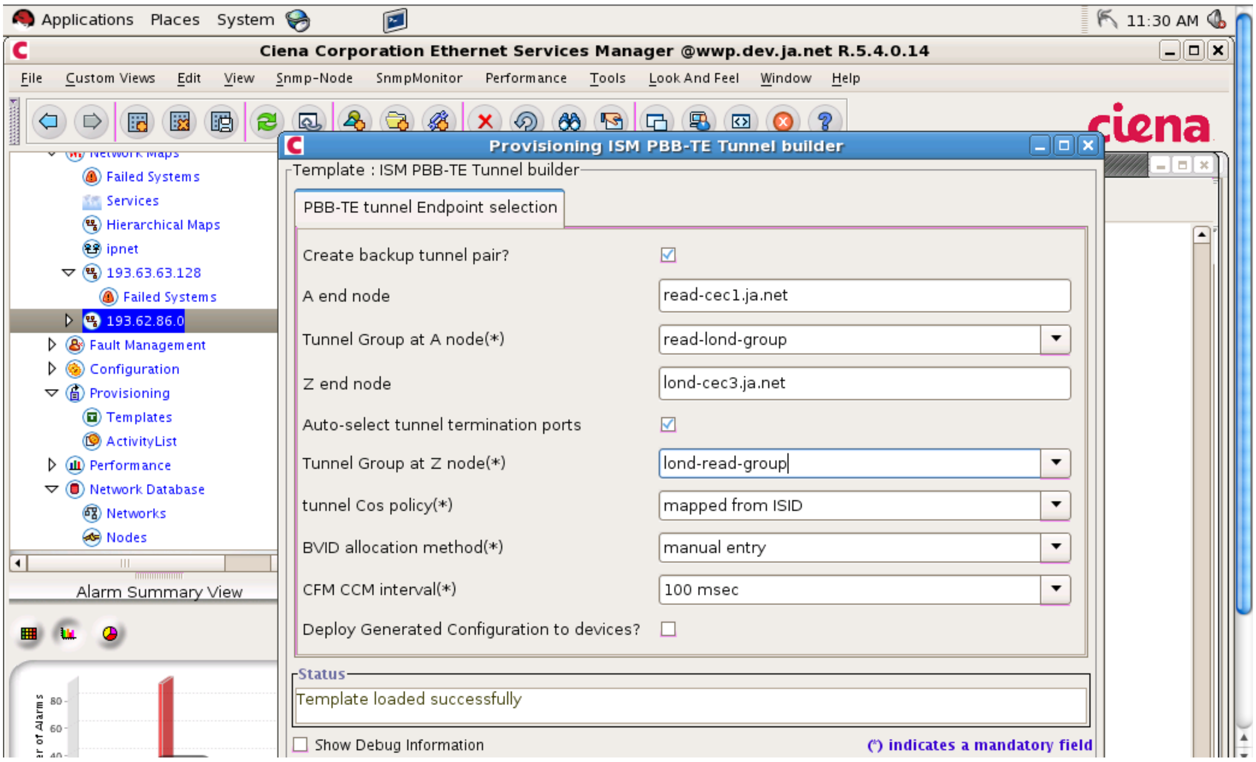


Figure 5.3: ESM screenshot – map of the World Wide Packets (WWP) switches

- Event and alarms generated by the testbed switches were properly registered, stored and visualised by ESM. However, the details of events and alarms were presented in a way that was not user-friendly and understandable for a network administrator.
- ESM supported automated provisioning of PBB-TE unprotected and protected tunnels (i.e. tunnel resilience) by means of a provisioning script. The provisioning script could accept the choice of tunnel end points as a result of the administrator clicking on the network map (see Figure 5.4).
- ESM supported automated provisioning of customer I-SID connections through existing PBB-TE tunnels. In the case when a tunnel between chosen end points did not exist, the connection provisioning script invoked the tunnel provisioning script, which provisioned the required tunnel.
- ESM can deploy the configuration information produced by a provisioning script by using telnet or SSH access to network nodes.
- Both ESM connection and tunnel provisioning scripts use a convenient method of automation, which allows administrator intervention before deploying the configuration information: the information can be checked by an administrator who can view it in familiar CLI form before confirming/denying its deployment on switches. Alternatively, the administrator can allow deployment without an intermediate check if he has enough experience of using the scripts and hence trusts them.
- Unfortunately, the versions of the tunnel provisioning script used during the project did not support the traffic engineering features of PBB-TE. In fact, the provisioning script could only choose the path that followed the usual Interior Gateway Protocol (IGP) route. Manual intervention in the path calculation, e.g. by specifying explicit and implicit routes, was also not supported. However, traffic engineering support in ESM was promised by Ciena for future versions of ESM, which were beyond the PBB-TE trial lifetime.



Template : ISM PBB-TE Tunnel builder

PBB-TE tunnel Endpoint selection

Create backup tunnel pair?	☑
A end node	read-cec1.ja.net
Tunnel Group at A node(*)	read-lond-group
Z end node	lond-cec3.ja.net
Auto-select tunnel termination ports	☑
Tunnel Group at Z node(*)	lond-read-group
tunnel Cos policy(*)	mapped from ISID
BVID allocation method(*)	manual entry
CFM CCM interval(*)	100 msec
Deploy Generated Configuration to devices?	□

Status

Template loaded successfully

□ Show Debug Information

(*) indicates a mandatory field

Figure 5.4: ESM Screenshot - Initial Form of Tunnel Provisioning Script

Below is an example of the tunnel provisioning script configuration produced for an unprotected tunnel starting at read-cec1 and terminating at lond-cec3:

```
pbt tunnel-group create group U_lond-cec3 tunnel-sync on
port set port 7/1 max-frame-size 2000
pbt encap-tunnel create static-encap U_lond-cec3 tunnel-group U_lond-cec3 pair-
      index 1 port 7/1 bvid 3224 dest-bridge-name lond-cec3 weight 6
pbt decap-tunnel create static-decap U_lond-cec3 tunnel-group U_lond-cec3 pair-
      index 1 port 7/1 bvid 3224 src-bridge-name read-cec1
cfm service create static-encap U_lond-cec3 name PBT-3224 next-mepid 1 level 2
      ccm-interval 100ms
cfm service enable service PBT-3224
```

A CFM service was created for this tunnel as an option chosen by an administrator.

Below is an example of the connection provisioning script configuration produced for a connection starting at lond-cec3 with Customer VLAN ID (C-VID) 777 and S-VID 150 going towards warr-cec1 using tunnel group lond-warr-group:

```
! create service CVID_00007308 on CN 5305 lond-cec3.ja.net
rstp disable port 7/3
aggregation set port 7/3 agg-mode manual
```

```
lldp set port 7/3 mode rx-only notification off
virtual-switch create vs CVID_00007308
! add CVID 777 on Sub port 7/3 to CVID_00007308
sub-port create sub-port CVID_00007308 parent-port 7/3 classifier-precedence
      777 ingress-l2-transform push-*.150.map egress-l2-transform pop
sub-port add sub-port CVID_00007308 class-element 1 vtag-stack 777
virtual-switch interface attach sub-port CVID_00007308 vs CVID_00007308
cpu-interface sub-interface create cpu-subinterface CVID_00007308 cpu-egress-
      l2-transform push-8100.777.7:push-8100.150.7
virtual-switch interface attach cpu-subinterface CVID_00007308 vs CVID_00007308
cfm service create vs CVID_00007308 name CVID_00007308 next 1 level 4
cfm service set service CVID_00007308 alarm-priority 1
cfm service set service CVID_00007308 alarm-time 10
cfm service set service CVID_00007308 reset-time 3000
cfm mep create service CVID_00007308 sub-port CVID_00007308 type up mepid 1
cfm service enable service CVID_00007308
cfm service set service CVID_00007308 ccm-interval 1s
! add PBB-TE Service CVID_00007308 to CVID_00007308
pbt service create service CVID_00007308 ingress-isid 205278 egress-isid 205278
      tunnel-group lond-warr-group
virtual-switch interface attach pbt-service CVID_00007308 vs CVID_00007308
```

Some of the configurations produced by the ESM provisioning scripts were successfully deployed in the core 5305 switches and used for transferring project partners’ traffic.

5.5.6.4 *Problems Encountered*

No problems were encountered.

5.6 Essex University Local PBB-TE Tests

This section covers the PBB-TE tests carried out by Essex University:

- PBT[1] connectivity.
- PBT resilience.
- CFM and PBT protection and resilience.

It begins by describing the Essex testbed setup, PB and PBB configurations, and configuration checks.

[1] Ciena technical documentation uses the Nortel proprietary acronym PBT instead of PBB-TE, which is the IEEE equivalent.

5.6.1 Essex Testbed Setup

The Essex University testbed for running Carrier Ethernet consisted of three Extreme BlackDiamond® (BD) 12k switches, which support PB, PBB and PBT protocols, meshed with 10 GE links. While the BD 12k boxes formed the core of the testbed, two Foundry Networks / Brocade FastIron switches were used on the edges of the core (connected to two of the Extreme boxes via 1G links), so they provided the user aggregation points.

This testbed was used for carrying out a variety of experiments such as connectivity tests, traffic engineering tests, video applications and so on, both on a local scale and in connection with other sites.

The network setup is shown in Figure 5.5 and Figure 5.6.

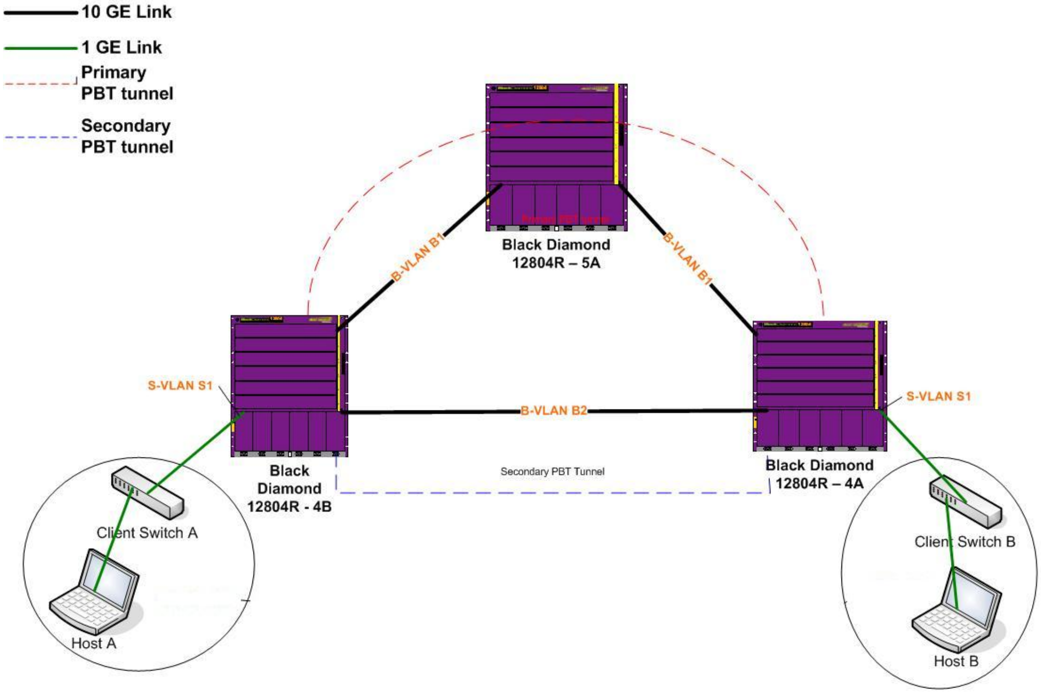


Figure 5.5: Local Essex testbed

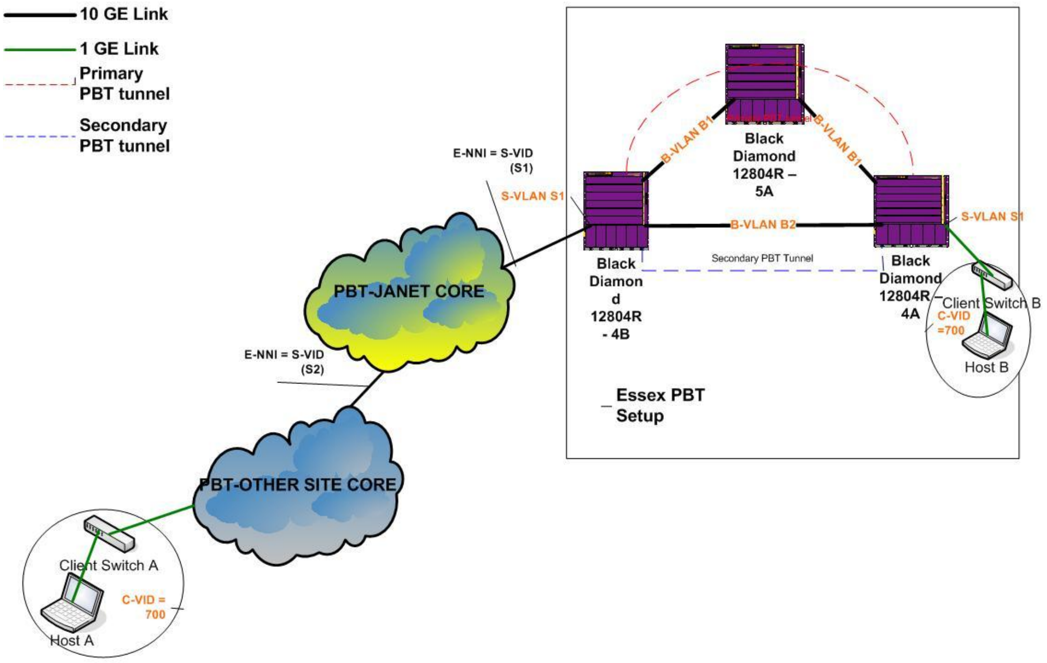


Figure 5.6: Essex testbed connected to other sites through JANET core

5.6.2 PB and PBB Configurations

For deploying PB connectivity, the port on the Extreme box was connected to the access switch using an untagged Service VLAN (S-VLAN). The VLAN tagged traffic was directed from the untagged S-VLAN to the tagged S-VLAN with an extended VLAN tag.

PB connection was setup both locally and with Lumen House.

5.6.2.1 *PB Configurations*

The three switches at Essex were configured to support provider bridging as follows:

Switch 4A:

```
creat svlan sjan
conf sjan tag 410
conf sjan add port 2:10 untagged
conf sjan add port 2:1 tagged
```

Switch 5A:

```
creat svlan sjan
```

```
conf sjan tag 410
conf sjan add port 5:1 tagged
conf sjan add port 2:1 tagged
```

Switch 4B:

```
creat svlan sjan
conf sjan tag 410
conf sjan add port 2:10 untagged
conf sjan add port 2:1 tagged
```

5.6.2.2 *PBB Configurations*

For local testing of PBB, Backbone VLANs (B-VLANs) encapsulated S-VLAN streams to transfer them over the CET core network.

The three switches were set up to run PBB between them as shown in Figure 5.5 and Figure 5.6. Switch 5A as the core had ports (2:1, 5:2) tagged connected to 4A and 4B. Switch 4B had access untagged ports (1:8) and network ports (5:2, 5:1). Switch 4A had network ports (2:1, 5:1) tagged, and customer port (2:10) untagged. Switches 4A and 4B were connected directly through port 5:1 on each to provide a secondary path for traffic delivery.

The switches were configured to provide PBB as follows:

BD4A:

```
create bvlan bjan
conf bjan tag 1410
conf bjan add port 2:1 tagged
creat svlan sjan
conf sjan tag 410
conf sjan add port 2:10 untagged
conf bvlan bjan add svlan sjan
```

BD5A:

```
create bvlan bjan
conf bvlan bjan tag 1410
conf bvlan bjan add port 2:1, 5:2 tagged
```

BD4B:

```
create bvlan bjan
conf bvlan bjan tag 1410
conf bvlan bjan add port 5:2 tagged
```

```
creat svlan sjan
conf svlan sjan tag 410
conf svlan sjan add port 1:8 untagged
conf bvlan bjan add svlan sjan
```

5.6.2.3 *Configuration Check*

The configurations were checked (using a `show the bvlan` command), resulting in the following output:

5A.2 # show the bvlan

```
BVLAN Interface with name bjan created by user
       Admin State:    Enabled         Tagging:        802.1Q Tag 1410
        Virtual router: VR-Default
        IPv6:           None
        STPD:           None
        Protocol:       Match all unfiltered protocols
        Loopback:       Disabled
        NetLogin:       Disabled
        QosProfile:     QP1
        Egress Rate Limit Designated Port: None configured
        Dot1ah Mode:    Backbone
        Service Count:  0
        Ports:   2.        (Number of active ports=2)
           Tag:      *2:1,   *5:2
        Flags:    (*) Active, (!) Disabled, (g) Load Sharing port
                  (b) Port blocked on the vlan, (m) Mac-Based port
                  (a) Egress traffic allowed for NetLogin
                  (u) Egress traffic unallowed for NetLogin
                  (t) Translate VLAN tag for Private-VLAN
                  (s) Private-VLAN System Port, (L) Loopback port
                  (e) Private-VLAN End Point Port
```

4A.1 # show the bvlan

```
BVLAN Interface with name bjan created by user
        Admin State:    Enabled         Tagging:        802.1Q Tag 1410
        Virtual router: VR-Default
        IPv6:           None
        STPD:           None
        Protocol:       Match all unfiltered protocols
        Loopback:       Disabled
        NetLogin:       Disabled
        QosProfile:     QP1
        Egress Rate Limit Designated Port: None configured
```

```
           Dot1ah Mode:    Backbone
           Service Count:  1
                           Service Name      VID   I-ISID
                           ================================
                           sjan              410
           Ports:   1.       (Number of active ports=1)
              Tag:      *2:1
           Flags:    (*) Active, (!) Disabled, (g) Load Sharing port
                     (b) Port blocked on the vlan, (m) Mac-Based port
                     (a) Egress traffic allowed for NetLogin
                     (u) Egress traffic unallowed for NetLogin
                     (t) Translate VLAN tag for Private-VLAN
                     (s) Private-VLAN System Port, (L) Loopback port
                     (e) Private-VLAN End Point Port
```

4B.1 # show the bvlan

```
   BVLAN Interface with name bjan created by user
           Admin State:    Enabled         Tagging:        802.1Q Tag 1410
           Virtual router: VR-Default
           IPv6:           None
           STPD:           None
           Protocol:       Match all unfiltered protocols
           Loopback:       Disabled
           NetLogin:       Disabled
           QosProfile:     QP1
           Egress Rate Limit Designated Port: None configured
           Dot1ah Mode:    Backbone
           Service Count:  1
                           Service Name      VID   I-ISID
                           ================================
                           sjan              410
           Ports:   1.       (Number of active ports=1)
              Tag:      *5:2
           Flags:    (*) Active, (!) Disabled, (g) Load Sharing port
                     (b) Port blocked on the vlan, (m) Mac-Based port
                     (a) Egress traffic allowed for NetLogin
                     (u) Egress traffic unallowed for NetLogin
                     (t) Translate VLAN tag for Private-VLAN
                     (s) Private-VLAN System Port, (L) Loopback port
                     (e) Private-VLAN End Point Port
```

5.6.3 Test 1: PBT Connectivity

5.6.3.1 *Testbed Setup*

The PBT connection between Essex University and LH was set up for compatibility and performance tests. This connection was used for HD/UHD video streaming between Essex and LH.

The testbed remained the same as in the PBB test bed. For the PBT setting, learning and flooding of the core switches were disabled and permanent forwarding database (FDB) entries as MAC tables were defined for edge-to-edge communication over the core of the network.

5.6.3.2 *Configurations*

The PBT setup followed the PBB setups. However, flooding and learning on the switches were disabled, and FDB entries had to be made manually.

The PBT configuration (bvlan-svlan, FDB entries and disabling flooding and learning) of each of the switches was as follows:

Switch 4A:

```
create bvlan bjan
conf bjan tag 1410
conf bjan add port 2:1 tagged
create svlan sjan
conf svlan sjan tag 410
conf svlan sjan add port 2:10 untagged
conf bvlan bjan add svlan sjan

create fdb 00:04:96:3b:23:10 bvlan bjan port 2:1
```

Switch 5A:

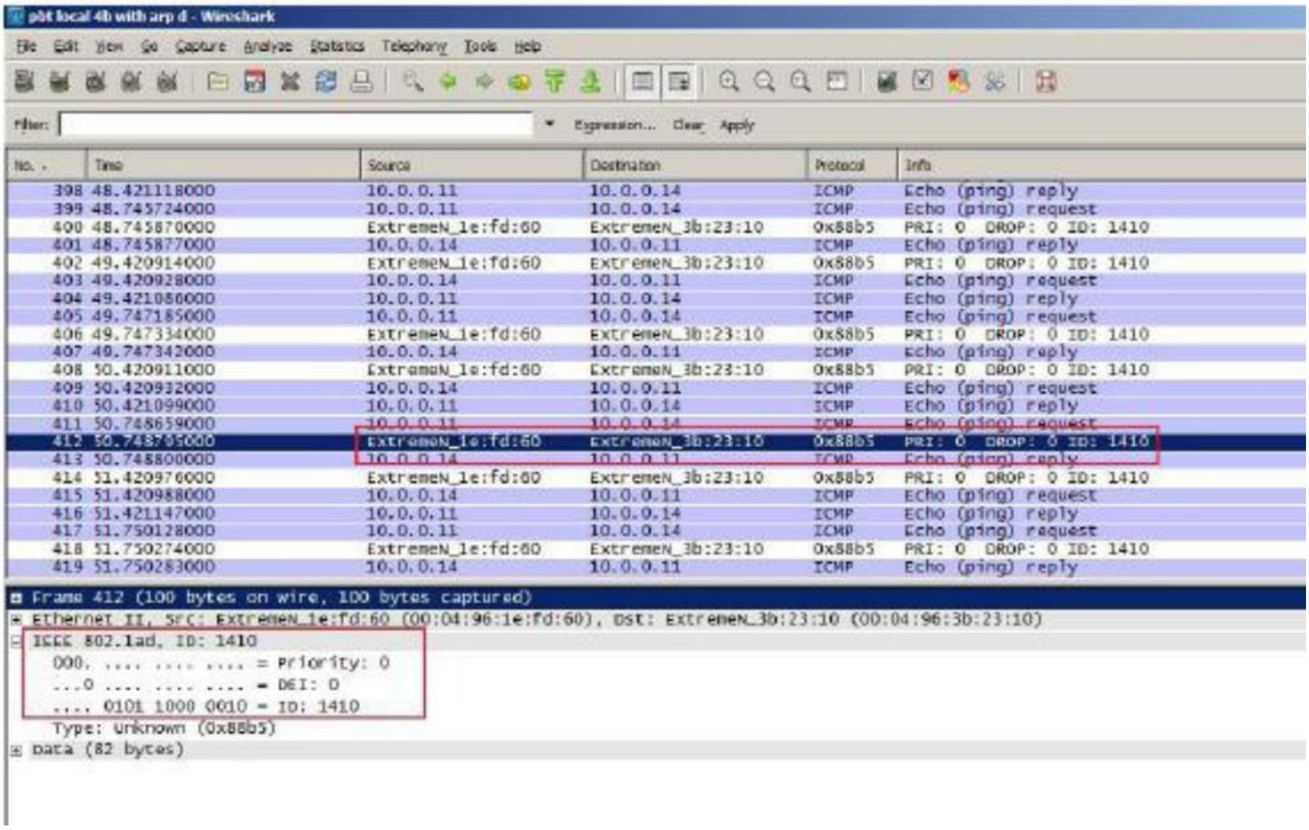
```
create bvlan bjan
conf bvlan bjan tag 1410
conf bvlan bjan add port 2:1, 5:2 tagged

create fdb 00:04:96:3b:23:10 bvlan bjan port 5:2
create fdb 00:04:96:1e:fd:60 bvlan bjan port 2:1
disable flooding bjan
disable learning bjan
```

Switch 4B:

```
create bvlan bjan
conf bvlan bjan tag 1410
conf bvlan bjan add port 5:2 tagged
create svlan sjan
conf svlan sjan tag 410
conf svlan sjan add port 1:8 untagged
conf bvlan bjan add svlan sjan
create fdb 00:04:96:1e:fd:60  bvlan bjan port 5:2
```

5.6.3.3 *Results*

The screenshot in Figure 5.7 below shows the captured packets in the connectivity test.


pbt local 4b with arp d - Wireshark
File Edit View Go Capture Analyze Statistics Telephony Tools Help
Filter:
Expression... Clear Apply

```
No. .  Time            Source               Destination          Protocol  Info
   398 48.421118000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) reply
   399 48.745724000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) request
   400 48.745870000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   401 48.745877000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) reply
   402 49.420914000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   403 49.420928000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) request
   404 49.421086000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) reply
   405 49.747185000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) request
   406 49.747334000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   407 49.747342000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) reply
   408 50.420911000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   409 50.420932000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) request
   410 50.421099000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) reply
   411 50.748659000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) request
   412 50.748795000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   413 50.748800000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) reply
   414 51.420976000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   415 51.420988000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) request
   416 51.421147000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) reply
   417 51.750128000    10.0.0.11            10.0.0.14            ICMP      Echo (ping) request
   418 51.750274000    ExtremeN_1e:fd:60    ExtremeN_3b:23:10    0x88b5    PRI: 0  DROP: 0 ID: 1410
   419 51.750283000    10.0.0.14            10.0.0.11            ICMP      Echo (ping) reply

⊞ Frame 412 (100 bytes on wire, 100 bytes captured)
⊞ Ethernet II, Src: ExtremeN_1e:fd:60 (00:04:96:1e:fd:60), Dst: ExtremeN_3b:23:10 (00:04:96:3b:23:10)
⊟ IEEE 802.1ad, ID: 1410
    000. .... .... .... = Priority: 0
    ...0 .... .... .... = DEI: 0
    .... 0101 1000 0010 = ID: 1410
    Type: Unknown (0x88b5)
⊞ Data (82 bytes)
```

Figure 5.7: Packet capture in PBT test

5.6.4 Test 2: PBT Resilience

In order to restore the network in the event of failure, a secondary path can be defined beside the working path to restore the network as soon as possible, as shown in Figure 5.8).

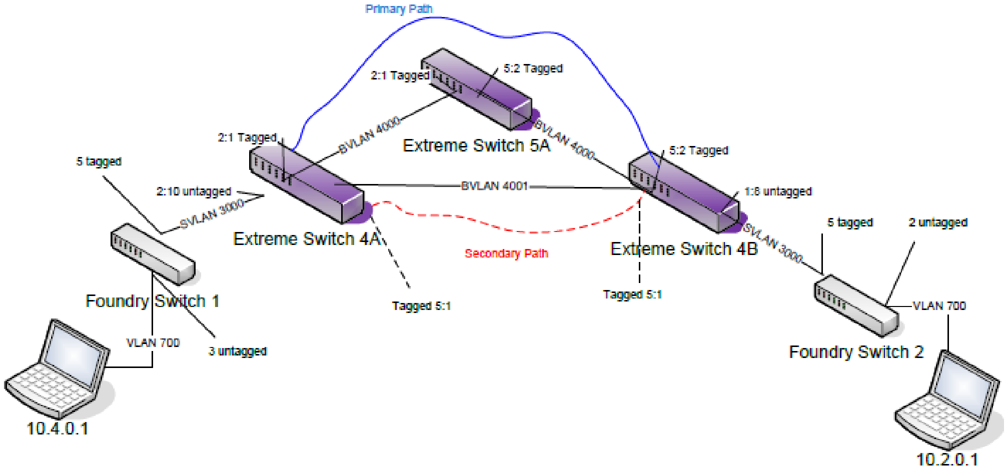


Figure 5.8: PBT secondary tunnel

5.6.4.1 *Configurations and Results*

The commands for defining and manually changing to the secondary path are shown below.

Switch 4A:

```
create bvlan ProBjan
conf bvlan ProBjan tag 2410
conf bvlan ProBjan add port 5:1 tagged
```

Change to the secondary path:

```
conf bvlan bjan del svlan sjan
conf bvlan ProBjan add svlan sjan

del fdb 00:04:96:3b:23:10 bvlan bjan
create fdb 00:04:96:3b:23:10 bvlan Probjan port 5:1
```

Change back to the primary path:

```
conf bvlan Probjan del svlan sjan
conf bvlan Bjan add svlan sjan
```

```
del fdb 00:04:96:3b:23:10 bvlan Probjan
create fdb 00:04:96:3b:23:10 bvlan bjan port 2:1
```

Switch 4B:

```
create bvlan ProBjan
conf bvlan ProBjan tag 2410
conf bvlan ProBjan add port 5:1 tagged
```

Change to the secondary path:

```
conf bvlan bjan del svlan sjan
conf bvlan ProBjan add svlan sjan

del fdb 00:04:96:1e:fd:60 bvlan bjan
create fdb 00:04:96:1e:fd:60  bvlan Probjan port 5:1
```

Change back to the primary path:

```
conf bvlan Probjan del svlan sjan
conf bvlan Bjan add svlan sjan

del fdb 00:04:96:1e:fd:60  bvlan Probjan
create fdb 00:04:96:1e:fd:60  bvlan bjan port 5:2
```

The test results showed that a manual switching between primary and backup PBT tunnels worked in an expected and straightforward way.

5.6.5 Test 3: CFM and PBT Protection and Resilience

5.6.5.1 *Testbed Setup*

CFM end points were set up on switches 4A and 4B at the edges of the PBT network. A PBT trunk group was established between the two edge switches, which consisted of two paths or Backbone VLANs (B-VLANs), the primary path being B-VLAN 1410, and the secondary path B-VLAN 2410. The setup is shown in Figure 5.9 below.

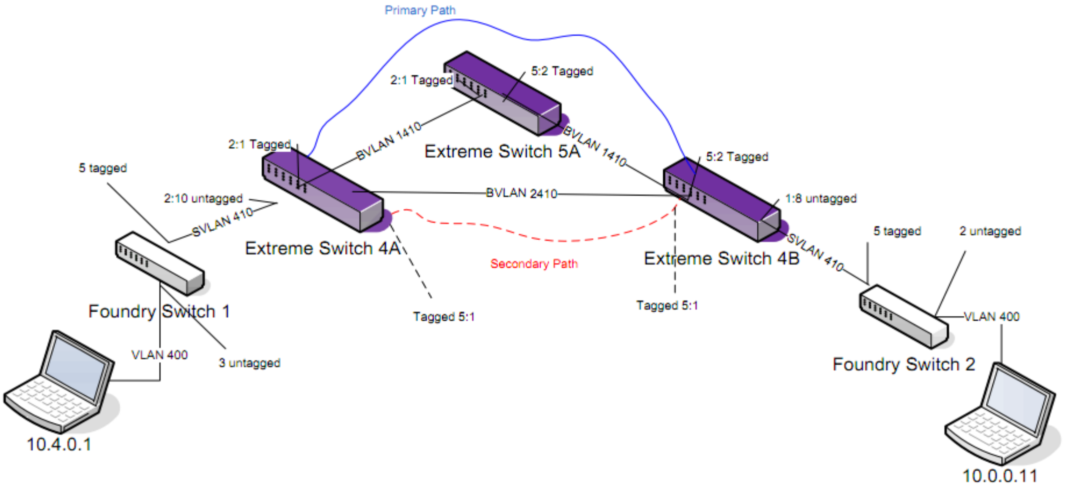


Figure 5.9: Essex testbed setup – CFM and PBT protection and resilience

5.6.5.2 *Configurations*

The testbed elements were configured as follows:

Switch (4A) Configurations

B-VLAN and S-VLAN:

```
//Configuration of the corresponding B vlan and S vlans.
conf bvlan bjan add port 2:1 tag
conf bvlan probjan add port 5:1 tag
conf svlan sjan add port 2:10 untagged

conf bvlan bjan del svlan sjan
conf bvlan probjan del svlan sjan
```

FDB:

```
//manual entering the edge addresses for the PBT tunnel set up
creat fdb 00:04:96:3b:23:10 bvlan bjan port 2:1
creat fdb 00:04:96:3b:23:10 bvlan probjan port 5:1
```

CFM:

```
//CFM session was set up to monitor the connection, and to see if it is alive
      or not.

create cfm domain string "pbt-d2" md-level 2
```

```
configure cfm domain "pbt-d2" add association string "pbt-d2-protecting" bvlan
      probjan
configure cfm domain "pbt-d2" association "pbt-d2-protecting" destination-mac-
      type unicast

configure cfm domain "pbt-d2" add association string "pbt-d2-working" bvlan
      bjan

configure cfm domain "pbt-d2" association "pbt-d2-working" destination-mac-type
      unicast

configure cfm domain "pbt-d2" association "pbt-d2-protecting" add remote-mep
      250 mac-address 00:04:96:3b:23:10

configure cfm domain "pbt-d2" association "pbt-d2-working" add remote-mep 150
      mac-address 00:04:96:3b:23:10

configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 add
      end-point down 200
configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 end-
      point down transmit-interval 100

configure cfm domain "pbt-d2" association "pbt-d2-working" ports 2:1 add end-
      point down 100
configure cfm domain "pbt-d2" association "pbt-d2-working" ports 2:1 end-point
      down transmit-interval 100
```

PBT trunk:

```
// a trunk group of two PBT tunnels have been installed, so one of them was the
      primary tunnel
//for data transfer, whislt the other one was kept as a backup path. After each
      failer in the
//primary path, the trunk changed to the protection path, until the first path
      was restored.
create pbt trunk-group "pbt-test1"
configure pbt trunk-group "pbt-test1" peer-bridge-mac 00:04:96:3b:23:10
configure pbt trunk-group "pbt-test1" working-path add bvlan bjan
configure pbt trunk-group "pbt-test1" protecting-path add bvlan probjan
configure pbt trunk-group "pbt-test1" working-path cfm "pbt-d2" association
      "pbt-d2-working" "100"
configure pbt trunk-group "pbt-test1" protecting-path cfm "pbt-d2" association
      "pbt-d2-protecting" "200"
configure pbt trunk-group "pbt-test1" add svlan sjan
```

Switch (4B) Configurations

B-VLAN and S-VLAN:

```
//Configuration of the corresponding B vlan and S vlans.
conf bvlan bjan add port 5:2 tag
conf bvlan probjan add port 5:1 tag
conf svlan sjan del poty 1:17
conf svlan sjan add port 1:8 untagged

conf bvlan bjan del svlan sjan
conf bvlan probjan del svlan sjan
```

FDB:

```
//manual entering the edge addresses for the PBT tunnel set up
creat fdb 00:04:96:1e:fd:60 bvlan bjan port 5:2
creat fdb 00:04:96:1e:fd:60 bvlan probjan port 5:1
```

CFM:

```
//CFM session was set up to monitor the connection, and to see if it is alive
      or not.
create cfm domain string "pbt-d2" md-level 2
configure cfm domain "pbt-d2" add association string "pbt-d2-protecting" bvlan
      probjan
configure cfm domain "pbt-d2" association "pbt-d2-protecting" destination-mac-
      type unicast

configure cfm domain "pbt-d2" add association string "pbt-d2-working" bvlan
      bjan
configure cfm domain "pbt-d2" association "pbt-d2-working" destination-mac-type
      unicast

configure cfm domain "pbt-d2" association "pbt-d2-protecting" add remote-mep
      250 mac-address 00:04:96:1e:fd:60

configure cfm domain "pbt-d2" association "pbt-d2-working" add remote-mep 100
      mac-address 00:04:96:1e:fd:60

configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 add
      end-point down 250
configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 end-
      point down transmit-interval 100
```

```
configure cfm domain "pbt-d2" association "pbt-d2-working" ports 5:2 add end-
      point down 150
configure cfm domain "pbt-d2" association "pbt-d2-working" ports 5:2 end-point
      down transmit-interval 100
```

PBT trunks with protection paths:

```
// a trunk group of two PBT tunnels have been installed, so one of them was the
      primary tunnel
//for data transfer, twhilst the other one was kept as a backup path. After
      each failer in the
//primary path, the trunk changed to the protection path, until the first path
      was restored.

create pbt trunk-group "pbt-test1"
configure pbt trunk-group "pbt-test1" peer-bridge-mac 00:04:96:1e:fd:60
configure pbt trunk-group "pbt-test1" working-path add bvlan bjan
configure pbt trunk-group "pbt-test1" protecting-path add bvlan probjan
configure pbt trunk-group "pbt-test1" working-path cfm "pbt-d2" association
      "pbt-d2-working" "150"
configure pbt trunk-group "pbt-test1" protecting-path cfm "pbt-d2" association
      "pbt-d2-protecting" "250"
configure pbt trunk-group "pbt-test1" add svlan sjan
```

5.6.5.3 *Expected Results*

Connectivity messages between the two edge switches should maintain the connectivity by changing the working path from the primary to the secondary path in the event of failure in the primary path, and changing back again whenever the primary path becomes available.

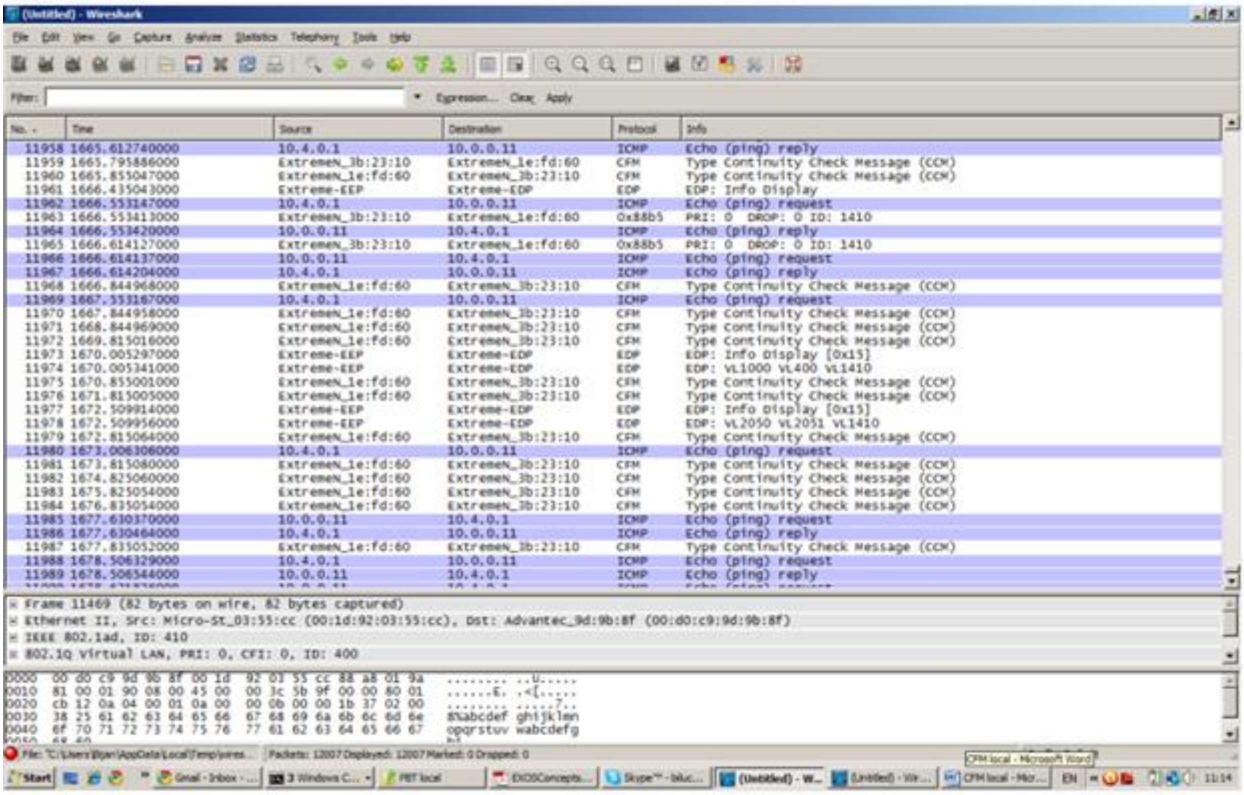
5.6.5.4 *Results*

The test results were as expected, as shown by the screen captures in Figure 5.10 to Figure 5.12 below.

The upper part of the screen capture in Figure 5.11 was taken when all ports along the path with B-VID 1410 were up and hence a primary PBT tunnel with B-VID 1410 was in use.

The bottom part of the screenshot in Figure 5.11 was taken when the port 2:1 of switch 5A was put down. The output of the `show pbt` command shows that the backup PBT tunnel with B-VID 2410 became the path in use.

The screenshot in Figure 5.12 shows that after switching, port 2:1 on the primary PBT tunnel with B-VID 1410 returned to use.

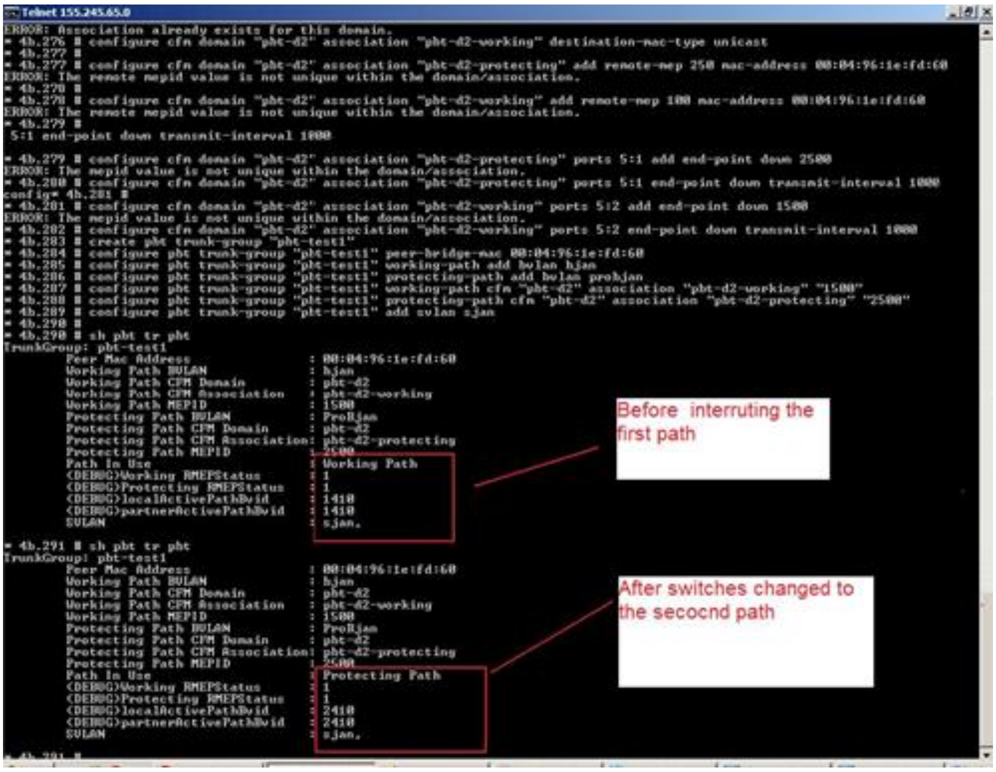

(Untitled) - Wireshark
File Edit View Go Capture Analyze Statistics Telephony Tools Help
Filter:
Expression... Clear Apply

```
No. -  Time               Source               Destination          Protocol  Info
11958 1665.612740000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) reply
11959 1665.795886000      ExtremeN_3b:23:10    ExtremeN_1e:fd:60    CFM       Type Continuity Check Message (CCM)
11960 1665.855047000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11961 1666.435043000      Extreme-EEP          Extreme-EDP          EDP       EDP: Info Display
11962 1666.553147000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) request
11963 1666.553413000      ExtremeN_3b:23:10    ExtremeN_1e:fd:60    0x88b5    PRI: 0  DROP: 0 ID: 1410
11964 1666.553420000      10.0.0.11            10.4.0.1             ICMP      Echo (ping) reply
11965 1666.614127000      ExtremeN_3b:23:10    ExtremeN_1e:fd:60    0x88b5    PRI: 0  DROP: 0 ID: 1410
11966 1666.614137000      10.0.0.11            10.4.0.1             ICMP      Echo (ping) request
11967 1666.614204000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) reply
11968 1666.844968000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11969 1667.553167000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) request
11970 1667.844958000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11971 1668.844969000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11972 1669.815016000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11973 1670.005297000      Extreme-EEP          Extreme-EDP          EDP       EDP: Info Display [0x15]
11974 1670.005341000      Extreme-EEP          Extreme-EDP          EDP       EDP: VL1000 VL400 VL1410
11975 1670.855001000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11976 1671.815005000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11977 1672.509914000      Extreme-EEP          Extreme-EDP          EDP       EDP: Info Display [0x15]
11978 1672.509956000      Extreme-EEP          Extreme-EDP          EDP       EDP: VL2050 VL2051 VL1410
11979 1672.815064000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11980 1673.006306000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) request
11981 1673.815080000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11982 1674.825060000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11983 1675.825054000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11984 1676.835054000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11985 1677.630370000      10.0.0.11            10.4.0.1             ICMP      Echo (ping) request
11986 1677.630464000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) reply
11987 1677.835052000      ExtremeN_1e:fd:60    ExtremeN_3b:23:10    CFM       Type Continuity Check Message (CCM)
11988 1678.506329000      10.4.0.1             10.0.0.11            ICMP      Echo (ping) request
11989 1678.506544000      10.0.0.11            10.4.0.1             ICMP      Echo (ping) reply
```

```
⊞ Frame 11469 (82 bytes on wire, 82 bytes captured)
⊞ Ethernet II, Src: Micro-St_03:55:cc (00:1d:92:03:55:cc), Dst: Advantec_9d:9b:8f (00:d0:c9:9d:9b:8f)
⊞ IEEE 802.1ad, ID: 410
⊞ 802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 400
```

```
0000  00 d0 c9 9d 9b 8f 00 1d  92 03 55 cc 88 a8 01 9a   ........ ..U.....
0010  81 00 01 90 08 00 45 00  00 3c 5b 9f 00 00 80 01   ......E. .<[.....
0020  cb 12 0a 04 00 01 0a 00  00 0b 00 00 1b 37 02 00   ........ .....7..
0030  38 25 61 62 63 64 65 66  67 68 69 6a 6b 6c 6d 6e   8%abcdef ghijklmn
0040  6f 70 71 72 73 74 75 76  77 61 62 63 64 65 66 67   opqrstuv wabcdefg
```

Figure 5.10: Packet capture showing the operation of the secondary path

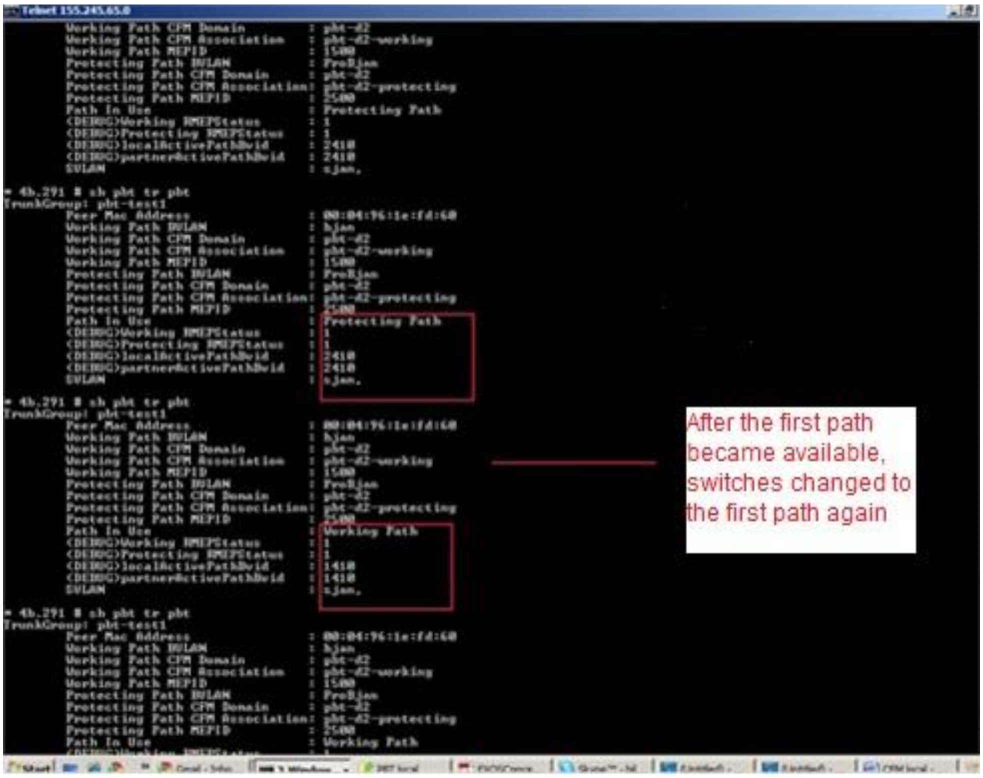


```
ERROR: Association already exists for this domain.
* 4b.276 # configure cfm domain "pbt-d2" association "pbt-d2-working" destination-mac-type unicast
* 4b.277 #
* 4b.277 # configure cfm domain "pbt-d2" association "pbt-d2-protecting" add remote-mep 250 mac-address 00:04:96:1e:fd:60
ERROR: The remote mepid value is not unique within the domain/association.
* 4b.278 #
* 4b.278 # configure cfm domain "pbt-d2" association "pbt-d2-working" add remote-mep 100 mac-address 00:04:96:1e:fd:60
ERROR: The remote mepid value is not unique within the domain/association.
* 4b.279 #
 5:1 end-point down transmit-interval 1000

* 4b.279 # configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 add end-point down 2500
ERROR: The mepid value is not unique within the domain/association.
* 4b.280 # configure cfm domain "pbt-d2" association "pbt-d2-protecting" ports 5:1 end-point down transmit-interval 1000
config* 4b.281 #
* 4b.281 # configure cfm domain "pbt-d2" association "pbt-d2-working" ports 5:2 add end-point down 1500
ERROR: The mepid value is not unique within the domain/association.
* 4b.282 # configure cfm domain "pbt-d2" association "pbt-d2-working" ports 5:2 end-point down transmit-interval 1000
* 4b.283 # create pbt trunk-group "pbt-test1"
* 4b.284 # configure pbt trunk-group "pbt-test1" peer-bridge-mac 00:04:96:1e:fd:60
* 4b.285 # configure pbt trunk-group "pbt-test1" working-path add bvlan bjan
* 4b.286 # configure pbt trunk-group "pbt-test1" protecting-path add bvlan probjan
* 4b.287 # configure pbt trunk-group "pbt-test1" working-path cfm "pbt-d2" association "pbt-d2-working" "1500"
* 4b.288 # configure pbt trunk-group "pbt-test1" protecting-path cfm "pbt-d2" association "pbt-d2-protecting" "2500"
* 4b.289 # configure pbt trunk-group "pbt-test1" add svlan sjan
* 4b.290 #
* 4b.290 # sh pbt tr pbt
TrunkGroup: pbt-test1
        Peer Mac Address                : 00:04:96:1e:fd:60
        Working Path BVLAN              : bjan
        Working Path CFM Domain         : pbt-d2
        Working Path CFM Association    : pbt-d2-working
        Working Path MEPID              : 1500
        Protecting Path BVLAN           : ProBjan
        Protecting Path CFM Domain      : pbt-d2
        Protecting Path CFM Association: pbt-d2-protecting
        Protecting Path MEPID           : 2500
        Path In Use                     : Working Path
        <DEBUG>Working RMEPStatus       : 1
        <DEBUG>Protecting RMEPStatus    : 1
        <DEBUG>localActivePathBvid      : 1410
        <DEBUG>partnerActivePathBvid    : 1410
        SVLAN                           : sjan.

* 4b.291 # sh pbt tr pbt
TrunkGroup: pbt-test1
        Peer Mac Address                : 00:04:96:1e:fd:60
        Working Path BVLAN              : bjan
        Working Path CFM Domain         : pbt-d2
        Working Path CFM Association    : pbt-d2-working
        Working Path MEPID              : 1500
        Protecting Path BVLAN           : ProBjan
        Protecting Path CFM Domain      : pbt-d2
        Protecting Path CFM Association: pbt-d2-protecting
        Protecting Path MEPID           : 2500
        Path In Use                     : Protecting Path
        <DEBUG>Working RMEPStatus       : 1
        <DEBUG>Protecting RMEPStatus    : 1
        <DEBUG>localActivePathBvid      : 2410
        <DEBUG>partnerActivePathBvid    : 2410
        SVLAN                           : sjan.

* 4b.291 #
```

Before interruting the first path

After switches changed to the secocnd path

Figure 5.11: PBT trunk status before and after interruption in the primary path



```
        Working Path CFM Domain          : pbt-d2
        Working Path CFM Association     : pbt-d2-working
        Working Path MEPID               : 1500
        Protecting Path BVLAN            : ProBjan
        Protecting Path CFM Domain       : pbt-d2
        Protecting Path CFM Association  : pbt-d2-protecting
        Protecting Path MEPID            : 2500
        Path In Use                      : Protecting Path
        <DEBUG>Working RMEPStatus        : 1
        <DEBUG>Protecting RMEPStatus     : 1
        <DEBUG>localActivePathBvid       : 2410
        <DEBUG>partnerActivePathBvid     : 2410
        SVLAN                            : sjan,

* 4b.291 # sh pbt tr pbt
TrunkGroup: pbt-test1
        Peer Mac Address                 : 00:04:96:1e:fd:60
        Working Path BVLAN               : bjan
        Working Path CFM Domain          : pbt-d2
        Working Path CFM Association     : pbt-d2-working
        Working Path MEPID               : 1500
        Protecting Path BVLAN            : ProBjan
        Protecting Path CFM Domain       : pbt-d2
        Protecting Path CFM Association  : pbt-d2-protecting
        Protecting Path MEPID            : 2500
        Path In Use                      : Protecting Path
        <DEBUG>Working RMEPStatus        : 1
        <DEBUG>Protecting RMEPStatus     : 1
        <DEBUG>localActivePathBvid       : 2410
        <DEBUG>partnerActivePathBvid     : 2410
        SVLAN                            : sjan,

* 4b.291 # sh pbt tr pbt
TrunkGroup: pbt-test1
        Peer Mac Address                 : 00:04:96:1e:fd:60
        Working Path BVLAN               : bjan
        Working Path CFM Domain          : pbt-d2
        Working Path CFM Association     : pbt-d2-working
        Working Path MEPID               : 1500
        Protecting Path BVLAN            : ProBjan
        Protecting Path CFM Domain       : pbt-d2
        Protecting Path CFM Association  : pbt-d2-protecting
        Protecting Path MEPID            : 2500
        Path In Use                      : Working Path
        <DEBUG>Working RMEPStatus        : 1
        <DEBUG>Protecting RMEPStatus     : 1
        <DEBUG>localActivePathBvid       : 1410
        <DEBUG>partnerActivePathBvid     : 1410
        SVLAN                            : sjan,

* 4b.291 # sh pbt tr pbt
TrunkGroup: pbt-test1
        Peer Mac Address                 : 00:04:96:1e:fd:60
        Working Path BVLAN               : bjan
        Working Path CFM Domain          : pbt-d2
        Working Path CFM Association     : pbt-d2-working
        Working Path MEPID               : 1500
        Protecting Path BVLAN            : ProBjan
        Protecting Path CFM Domain       : pbt-d2
        Protecting Path CFM Association  : pbt-d2-protecting
        Protecting Path MEPID            : 2500
        Path In Use                      : Working Path
```

After the first path became available, switches changed to the first path again

Figure 5.12: After the primary path has been fixed, it is used again

5.7 JANET Lumen House Local PBB-TE Tests

This section covers the PBB-TE tests carried out by JANET Lumen House:

- Unprotected tunnels.
- CFM.
- Performance monitoring.
- Path protection.
- Traffic policing.

It begins by describing the Lumen House testbed setup.

5.7.1 Lumen House Testbed Setup

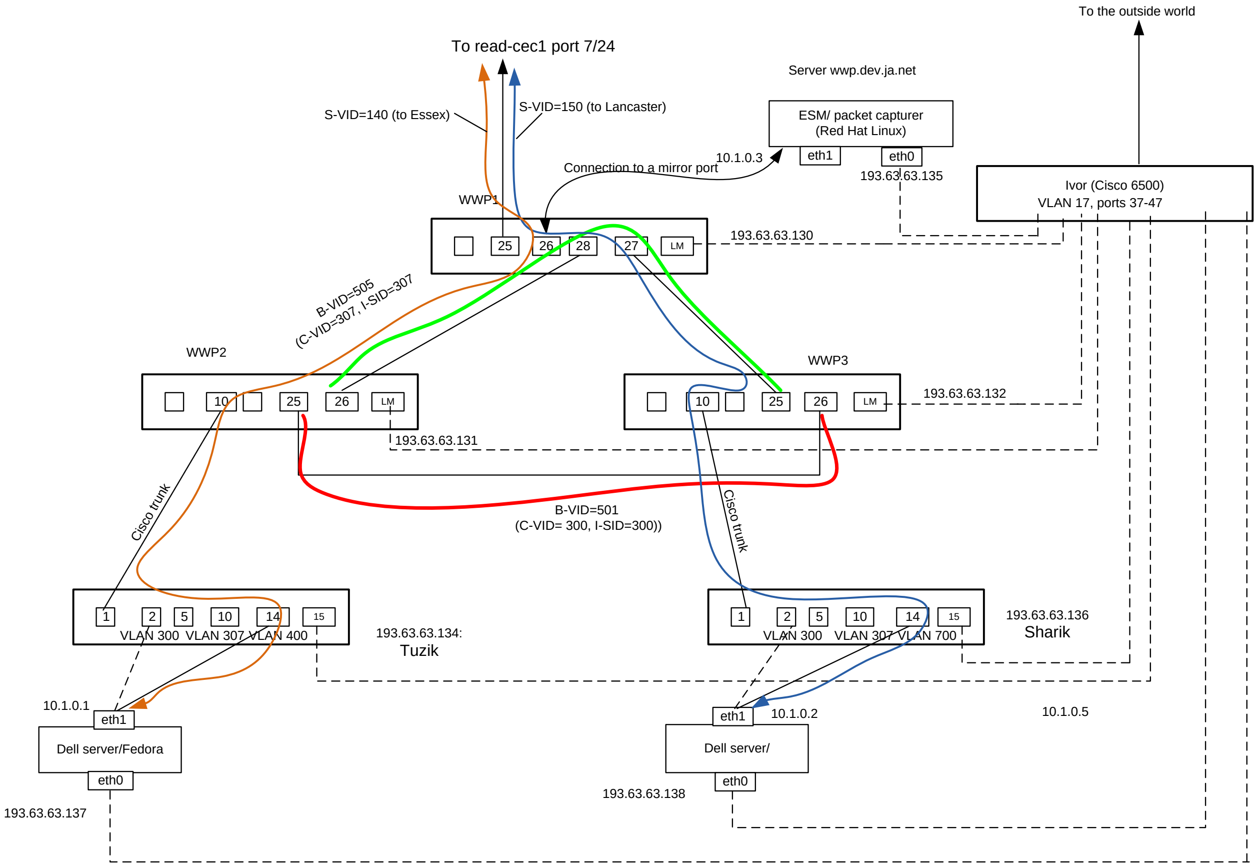


Figure 5.13: Lumen House testbed

The Lumen House testbed used in the lab trial is shown in Figure 5.13. The equipment and software are described below.

Three Ciena LE-311v boxes (WWP1, WWP2 and WWP3) were used for establishing PBT tunnels and circuits within them. The triangular topology of the testbed was created by connecting the 1 GE optical ports of the LE-311v boxes. (A triangular topology is the minimum configuration that allows the TE and resilience capabilities of a network to be checked.) All the boxes were bought with LightningEdge Operating System (LEOS) v.4.4, which was upgraded to v.4.6 during the tests.

Two Cisco 2950 switches and two Dell servers were used to emulate customer networks. Each 2950 switch supported two customer VLANs with C-VID 300 and C-VID 307.

The Dell servers had two Ethernet interfaces, one of which was used to emulate a customer node belonging to the 10.1.0.0 subnet. The servers’ second Ethernet interface servers was used for management; the subnet 193.63.63.128/26, accessible from outside the JANET development LAN, was used for this purpose.

A management server (Red Hat Linux Enterprise 5.0 platform) with Ethernet Service Manager (ESM) from Ciena was installed (Server wwp.dev.ja.net in Figure 5.13). The management server was also used for capturing traffic on mirrored ports of the LE-311v boxes (Wireshark was installed on the server for this purpose).

5.7.2 Test 1: Unprotected Tunnels

5.7.2.1 *Test Objective*

The aim of this test was to establish two unprotected PBT tunnels between ports on WWP boxes and use them for connecting customer services (customer VLANs 300 and 307). This would also allow the trial participants to gain basic experience in setting up and managing PBT tunnels.

5.7.2.2 *Test Setup*

Two PBT tunnels were set up:

- An indirect tunnel with B-VID 505 between WWP2 port 26 and WWP3 port 25, going through the WWP1 ports 28 and 27.
- A direct tunnel with B-VID 501 between WWP2 port 25 and WWP3 port 26.

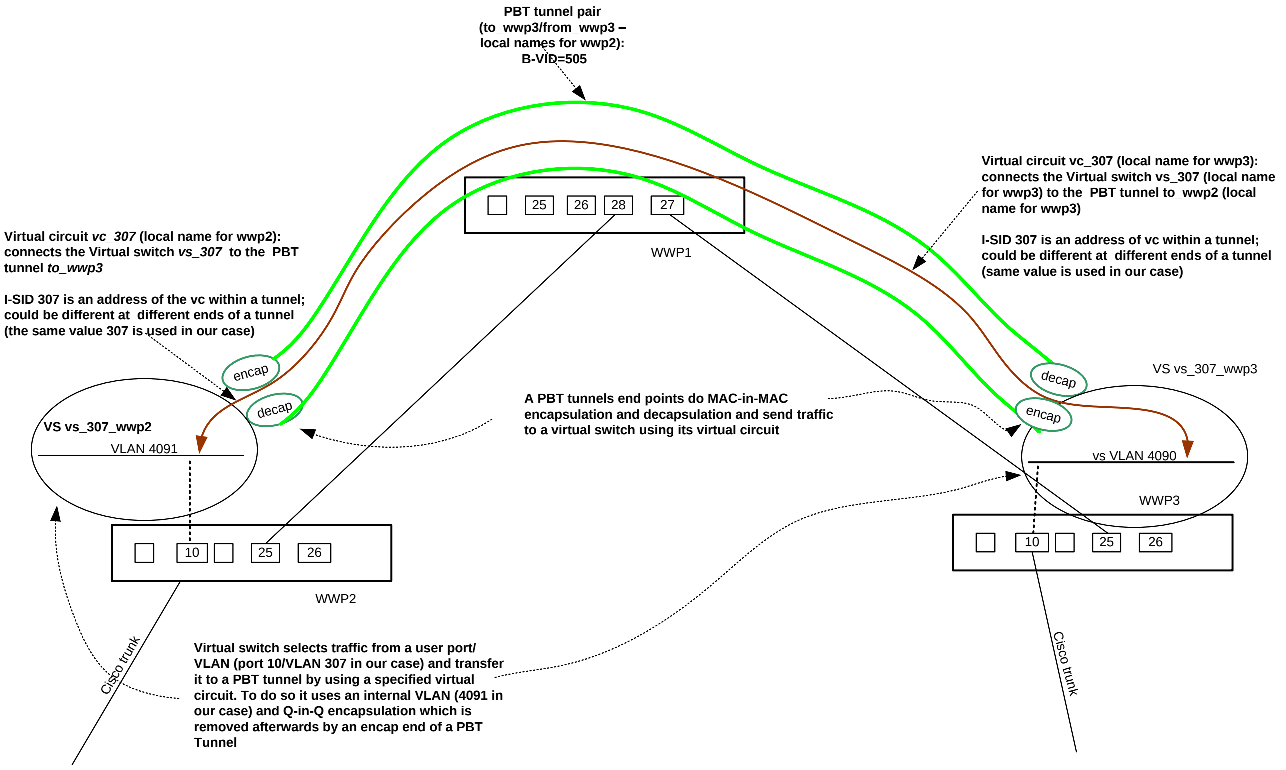
Tunnel 505 was to be used for carrying the customer service VLAN 307 while tunnel 501 was for the customer service VLAN 300.

5.7.2.3 *Configuration*

As shown in Figure 5.14, a PBT tunnel in the WWP implementation consisted of two uni-directional parts:

- **encap**, which encapsulates Ethernet frames entering the tunnel using 802.1ah MAC-in-MAC encapsulation and sends them down the tunnel.
- **decap**, which decapsulates Ethernet frames leaving the tunnel.

Generally, each part of a tunnel can use a separate route through a network, but this option was not tried in the present tests.



PBT tunnel pair (to_wwp3/from_wwp3 – local names for wwp2): B-VID=505

Virtual circuit vc_307 (local name for wwp3): connects the Virtual switch vs_307 (local name for wwp3) to the PBT tunnel to_wwp2 (local name for wwp3)

I-SID 307 is an address of vc within a tunnel; could be different at different ends of a tunnel (same value is used in our case)

Virtual circuit *vc_307* (local name for wwp2): connects the Virtual switch *vs_307* to the PBT tunnel *to_wwp3*

I-SID 307 is an address of the vc within a tunnel; could be different at different ends of a tunnel (the same value 307 is used in our case)

A PBT tunnels end points do MAC-in-MAC encapsulation and decapsulation and send traffic to a virtual switch using its virtual circuit

Virtual switch selects traffic from a user port/ VLAN (port 10/VLAN 307 in our case) and transfer it to a PBT tunnel by using a specified virtual circuit. To do so it uses an internal VLAN (4091 in our case) and Q-in-Q encapsulation which is removed afterwards by an encap end of a PBT Tunnel

Figure 5.14: Tunnel and service elements

The WWP2 configuration describing the tunnels was as follows:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! TUNNEL CONFIG:
!
!The interaction of tunnels, virtual circuits and vrtual switches are shown on
!the PBT config elements diagram
tunnel encap create static-pbt to_wwp3 b-vid 505 dest-bridge-name wwp3 port 26
tunnel encap create static-pbt 501_to_wwp3 b-vid 501 dest-bridge-name wwp3 port
      25
tunnel decap create static-pbt from_wwp3 b-vid 505 src-bridge-name wwp3 port 26
tunnel decap create static-pbt 501_from_wwp3 b-vid 501 src-bridge-name wwp3
      port 25
!
tunnel pair create tnl-pair wwp3 encap-pbt to_wwp3 decap-pbt from_wwp3
tunnel pair create tnl-pair 501_tunnel encap-pbt 501_to_wwp3 decap-pbt
      501_from_wwp3
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

The names of tunnel parts, such as to_wwp3, are arbitrary; the names of tunnel destinations (such as wwp3) were specified by configuring the loop-back style MAC addresses on the boxes and then binding these MAC addresses to boxes' names.

For example, on WWP3 the following CLI command was entered for specifying its loopback MAC address:

```
pbt set bridge-mac 04:00:00:00:00:03
```

On WWP2 the command

```
pbt remote-bridge create bridge-name wwp3 bridge-mac 04:00:00:00:00:03
```

bound the WWP3 loopback address to the name wwp3.

The `tunnel pair create` command is not mandatory but it is useful as it couples the encap and decap parts of a tunnel into a single entity, which is useful in some cases: for example, when one part of a tunnel goes down and another does not, a coupled tunnel declares both as down.

A virtual circuit and a virtual switch are elements that WWP boxes use to create a customer service within a PBT tunnel to carry selected customer traffic.

- A virtual circuit (WWP's term) is what IEEE terms a service instance, which is specified by the I-SID tag within 802.1ah encapsulation. It is actually a logical connection between user interfaces (physical or VLAN-based) and tunnelled by a PBT tunnel. Each PBT tunnel can carry up to 16 million customer connections as the I-SID field is 24 bits. However, the present tests used one I-SID connection per PBT tunnel. An I-SID connection can be thought of as the rough equivalent of an EoMPLS pseudowire.
- A virtual switch selects specific user traffic entering a WWP switch port.

The WWP2 configuration specifying virtual circuits and virtual switches was as follows:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! VIRTUAL-CIRCUIT CONFIG:  virtual circuits
!
virtual-circuit pbt create static-vc vs_307  egress-isid 307 ingress-isid 307
      tunnel to_wwp3
virtual-circuit pbt create static-vc 501_vc  egress-isid 300 ingress-isid 300
      tunnel
501_to_wwp3
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! VIRTUAL-SWITCH CONFIG:
!
virtual-switch add reserved-vlan 4090-4094
!
virtual-switch ethernet create vs vs_307 vc vs_307
```

```
virtual-switch ethernet create vs vs_300_501 vc 501_vc
!
virtual-switch ethernet add vs vs_307 port 10 vlan 307
virtual-switch ethernet add vs vs_300_501 port 10 vlan 300
!!!!!!!
```

Both virtual switches created select user traffic from port 10 of WWP2 but with different C-VID values.

A virtual switch creation command binds a switch to a virtual circuit while a virtual circuit creation command binds a circuit with a particular PBT tunnel (its encap part) and gives it an I-SID value.

In the test the I-SID values chosen coincided with C-VIDs.

Please note that WWP boxes can translate ingress and egress I-SIDs. However, in the tests the same value was used for both ends of virtual circuits; these values coincided with user VLAN IDs for simplicity.

A virtual switch in the WWP implementation needs a VLAN ID to carry out its internal operations. A switch uses this VLAN ID for intermediate 802.1ad (also known as QinQ) encapsulation of a customer frame; that encapsulation is then removed by the encap (ingress) end of a PBT tunnel so that a customer frame within a PBT tunnel has its original header plus an 802.1ah header but no QinQ header in between.

In the tests the range of VLAN IDs from 4090 to 4094 was reserved for use by virtual switches.

5.7.2.4 *Results*

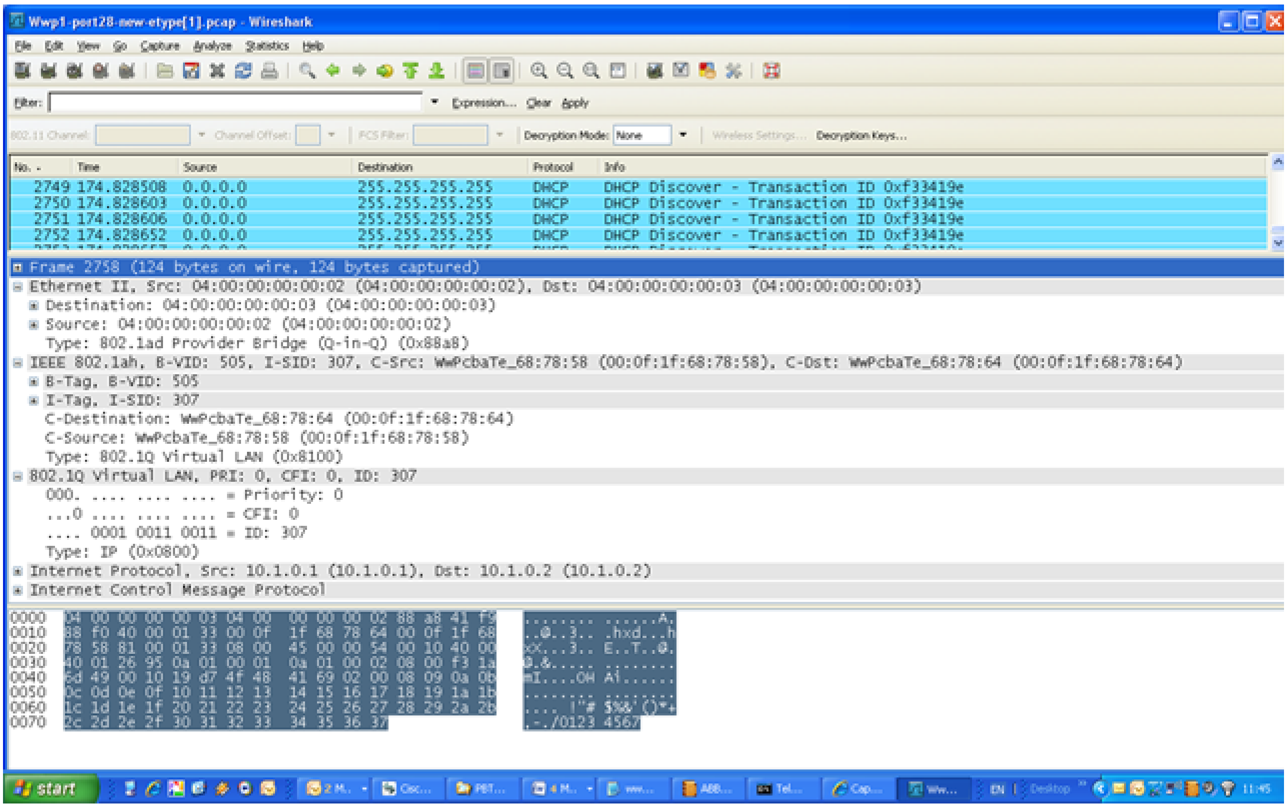
Two PBT tunnels and two customer connections were tested by periodically pinging server 10.1.0.2 from server 10.1.0.1 and vice versa (over long periods of time with an interval of 1 sec).

The pinging showed 100% performance (no dropped frames) for both C-VIDs, 300 and 307.

The WWP CLI command `tunnel show` correctly reflected the status of both tunnels:

```
wwp2> tunnel show
+------------------------ ENCAP TUNNEL TABLE --------------------------------+
|               |            |  State   |               |      |    |        |
|Name           |Type        |Oper|Admin|Destination    |B-VID |Role|Active  |
+---------------+------------+----------+---------------+------+----+--------+
|to_wwp3        |encap-pbt   |En  |En   |wwp3           |505   |pri |Yes     |
|501_to_wwp3    |encap-pbt   |En  |En   |wwp3           |501   |pri |Yes     |
+---------------+------------+----------+---------------+------+----+--------+
+------------------------ DECAP TUNNEL TABLE --------------------------------+
|                 |              | Oper  |                                   |
| Name            | Type         | State | B-VID                             |
+-----------------+--------------+-------+-----------------------------------+
| from_wwp3       | decap-pbt    | En    | 505                               |
| 501_from_wwp3    | decap-pbt     | En     | 501                              |
+------------------+---------------+--------+----------------------------------+
```

Frames captured by Wireshark software [Wireshark] on port 28 of WWP1 were correctly encapsulated according to the 802.1ah standard, as shown in Figure 5.15:



```
No. -  Time        Source    Destination      Protocol  Info
  2749 174.828508  0.0.0.0   255.255.255.255  DHCP      DHCP Discover - Transaction ID 0xf33419e
  2750 174.828603  0.0.0.0   255.255.255.255  DHCP      DHCP Discover - Transaction ID 0xf33419e
  2751 174.828606  0.0.0.0   255.255.255.255  DHCP      DHCP Discover - Transaction ID 0xf33419e
  2752 174.828652  0.0.0.0   255.255.255.255  DHCP      DHCP Discover - Transaction ID 0xf33419e
```

```
⊞ Frame 2758 (124 bytes on wire, 124 bytes captured)
⊟ Ethernet II, Src: 04:00:00:00:00:02 (04:00:00:00:00:02), Dst: 04:00:00:00:00:03 (04:00:00:00:00:03)
  ⊞ Destination: 04:00:00:00:00:03 (04:00:00:00:00:03)
  ⊞ Source: 04:00:00:00:00:02 (04:00:00:00:00:02)
    Type: 802.1ad Provider Bridge (Q-in-Q) (0x88a8)
⊟ IEEE 802.1ah, B-VID: 505, I-SID: 307, C-Src: WwPcbaTe_68:78:58 (00:0f:1f:68:78:58), C-Dst: WwPcbaTe_68:78:64 (00:0f:1f:68:78:64)
  ⊞ B-Tag, B-VID: 505
  ⊞ I-Tag, I-SID: 307
    C-Destination: WwPcbaTe_68:78:64 (00:0f:1f:68:78:64)
    C-Source: WwPcbaTe_68:78:58 (00:0f:1f:68:78:58)
    Type: 802.1Q Virtual LAN (0x8100)
⊟ 802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 307
    000. .... .... .... = Priority: 0
    ...0 .... .... .... = CFI: 0
    .... 0001 0011 0011 = ID: 307
    Type: IP (0x0800)
⊞ Internet Protocol, Src: 10.1.0.1 (10.1.0.1), Dst: 10.1.0.2 (10.1.0.2)
⊞ Internet Control Message Protocol
```

```
0000  04 00 00 00 00 03 04 00  00 00 00 02 88 a8 41 f9   ........ ......A.
0010  88 f0 40 00 01 33 00 0f  1f 68 78 64 00 0f 1f 68   ..@..3.. .hxd...h
0020  78 58 81 00 01 33 08 00  45 00 00 54 00 10 40 00   xX...3.. E..T..@.
0030  40 01 26 95 0a 01 00 01  0a 01 00 02 08 00 f3 1a   @.&..... ........
0040  6d 49 00 10 19 d7 4f 48  41 69 02 00 08 09 0a 0b   mI....OH Ai......
0050  0c 0d 0e 0f 10 11 12 13  14 15 16 17 18 19 1a 1b   ........ ........
0060  1c 1d 1e 1f 20 21 22 23  24 25 26 27 28 29 2a 2b   .... !"# $%&'()*+
0070  2c 2d 2e 2f 30 31 32 33  34 35 36 37               ,-./0123 4567
```

Figure 5.15: MAC-in-MAC (802.1ah) encapsulation of frames within a PBT tunnel

5.7.2.5 *Problems Encountered*

- On WWP LE-311v boxes, PBT tunnels can be created only on gigabit ports (25-27). Attempts to create a tunnel starting at a Fast Ethernet port resulted in a message “ERROR: creating encap tunnel ‘abcd’, Port is subscriber facing”. The reason is that only Field-Programmable Gate Array (FPGA) ports support PBT operations; LE-311v boxes have only 1 GE ports of that type.
- It is not possible to change a B-VID value for a tunnel that was specified during tunnel creation. The `tunnel set` option does not allow it; it is not even possible to delete the mistakenly configured tunnel, as LEOS says that there is a virtual circuit related to this tunnel.

5.7.3 Test 2: Connectivity Fault Management (CFM)

5.7.3.1 *Test Objective*

The objective of this test was to understand how CFM Continuity Check Messages (CCMs), or heartbeat messages[2], can be used for monitoring the state of PBT tunnels and I-SID connections between virtual switches.

5.7.3.2 *Test Setup*

CFM sessions for two PBT tunnels were established and for two I-SID connections within them.

5.7.3.3 *Configuration*

Only the CCM function, which permanently monitors the state of a connection with heartbeat messages, is available for PBT tunnels and connections between virtual switches on WWP LE-311v boxes. The CFM standard also includes linktrace and loopback functions, which are useful for troubleshooting; they are implemented on LE-311v boxes but are only available for other types of LE-311v services, e.g. for plain VLANs.

CCMs are generated by each end of a PBT tunnel or a virtual switch independently. However, they must have both a common name and the same Maintenance Domain (MD) level value to be recognised by the remote end of a PBT tunnel or a virtual switch. They must also establish a Maintenance Association (MA) between the end points of a connection.

The MD value for the virtual switches should be greater than the MD value for the PBT tunnel through which virtual switches communicate, as this reflects the place of these entities in the network hierarchy. The default value of an MD is 3 and this was used for PBT tunnels in the test; for virtual switches the MD value 4 was used.

Each end of a PBT tunnel or a virtual switch connection must have a unique number for a MEP. Establishing MIPs for PBT tunnels and virtual switches does not make sense as they are treated as a single hop entity.

The WWP2 configuration regarding the CCM configuration looked as follows:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! CFM CONFIG: global attributes
!
cfm enable
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! CFM CONFIG: services
!
```

[2] End point A assumes that end point B is alive if A regularly receives messages from B, hence these are referred to as “heartbeat messages”.

```
cfm service create static-pbt to_wwp3 name pbt_505_cfm next-mepid 100
cfm service enable service pbt_505_cfm
cfm service create static-pbt 501_to_wwp3 name pbt_501_cfm next-mepid 101
cfm service enable service pbt_501_cfm
cfm service create vs vs_300_501 name cfm_300 level 4 next-mepid 501
cfm service enable service cfm_300
cfm service create vs vs_307 name vs_307_b level 4 next-mepid 500
cfm service enable service vs_307_b
!!!!!!!!!
```

Four MAs were established:

- pbt_505_cfm for PBT tunnel to_wwp3 with B-VID 505.
- pbt_501_cfm for PBT tunnel 501_to_wwp3 with B-VID 501.
- cfm_300 for virtual switch vs_300_501 (the service instance going through PBT tunnel to_wwp3).
- vs_307_b for virtual switch vs_307 (the service instance going through PBT tunnel 501_to_wwp3).

The interval between CCMs was 1 second (the default).

5.7.3.4 *Results*

CCM monitoring correctly reflected changes in the state of PBT tunnels and virtual switch connections when the ports of the WWP boxes concerned were disconnected.

For example, when the ports were connected the WWP CLI showed that both tunnels had both administrative and operational states enabled (en) and no faults were indicated:

```
wwp3> cfm remote-mep show
+------------------------------ CFM REMOTE MEPS -------------------------------+
|                |     |                 |State|Total     |Seq  |Last     Fault|
|Service         |Mepid|Mac Address      |Ad|Op|Rx CCM    |Error|Seq Num  F|P|R|
+----------------+-----+-----------------+--+--+----------+-----+---------+-+-++
|pbt_505_cfm     |100  |04:00:00:00:00:02|en|en|3463227   |0    |3463227  | | ||
+----------------+-----+-----------------+--+--+----------+-----+---------+-+-++
|pbt_501_cfm     |101  |04:00:00:00:00:02|en|en|3463226   |0    |3463226  | | ||
+----------------+-----+-----------------+--+--+----------+-----+---------+-+-++
```

When the ports of both tunnels were disconnected, the CLI showed the Fault state of the MAs:

```
wwp3> cfm remote-mep show
+------------------------------ CFM REMOTE MEPS -------------------------------+
|                |     |                 |State|Total     |Seq  |Last     Fault|
|Service         |Mepid|Mac Address      |Ad|Op|Rx CCM    |Error|Seq Num |F|P|R|
+----------------+-----+-----------------+--+--+----------+-----+--------+-+-+-+
|pbt_505_cfm     |100  |04:00:00:00:00:02|en|en|3474064  |0    |3474064 |X| | |
+----------------+-----+-----------------+--+--+---------+-----+--------+-+-+-+
|pbt_501_cfm     |101  |04:00:00:00:00:02|en|en|3473941  |0    |3473941 |X| | |
+----------------+-----+-----------------+--+--+---------+-----+--------+-+-+-+
```

Mutual pinging of servers 10.1.0.1 and 10.1.0.2 confirmed the state information presented by the CFM CLI.

5.7.3.5 *Problems Encountered*

CFM for virtual switches stops working after rebooting despite the fact that it is configured properly and worked before rebooting.

The only way to fix this is to remove the proper CFM CONFIG for the virtual switch, reboot the box and then enter the same CFM CONFIG again, after which it starts working (unfortunately, only until the next reboot). WWP/Ciena were informed about the problem and suggested that it was an unknown bug in LEOS v.4.6. The case has been escalated within WWP/Ciena; the results of the investigation are as yet unknown. The newer version of the switch software turned out to be free of the bug.

5.7.4 Test 3: Performance Monitoring

5.7.4.1 *Test Objectives*

The aim of this test was to evaluate Y.1731 functionality in monitoring the performance (frame delays and jitter) of the boxes. Loss monitoring is also supported by the LE-311v switches, but this was not tested due to lack of time.

5.7.4.2 *Test Setup*

Delay and jitter measurement monitoring was initiated for two existing PBT tunnels.

5.7.4.3 *Configuration*

Performance monitoring functions specified by the Y.1731 standard are configured in LE-311v switches by the CFM CLI.

To configure repeated monitoring of the delay and jitter experienced by services with PBT tunnels, the following command lines were added to the CFM configuration of the WWP2 switch:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! CFM CONFIG: meps
!
cfm delay send service pbt_505_cfm port 26 mepid 200 repeat 1
cfm delay send service pbt_501_cfm port 25 mepid 201 repeat 1
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

The first line of the configuration initiated the generating of Delay Measurement Messages (DMMs) for CFM service pbt_505_cfm (which was enabled for the PBT tunnel with B-VID 505) through port 26. The destination MEP was defined by MEP ID 200. The interval value between DMMs was 100 ms (the default) and the repeat interval value between series was specified as 1 minute. DMMs carry timestamps to enable delay and jitter calculation.

According to the configuration, MEP 200 (located at switch WWP3) was supposed to receive DMMs and generate Delay Measurement Response (DMR) messages. DMR messages carry initial DMM timestamps and their own timestamps.

The second configuration line does the same job for monitoring CFM service pbt_501_cfm, which was enabled for the PBT tunnel with B-VID 501.

5.7.4.4 *Results*

The command `cfm delay show` reports instant results of delay/jitter measurement according to the latest series of DMR messages received.

The results of two successive `cfm delay show` commands applied to the WWP2 switch were as follows:

```
wwp2> cfm delay show

+-------------- MEP DELAY MEASUREMENT MESSAGE INFORMATION --------------------+
|            |    |Remote           |Remote|     |     | Delay  | Jitter |Rep |
|Service     |Port|Mac Address      |Mepid |DMM's|DMR's| in us  | in us  |Time|
+------------+----+-----------------+------+-----+-----+--------+--------+----+
|pbt_505_cfm |26  |04:00:00:00:00:03| 200  |10   |10   |3759    |81      |1   |
|pbt_501_cfm |25  |04:00:00:00:00:03| 201  |10   |10   |1043    |231     |1   |
+------------+----+-----------------+------+-----+-----+--------+--------+----+

wwp2> cfm delay show
+-------------- MEP DELAY MEASUREMENT MESSAGE INFORMATION --------------------+
|            |    |Remote           |Remote|     |     | Delay  | Jitter |Rep |
|Service     |Port|Mac Address      |Mepid |DMM's|DMR's| in us  | in us  |Time|
+------------+----+-----------------+------+-----+-----+--------+--------+----+
|pbt_505_cfm |26  |04:00:00:00:00:03| 200  |10   |10   |3787    |252     |1   |
|pbt_501_cfm |25  |04:00:00:00:00:03| 201  |10   |10   |853     |136     |1   |
+------------+----+-----------------+------+-----+-----+--------+--------+----+
```

The output shows that delay and jitter averages changed slightly from series to series but not by much.

The clocks of the WWP2 and WWP3 switches were synchronised by the local Network Time Protocol (NTP) server of the JANET(UK) development network edge router, which was synchronised with JANET NTP servers.

According to WWP/Ciena, DMM/DMR-based measurement provided by LE-311v switches is precise only to within milliseconds due to it being a relatively low-precision, software-based implementation. If implemented in hardware (as per WWP/Ciena's roadmap), then the precision should be within microseconds.

5.7.4.5 *Problems Encountered*

No problems were encountered

5.7.5 Test 4: Path Protection

5.7.5.1 *Test Objective*

The objective of this test was to check whether PBT supports fast switching from primary to backup tunnels in case of primary tunnel failure. Protection switching (also known as resilience) in SDH style is PBT's key distinguishing feature compared to, say, PBB, which can support traffic engineering like PBT, but does not support protection switching.

5.7.5.2 *Test Description*

The test consisted of the following steps:

1. Configure a backup tunnel for the primary compound tunnel wwp2-wwp1-wwp3 and initiate a CFM/CCM session for it.
2. Cause a connectivity break by physically disconnecting the inter-switch patch or by disabling a respective port.
3. Check the state of the primary and the backup tunnels by:
    a. Pinging connectivity between servers 10.1.0.1 and 10.1.0.2 in both directions with 0.1 sec intervals.
    b. Checking the status of each tunnel by the CLI command `tunnel show`.

5.7.5.3 *Configuration*

Configuring a PBT backup tunnel is very similar to configuring a primary one. A backup tunnel should be configured at both head-end switches and have the same B-VID as the primary tunnel. The backup tunnel should have the same name as the primary tunnel.

Configuration of WWP2 and WWP3 switches was as follows:

```
wwp2> config show
```

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! TUNNEL CONFIG:
!
tunnel encap create static-pbt to_wwp3 b-vid 505 dest-bridge-name wwp3 port 26
tunnel encap create static-pbt 501_to_wwp3 b-vid 501 dest-bridge-name wwp3 port
      25
tunnel encap create static-backup-pbt to_wwp3 b-vid 605 dest-bridge-name wwp3
      port 25
tunnel decap create static-pbt from_wwp3 b-vid 505 src-bridge-name wwp3 port 26
tunnel decap create static-pbt 501_from_wwp3 b-vid 501 src-bridge-name wwp3
      port 25
tunnel decap create static-pbt bkp-from-wwp3 b-vid 605 src-bridge-name wwp3
      port 25
!
tunnel pair create tnl-pair wwp3 encap-pbt to_wwp3 decap-pbt from_wwp3
tunnel pair create tnl-pair 501_tunnel encap-pbt 501_to_wwp3 decap-pbt
      501_from_wwp3
tunnel pair create tnl-pair wwp3-backup encap-backup-pbt to_wwp3 decap-pbt bkp-
      from-wwp3
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

wwp3> config show
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! TUNNEL CONFIG:
!
tunnel encap create static-pbt to_wwp2 b-vid 505 dest-bridge-name wwp2 port 25
tunnel encap create static-pbt 501_to_wwp2 b-vid 501 dest-bridge-name wwp2 port
      26
tunnel encap create static-backup-pbt to_wwp2 b-vid 605 dest-bridge-name wwp2
      port 26
tunnel decap create static-pbt from_wwp2 b-vid 505 src-bridge-name wwp2 port 25
tunnel decap create static-pbt 501_from_wwp2 b-vid 501 src-bridge-name wwp2
      port 26
tunnel decap create static-pbt bkp_from_wwp2 b-vid 605 src-bridge-name wwp2
      port 26
!
tunnel pair create tnl-pair wwp2 encap-pbt to_wwp2 decap-pbt from_wwp2
tunnel pair create tnl-pair 501_tunnel encap-pbt 501_to_wwp2 decap-pbt
      501_from_wwp2
tunnel pair create tnl-pair to_wwp2_backup encap-backup-pbt to_wwp2 decap-pbt
      bkp_from_wwp2
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

The CCM interval plays a crucial role in protection switching as switching starts after the third successive undelivered CCM. In the test the CCM interval was left at the default of 1 second.

5.7.5.4 *Results*

The protection switching mechanism worked during the test as expected. When port 28 of WWP1 was disabled, the status of the primary tunnel changed to Active-No, while the status of the backup tunnel changed to Active-Yes according to the output of the `tunnel show` command.

Pinging the connection with a 0.1 second interval showed a short interruption of connectivity. The protection switching caused 5% packet loss or 22 packets, which roughly corresponds to 2.2 seconds of switching time – less than the expected 3 seconds.

The pinging data was as follows:

```
--- 10.1.0.2 ping statistics ---
420 packets transmitted, 398 received, 5% packet loss, time 45117ms
rtt min/avg/max/mdev = 0.194/0.211/0.235/0.014 ms, pipe 2
```

5.7.5.5 *Problems Encountered*

No problems were encountered.

5.7.6 Test 5: Traffic Policing

5.7.6.1 *Test Objective*

The objective of this test was to investigate the ability of the WWP switches to do per-VLAN policing

5.7.6.2 *Test Description and Setup*

The test consisted of the following steps:

1. Configure a policing limit for ingress traffic defined by the port/VLAN pair.
2. Use the Agilent N2X traffic generator to send Ethernet frames to a chosen user port, starting from a level below the policing limit and gradually increasing it to exceed the policing limit.

The second Agilent port was used to collect frames that went through the policer and PBT tunnels.

The policing testbed is shown in Figure 5.16.

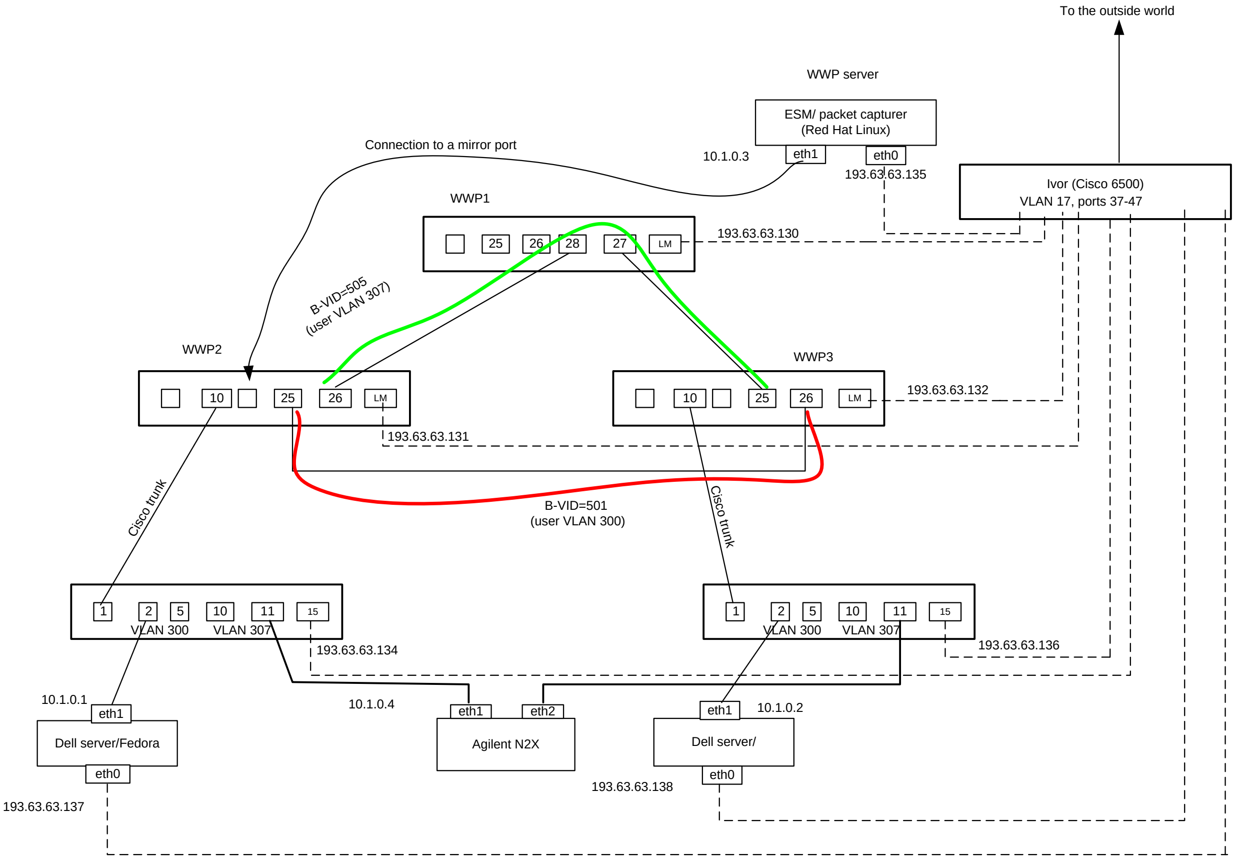


Figure 5.16: Policing testbed

5.7.6.3 *Configuration*

The configuring of a policer includes:

- Creating a policing profile on the basis of Committed Information Rate (CIR) and Peak Information Rate (PIR) for a port and VLAN ID.
- Enabling the profile.

The configuration of the WWP2 switch was as follows:

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! TRAFFIC PROFILING CONFIG:
!
traffic-profiling standard-profile create port 10 profile 1 cir 2048 pir 4032
      name tp_307 vlan 307
!
traffic-profiling set port 10 mode standard-vlan
traffic-profiling enable port 10
!
```

```
traffic-profiling enable
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

Policing was applied to ingress traffic at port 10 of the WWP2 switch with CIR = 2048 kbit/s and PIR = 4032 kbit/s. Traffic generated by the Agilent N2X box produced 64 octet long frames (including header) with 26 octets of IP payload. Such a small frame size was chosen due to it being more demanding for forwarding.

Policing was applied in VLAN mode which means that policed frames are selected on a user VLAN tag basis; this mode works only if an ingress port does MAC-in-MAC encapsulation.

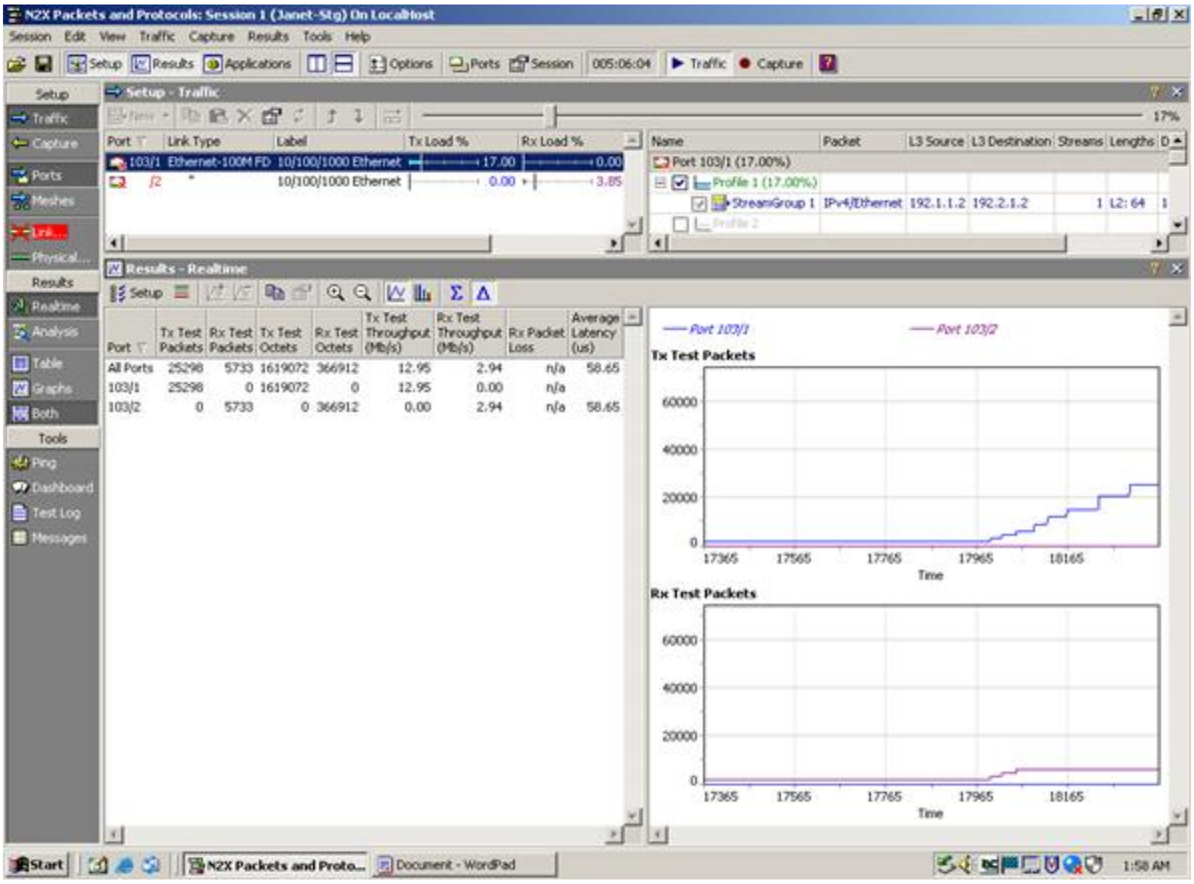
5.7.6.4 *Results*

When the offered load generated by the Agilent eth1 port ramped up from 1 Mbit/s to 4 Mbit/s (before the 17956 time mark on the screenshot in Figure 5.17), all traffic got through as both Wireshark (connected to port 26 of WWP1) and the Agilent eth2 port indicated the same rate of traffic.

When the offered load exceeded the PIR limit of 4 Mbit/s (after the 17965 time mark), the effect of frame dropping started to take place. The rate of traffic getting through the WWP1 switch (captured by Wireshark) and the rate of traffic received by the Agilent eth2 port both stayed unchanged, i.e. the traffic rate stalled at 4 Mbit/s.

Calculations showed that:

- CIR and PIR traffic meters count all the bytes of Ethernet frames, i.e. header and payload.
- Being applied in VLAN mode, a policer meter works on encapsulated packets so that it takes into account a full hierarchy of headers (i.e. all fields of a MAC-in-MAC header).



Port	Link Type	Label	Tx Load %	Rx Load %
103/1	Ethernet-100M FD	10/100/1000 Ethernet	17.00	0.00
/2	*	10/100/1000 Ethernet	0.00	3.85

Name	Packet	L3 Source	L3 Destination	Streams	Lengths	D
Port 103/1 (17.00%)						
☑ Profile 1 (17.00%)						
☑ StreamGroup 1	IPv4/Ethernet	192.1.1.2	192.2.1.2	1	L2: 64	1
□ Profile 2						

Port	Tx Test Packets	Rx Test Packets	Tx Test Octets	Rx Test Octets	Tx Test Throughput (Mb/s)	Rx Test Throughput (Mb/s)	Rx Packet Loss	Average Latency (us)
All Ports	25298	5733	1619072	366912	12.95	2.94	n/a	58.65
103/1	25298	0	1619072	0	12.95	0.00	n/a	
103/2	0	5733	0	366912	0.00	2.94	n/a	58.65

Figure 5.17: Agilent screenshot

5.7.6.5 *Problems Encountered*

No problems were encountered.

5.8 Conclusions

Generally, the tests conducted showed that the PBB-TE functionality of Ciena 5305, LE-311v and Extreme 12000 switches corresponded to expectations, which in turn were based on numerous presentations, white papers and demonstrations from the technology developers. The boxes demonstrated their proper behaviour in the following key areas:

- Manual establishment of TE point-to-point tunnels.
- Separation of customer and provider address spaces.
- CCM monitoring of state of tunnels.
- Fast protection switching.
- Per-VLAN traffic policing.

PBB-TE as a transport technology is better suited to single-domain applications, i.e. the environment for which it was designed. Multi-domain use of contiguous PBB-TE tunnels is possible, but it does not comply with the MEF specification of the External Network-to-Network Interface [MEF 26.1] and needs mutual knowledge of the MAC addresses of tunnel termination points, which goes beyond the normal method of management operations between domains. The vendors with whom the Task has met broadly agree with this assessment.

PBB-TE deployed in one domain can smoothly interoperate with EoMPLS deployed in other domains. The overlay model of this interoperation was tested and showed its robust nature.

Participants of the trial tend to agree with the current mainstream opinion that EoMPLS should be used in the core networks while PBB-TE could be used in access networks. (While EoMPLS only can be used in access networks, PBB-TE is much simpler and more easily understandable for people who deal mostly with LAN Ethernet, e.g. campus administrators.) PBB-TE could also be used in large campus networks, exploiting PBB-TE's traffic engineering features and fast protection switching whilst being applicable to the predominantly Ethernet-based expertise of many campus network managers.

6 MPLS-TP

6.1 Overview

This chapter describes the Multi-Protocol Label Switching Transport Profile (MPLS-TP) testing, which took place at the lab premises of Alcatel-Lucent (ALU), Ciena and Nokia Siemens Networks (NSN).

MPLS-TP is seen by some as a new technology and by others as simply the old, MPLS technology with additional features. It has proved divisive in other ways, too, notably in the standardisation process, which stalled because of disagreements between the IETF and ITU-T on Operation, Administration and Maintenance (OAM), leading to two different tracks of MPLS-TP OAM standards. The evolution and standardisation process is still ongoing.

The objectives of the trial were to test MPLS-TP and its features; identify the current status of the implementations, particularly with regard to OAM and control plane solutions; and identify areas for further study.

The aspects tested were MPLS-TP architecture, services, OAM, protection and control plane.

The objective of the MPLS-TP architecture test was to demonstrate the ability of the equipment to perform basic MPLS-TP functions, e.g. Label Switch Router and Label Edge Router functionality, and tunnel, Label Switched Path (LSP) and Pseudowire (PW) configuration. The results were positive, proving the ability of the equipment to perform basic MPLS-TP functionality according to the standard definitions. However, some functionality is still missing due to the early stage of the implementations. One example is ring protection, though it varies from vendor to vendor.

The objective of the services testing was to verify support of the three service types defined by the Metro Ethernet Forum: E-Line, E-LAN and E-Tree. Port-based services of the three service types were successfully configured and verified. It is recommended that testing is continued in GN3 Y4 to address VLAN-based services and to test the different protection schemes for every possible service.

The objective of the OAM tests was to verify the operation of both IETF and ITU-T versions of MPLS-TP OAM tools. The IETF tools used were ping and traceroute, in both LSPs and PWs; their functionality was successfully demonstrated and their importance for on-demand monitoring in production networks shown. The ITU-T tools used were Connectivity Verification (CV), Loopback Measurement and dual-ended Delay

Measurement, both proactive and on demand. The functionality of all three tools was verified in a lab demonstration; data from hands-on tests or live deployment is not available.

The objective of the protection tests was to verify 1+1 protection using Bi-directional Forward Detection (BFD), the IETF mechanism, and CV and Degraded Signal Defect (dDEG), which are ITU-T mechanisms. The results confirmed that the equipment supports 1+1 protection using all three mechanisms: ALU supports the ITU-T-based tools while Ciena and NSN support the IETF-based tools. The BFD polling rate is configurable, and is key to determining the protection switching time. Protection switching based on CV works in a similar way to protection switching with BFD. Protection switching based on signal degradation requires the preconfiguration of a threshold value; the protection mechanism is activated if the measured delay value exceeds this threshold.

The objective of the last test was to verify the equipment's control plane capabilities, namely, topology discovery, topology updates, and LSP/PW creation and deletion. These were demonstrated successfully, showing that MPLS-TP is able to operate with and without an NMS and that the use of the control plane is optional, complying with one of the main requirements of the MPLS-TP framework. The demonstration also proved that the implementation of key protocols (Open Shortest Path First, Resource Reservation Protocol – Traffic Engineering and Label Distribution Protocol) is mature enough for production environments.

The working sessions with the different vendors and the different demonstrations showed that MPLS-TP is a major focus in the industry. It is seen as the preferred technology to interface between OTN and the upper layers to deliver packet-based services providing transport features, and many vendors are integrating it into their transport equipment, albeit at different rates, with different priorities and with different (IETF or ITU-T) OAM solutions.

Areas identified for future work include MPLS/MPLS-TP interoperability; continued monitoring and investigation of new MPLS-TP developments; further services testing; and a more detailed investigation of control plane capabilities.

6.2 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. Despite the disputes around MPLS-TP, this technology has been a major focus for vendors (such as the ones who participated in the testing), 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 Operation, Administration and Maintenance (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.

The IETF is continuing its work on extensions of OAM tools such as BFD, while the ITU-T is working on the final approval of two standards: ITU-T G.8113.1 [ITU-T G.8113.1], which defines the OAM tools based on Y.1731, and ITU-T G.8113.2 [ITU-T G.8113.2], which defines OAM operations with IETF-based tools [ITU-T Newslog1, ITU-T Newslog2]. In the latest ITU-T meeting in Geneva (December 2011), the G.8113.1 standard was not approved and the decision was postponed until the next World Telecommunication Standardisation Assembly (WTSA-12). The approval requires a prior assignment by the IETF of an ACh code point, which would allow the identification of ITU-T OAM packets. However, the IETF's willingness to assign this code point depends on certain change requests regarding the title and content of the G.8113.1 standard [ITU-T Study Group 15: MPLS].

This chapter focuses on testing the feature set currently available, covering:

- MPLS-TP architecture.
- Services testing.
- OAM.
- Protection.
- Control plane.

For more information about the MPLS-TP technology, please refer to "Deliverable DJ1.1.1: Transport Network Technologies Study" [DJ1.1.1].

JRA1 Task 1 had difficulties obtaining access to equipment (due to the lack of budget in the project to buy test equipment and the lack of vendors willing to loan equipment to the project), but was able to participate in demonstrations at vendor premises. Due to time constraints the tests were not as exhaustive as originally planned. However, very valid and relevant information has been obtained during the process.

6.3 Test Objective

The main objective of this test was not only to test MPLS-TP and its features, but also to identify the current status of the implementations and to get a snapshot of the situation with regard to OAM and control plane solutions. In addition, it was the intention of the test to identify areas for further study during Y4 of the GN3 project, such as MPLS/MPLS-TP interoperability issues.

6.4 Test Infrastructure

6.4.1 Introduction

MPLS-TP was tested at the lab premises of Alcatel-Lucent (ALU), Ciena and Nokia Siemens Networks (NSN). In some cases the test format was more that of a lab demonstration. However, the information obtained was still highly relevant for the project. The test setups are described below. The level of detail has intentionally been kept to a minimum to comply with the Non-Disclosure Agreements (NDAs) signed with the vendors. A more detailed description of the test setups is not relevant for the purpose of the test.

6.4.2 Alcatel-Lucent Test Setup

The Alcatel-Lucent test platform was built with the 1850 Transport Service Switch (TSS) platform (Figure 6.2). The 1850 TSS is a multi-service platform capable of delivering legacy Synchronous Digital Hierarchy (SDH) services and MPLS-TP packet-based services. The test setup, shown in Figure 6.1, was built as a ring topology with 1 GE and 10 GE links. 1 GE and 10 GE interface were used as clients.

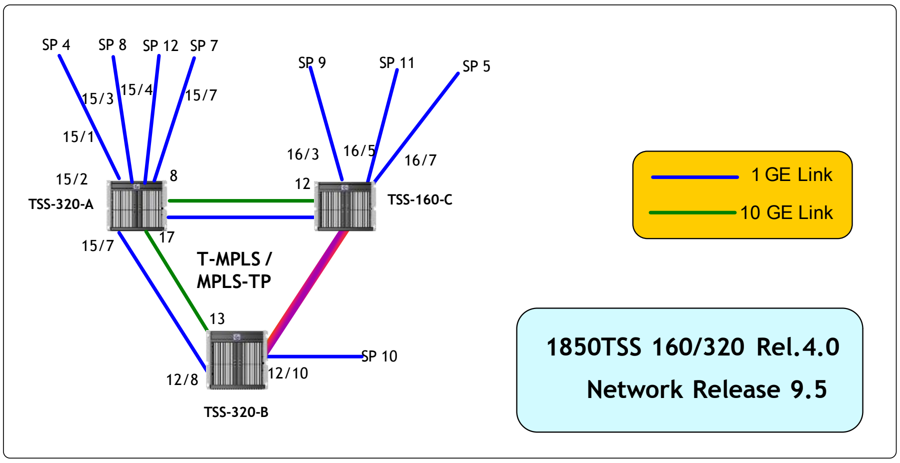
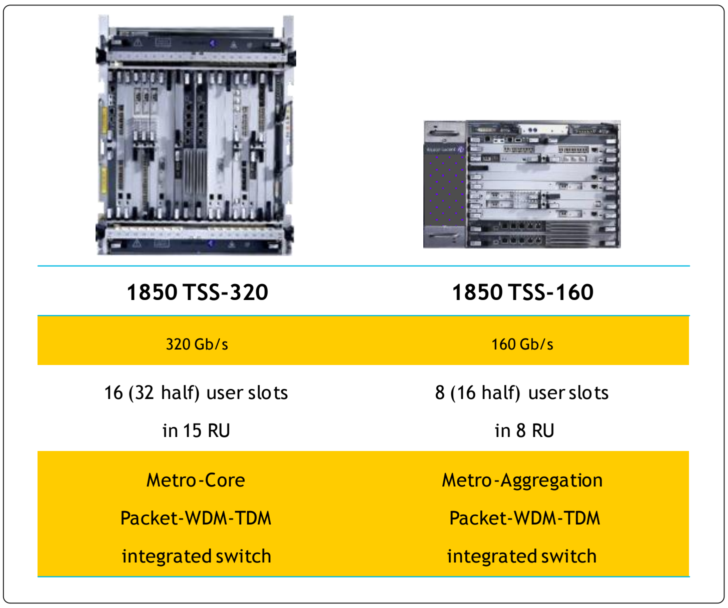


Figure 6.1: Alcatel-Lucent test setup



1850 TSS-320	1850 TSS-160
320 Gb/s	160 Gb/s
16 (32 half) user slots in 15 RU	8 (16 half) user slots in 8 RU
Metro-Core Packet-WDM-TDM integrated switch	Metro-Aggregation Packet-WDM-TDM integrated switch

Figure 6.2: Alcatel-Lucent TSS platform description

6.4.3 Ciena Test Setup

Ciena demonstrated MPLS-TP with their 5430 platform (Figure 6.3). The 5400 series is also a multi-service platform capable of providing Optical Transport Network (OTN)/SDH and packet-based switching. The platform was equipped with Hybrid Service Line Modules (HSLMs), which support both packet and Time-Division Multiplexing (TDM).

Figure 6.3: Ciena's 5430 platform

A logical representation of the setup used for the demo is shown in Figure 6.4. (The figure shows two nodes being connected with two links. However, the real setup was only a single node with loops in both links, which was nevertheless sufficient for the purpose of the demonstration.)

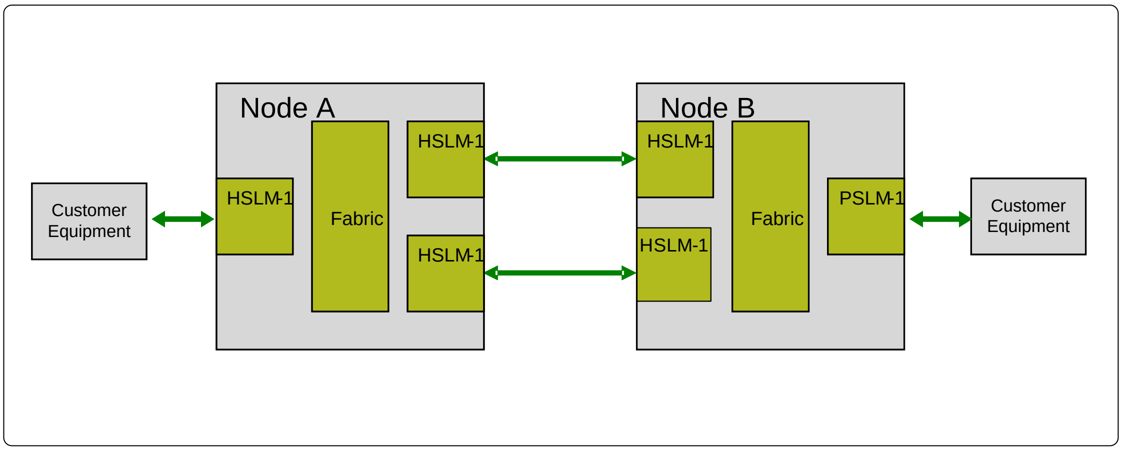


Figure 6.4: Logical representation of Ciena test setup

6.4.4 Nokia Siemens Networks (NSN) Test Setup

NSN used their new hiT 7100 platform (Figure 6.5) to demonstrate MPLS-TP. This platform provides optical and packet switching integration. The platform was equipped with the appropriate hardware for the purpose of the demonstration.

Figure 6.5: NSN's hiT 7100 platform

The test setup, shown in Figure 6.6, was a ring topology with dual GE links between the equipment. A tester was connected to one of the nodes. The nodes were remotely accessible.

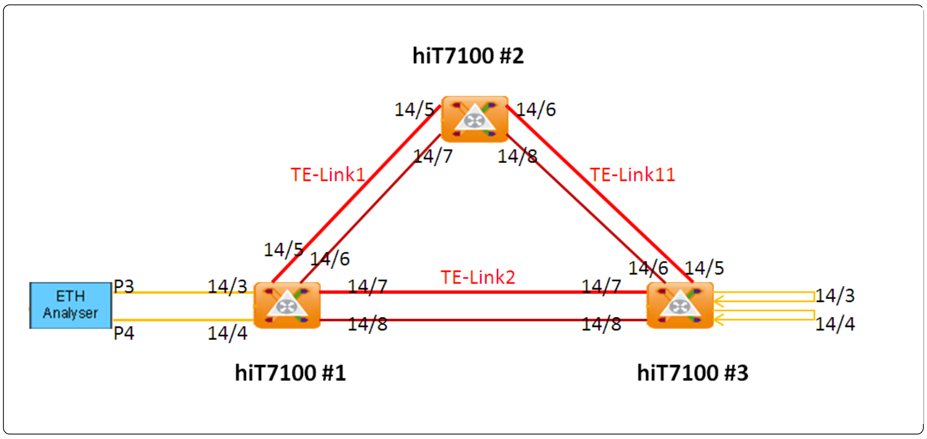


Figure 6.6: NSN test setup

6.5 MPLS-TP Architecture Test

6.5.1 Test Objective

The objective of the test was to demonstrate the ability of the equipment to perform basic MPLS-TP functions, e.g. Label Switch Router (LSR) and Label Edge Router (LER) functionality, tunnel (Label Switched Path (LSP)) configuration and Pseudowire (PW) configuration.

6.5.2 Test Setup

The setup used for this test is described in Section 6.4.

The equipment in the test setup is able to perform both LER and LSR functions. The LER is in charge of pushing the label into an incoming packet at the ingress of the network and popping it off at the egress. The LSR is a device capable of doing label switch routing, meaning that it routes the packets based on the labels.

When configuring a service it is necessary to configure the LSP (or tunnel) and the PW where the service, in this case a 1 GE, is carried (Figure 6.7). The actual procedure is in some cases vendor specific.

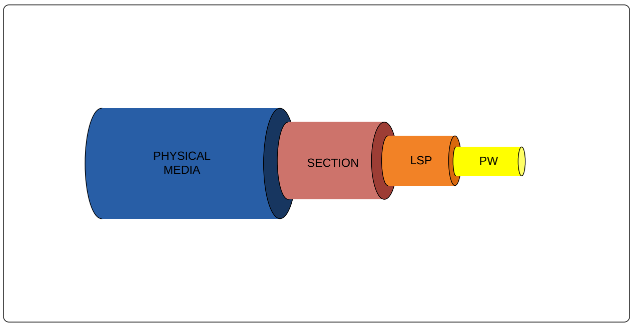


Figure 6.7: MPLS-TP architecture

6.5.3 Test Description

The test consisted of the following steps:

1. Create an LSP between the two end points.
2. Create a PW and map it to the LSP just created.
3. Create 1 GE service and map it to the PW just created.
4. Verify the configuration and assigned labels.
5. Verify the traffic flow in the testers.

6.5.4 Results

Basic MPLS-TP functionality was verified. It was possible to configure a service and verify the traffic flow between the two end points of the connection. It was also possible to verify the status of the LSP, the PW and the service via the Command Line Interface (CLI). Moreover, it was possible to see the label used by the specific LSP by looking at the information provided by the CLI.

As a second part of this test, a second service and PW were created and mapped to the same LSP. It was possible to verify this by checking the Label Forwarding Information Base (LFIB). Both PWs were multiplexed into the same LSP.

6.5.5 Test Conclusions

This test proved the ability of the equipment to perform basic MPLS-TP functionality according to the standard definitions. It was also demonstrated that the service configuration could be done via a Network Management System (NMS) or via a control plane (see Section 6.9).

6.6 Services Testing

6.6.1 Test Objective

The objective of this test was to verify the support of different types of services according to Metro Ethernet Forum (MEF) Technical Specification 6.1 “Ethernet Services Definitions – Phase 2” [MEF 6.1].

6.6.2 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 6.8 shows the three services. For E-LAN and E-Tree services a Virtual Bridge (VB) function is needed. This function is in charge of doing 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.

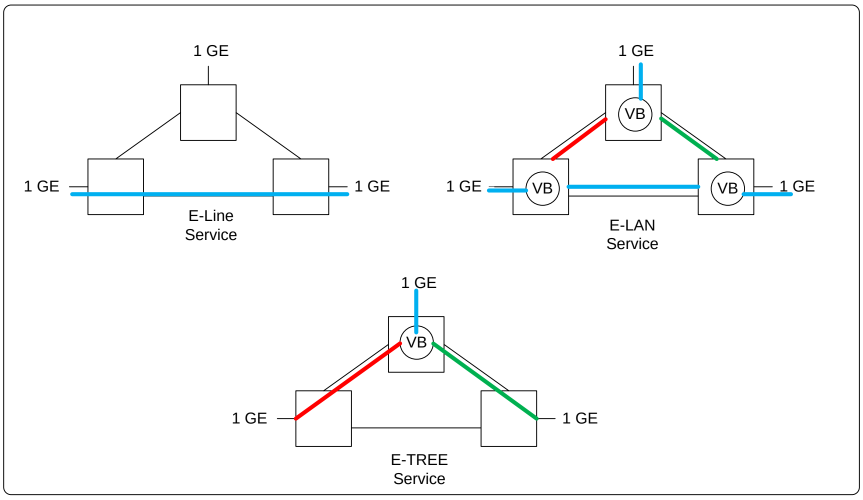


Figure 6.8: MEF services

6.6.3 Test Setup

The setup used for this test is described in Section 6.4.

6.6.4 Test Description

The test consisted of the following steps:

1. Configure an E-Line Service. This includes configuration of corresponding LSPs and PWs.
2. Verify that the traffic flows between the two end points.
3. Configure an E-LAN service. This includes configuration of corresponding LSPs and PWs.
4. Verify that traffic flows between the three end points.
5. Configure an E-Tree service. This includes configuration of corresponding LSPs and PWs.
6. Verify that traffic flows between the points.
7. Verify that traffic between the two spoke network elements goes via the hub (E-Tree service).

6.6.5 Results

The three services were configured and verified. Only port-based services were configured due to time restrictions. It was also possible to verify the FIB tables via CLI. This allowed verifying the direction in which the traffic was expected to flow depending on the destination MAC address.

6.6.6 Test Conclusions

While the port-based services of all three service types were successfully tested and verified, this test should be repeated during the next phase (i.e. in Year 4 of the GN3 project) with all the possible cases. Another interesting test would be to test the different protection schemes for every possible service.

6.7 OAM

The OAM tests covered both IETF and ITU-T versions of MPLS-TP OAM tools.

6.7.1 MPLS-TP OAM Tools – IETF Version

6.7.1.1 *Test Objective*

The objective of this test was to verify the operation of ping and traceroute tools in both LSPs and PWs.

6.7.1.2 *Technology Briefing*

LSP and PW ping will show whether the point in the network (Maintenance Entity Group End Point (MEP) or MEG Intermediate Point (MIP)) that is pinged is reachable; if it is, the monitoring point will echo a response to the originating point. The traceroute tool, however, will receive an echo of every monitoring point between the end points. The operation of these tools is similar to the well-known IP ping and traceroute tools [RFC4379].

6.7.1.3 *Test Setup*

The setup used for this test is shown in Figure 6.9. An LSP and a PW were created between the two end nodes to demonstrate OAM features. Then two MEPs and a MIP were created. These are the monitoring points in the network. A tester was connected to one end of the network and a loop was put in place at the other end.

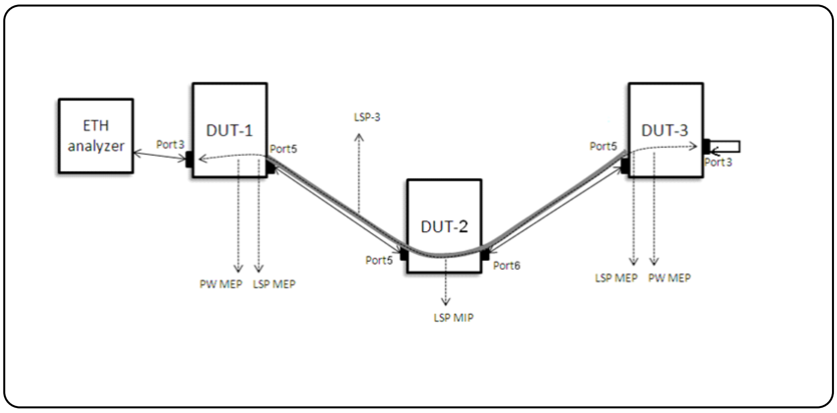


Figure 6.9: Setup for OAM testing

6.7.1.4 *Test Description*

The test consisted of the following steps:

1. Create a PW/LSP service between DUT-1 and DUT-3.
2. Enable MEP/MIP according to the topology.
3. Initiate PW ping from DUT-1 to DUT-3.
4. Initiate LSP ping from DUT-1 to DUT-3.
5. Initiate LSP traceroute from DUT-1 to DUT-3.
6. Initiate LSP traceroute from DUT-1 to DUT-3.

6.7.1.5 *Results*

The results were positive for all of the steps. PW and LSP ping showed the reachability of the MEP at the other end of the network. The traceroute tool for the LSP showed an echo from the MIP, while the PW traceroute only showed an echo from the last MEP as there was not a MIP at the PW layer.

6.7.1.6 *Test Conclusions*

The functionality of the IETF OAM tools was successfully demonstrated and their importance for on-demand monitoring in production networks shown.

6.7.2 MPLS-TP OAM – ITU-T Version

6.7.2.1 *Test Objective*

The objective of this test was to verify the operation of three ITU-T OAM tools: Connectivity Verification (CV), Loopback Measurement (LM) and dual-ended Delay Measurement (DM), both proactive and on demand.

6.7.2.2 *Technology Briefing*

CV, LM and DM belong to the other flavour of OAM tools that can be found in vendor implementations: those that are ITU-T based [ITU-T Y.1731].

CV (Figure 6.10) is a proactive OAM function running between two MEPs at a configured polling rate. CV can be used for different purposes like fault management, continuity supervision, performance monitoring and protection switching, as will be seen in Section 6.8.

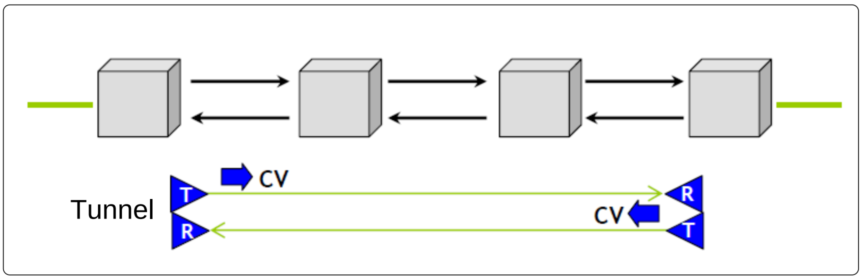


Figure 6.10: CV concept

LM (Figure 6.11) is an on-demand tool for packet loss measurement operated between two peer MEPs. A MEP periodically transmits LM packets indicating the amount of transmitted frames. The peer MEP answers with information about the amount of frames received and about the information received from the other MEP concerning transmitted frames and, finally, about the amount of frames that it is transmitting in the backward direction. The main purpose of this tool is performance monitoring.

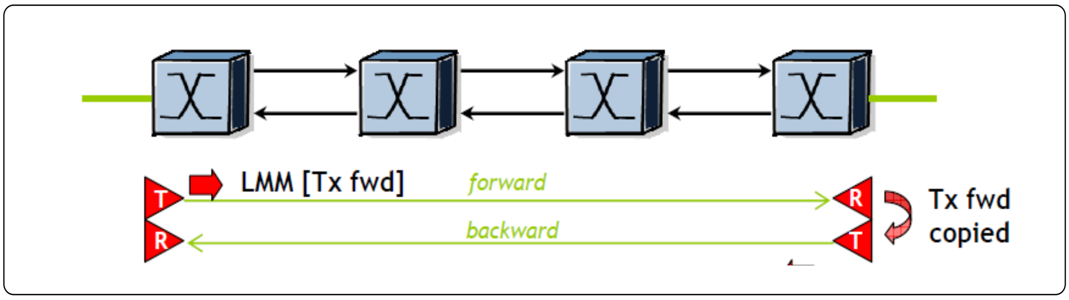


Figure 6.11: LM concept

Dual-ended DM (Figure 6.12) 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 measurements gives the delay variation. This tool is used for performance monitoring.

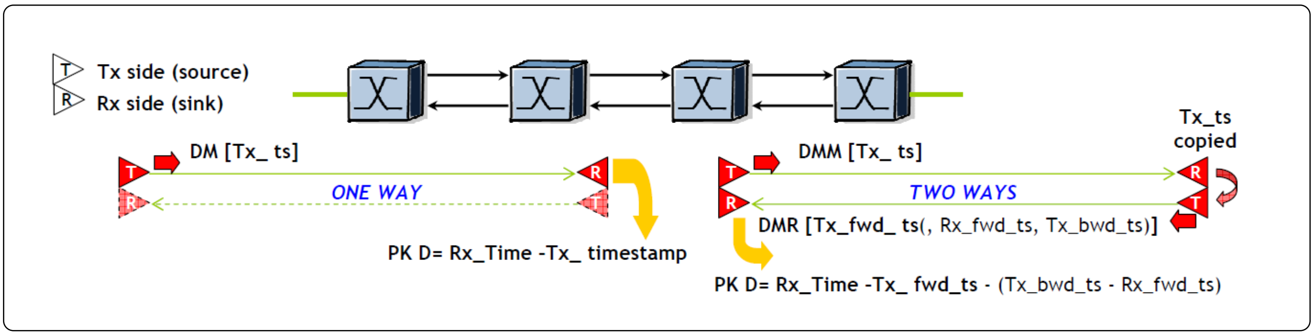

T Tx side (source)
R Rx side (sink)
DM [Tx_ ts]
ONE WAY
PK D= Rx_Time -Tx_ timestamp
DMM [Tx_ ts]
TWO WAYS
Tx_ts
copied
DMR [Tx_fwd_ ts(, Rx_fwd_ts, Tx_bwd_ts)]
PK D= Rx_Time -Tx_ fwd_ts - (Tx_bwd_ts - Rx_fwd_ts)

Figure 6.12: Dual-ended DM concept

6.7.2.3 *Test Setup*

The setup used in this test is described in Section 6.4.

A Network Management System (NMS) was used for the test. The different tools were triggered from the NMS GUI.

6.7.2.4 *Test Description*

The test consisted of the following steps:

1. Configure a service with the corresponding LSP and PW.
2. Configure the appropriate monitoring points (MEPs).

3. Test CV.
4. Test LM.
5. Test DM.

6.7.2.5 *Test Results*

The CV tool allowed end-to-end connectivity to be checked. The LM and DM tools returned the packet loss and delay measurement values respectively.

6.7.3 Test Conclusions

The functionality of three tools was verified. However, it was a lab demonstration and therefore no real data is available. CV and DM were also demonstrated in relation to 1 + 1 protection (described in Section 6.8).

6.8 Protection

The MPLS-TP standard defines protection mechanisms for both linear and ring topologies. However, at this point only 1+1 linear protection mechanisms have been implemented in the equipment to which JRA1 Task 1 had access. There are different implementations in the market with regard to the tools that are used to detect network failures. Moreover, there are also different implementations for switching traffic from the protected path to the protection path. One of these mechanisms is the well-known Automatic Protection Switching (APS) used in SONET and SDH networks.

6.8.1 1+1 Protection Using BFD

6.8.1.1 *Test Objective*

The objective of the test was to verify 1+1 protection using Bi-directional Forward Detection (BFD).

6.8.1.2 *Test Setup*

A logical representation of the setup used for this test is shown in Figure 6.13. (In reality there was only one network element with two loops on the network side.) Two tunnels were configured, the first one acting as the main path and the second one acting as the protection path.

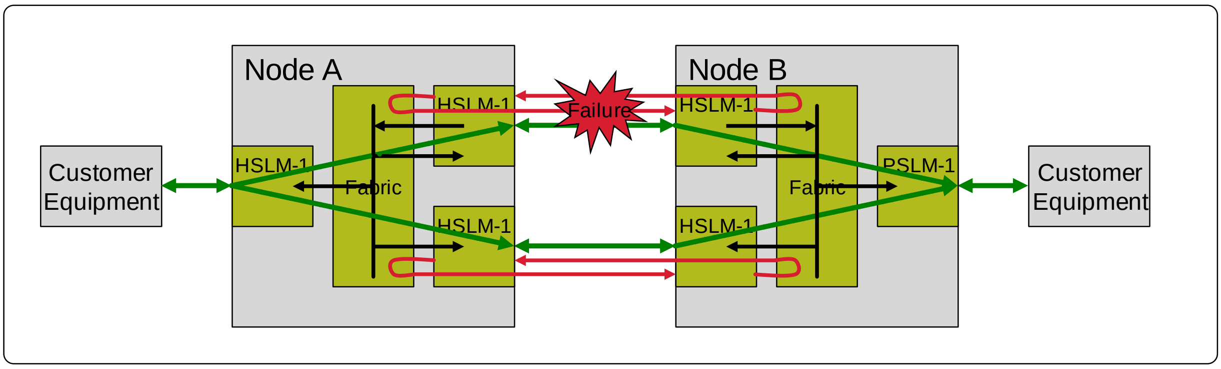


Figure 6.13: Logical representation of test setup for 1+1 protection with BFD

BFD was configured between the two monitored points. BFD acts as a heartbeat monitor, checking that the link is up. The BFD polling rate is configurable, and determines the protection switching time. It is therefore important to control this parameter.

Test equipment was used to check the traffic flow. To simulate the network outage, the fibre in the active path was removed.

6.8.1.3 *Test Description*

The test consisted of the following steps:

1. Configure an LSP.
2. Configure the protection LSP.
3. Configure the MEPs.
4. Configure BFD.
5. Check traffic is running in the active tunnel.
6. Simulate a network failure by removing the fibre.
7. Check traffic is running in the protection tunnel.

6.8.1.4 *Results*

1+1 protection worked as expected. The traffic was switched to the protection path upon detection via BFD of a failure. Because BFD was configured to a rate of 3.3 ms, protection switching took place in less than the 50 ms requirement inherited from SDH legacy networks.

6.8.2 1+1 Protection Using CV and dDEG

6.8.2.1 *Test Objective*

The objective of the test was to verify 1+1 protection using Connectivity Verification (CV) and Degraded Signal Defect (dDEG).

6.8.2.2 *Technology Briefing*

In this test, the 1+1 protection mechanism was implemented in a different manner. The OAM tools used to detect the failure in the network were CV and dDEG, which are ITU-T based [ITU-T Y.1731]. CV works in a similar way to BFD. The protection mechanism is activated upon detection of a lack of connectivity. dDEG, on the other hand, is based on delay measurements. A threshold value is preconfigured, and the switch to the protection path is triggered if the delay measurements detect degradation in the signal, i.e. if the measured delay value exceeds the pre-established threshold. DM is used for this purpose, to measure the delay on a specific connection proactively.

6.8.2.3 *Test Setup*

The setup used for this test is described in Section 6.4.2 and also shown in Figure 6.14 below.

Two different tests were performed. The first one used CV to detect a network failure. The second test used dDEG. The signal degradation was achieved by introducing a variable attenuator in the fibre between the two nodes.

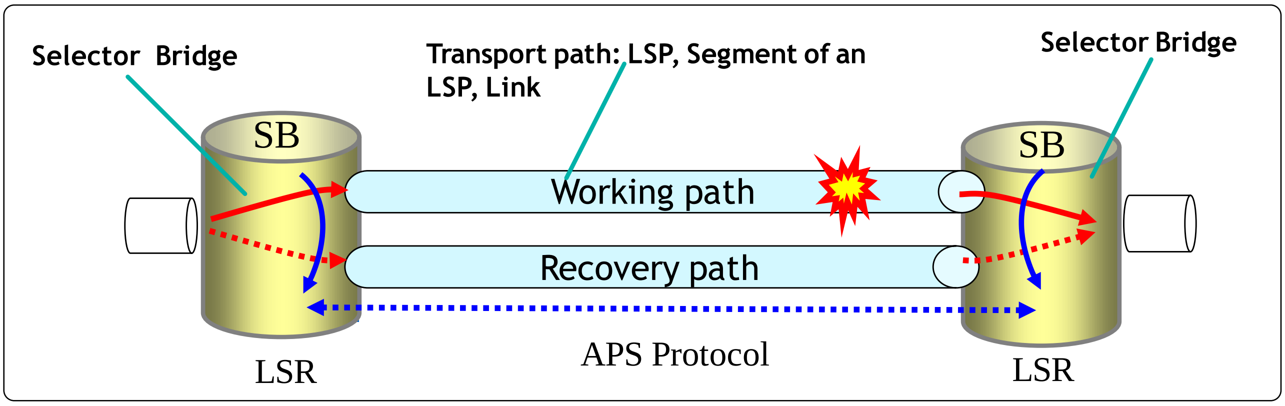


Figure 6.14: Setup for 1+1 protection with CV and dDEG

6.8.2.4 *Test Description*

The test consisted of the following steps:

1. Configure an LSP between two end nodes.
2. Configure the protection LSP.
3. Configure the MEPs.
4. Configure the OAM tools.
5. Verify the traffic flow in the active path.
6. Remove fibre to simulate an outage.
7. Verify traffic flow in protection path.
8. Force traffic to main path.
9. Simulate degradation in the signal by increasing the attenuation.
10. Verify that the traffic is switched to the protection path.

6.8.2.5 *Test Results*

Both tests gave positive results: 1+1 protection switching was successfully demonstrated using both CV and dDEG OAM tools.

6.8.3 Test Conclusions

The equipment supports 1+1 protection using BFD (Ciena and NSN), CV and dDEG (ALU). The BFD polling rate is configurable, and is key to determining the protection switching time. Protection switching based on CV works in a similar way to protection switching with BFD. Protection switching based on signal degradation requires the preconfiguration of a threshold value; the protection mechanism is activated if the measured delay value exceeds this threshold. These protections mechanisms are comparable to the 1+1 protection in OTN and SDH triggered by Signal Degradation (SD) and Signal Failure (SF) events.

6.9 Control Plane

6.9.1 Test Objective

The objective of this test was to verify the control plane capabilities.

6.9.2 Test Setup

The setup used for this test is shown in Figure 6.15. Three nodes were connected in a ring topology, with a tester for traffic verification.

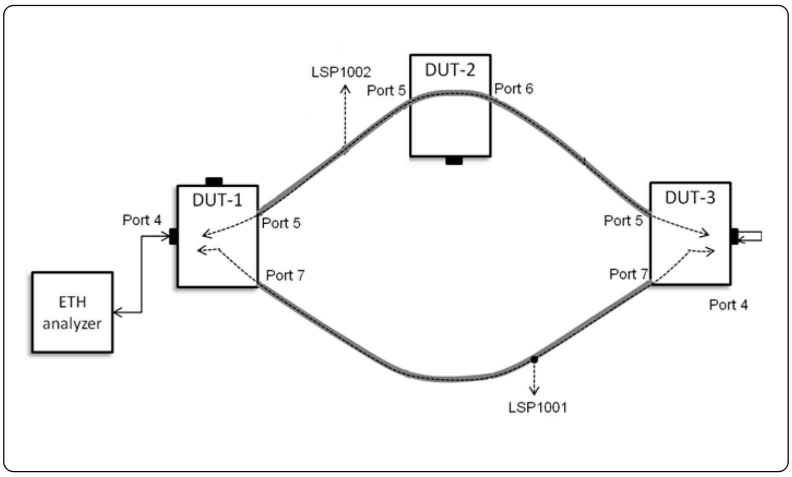


Figure 6.15: Control plane test setup

The test required previous configuration of the control plane. This included enabling the in-band Signalling Communication Network (SCN) and configuring Open Shortest Path First (OSPF).

6.9.3 Test Description

The test had four different parts and consisted of the following steps:

1. Topology discovery test:
    a. Enable SCN.
    b. Configure OSPF in all the nodes.
    c. Verify, via CLI, the routing table and that all nodes have topology information about the whole network.
2. Topology update:
    a. Configure a TE link between two nodes.
    b. Verify that the new TE link appears in the list of available resources in the OSPF database.
3. Tunnel/PW creation:
    a. Create an LSP and PW through Resource Reservation Protocol (RSVP) or Label Distribution Protocol (LDP).
    b. Verify LSP and PW via CLI.
    c. Check traffic flow with the tester.
4. Tunnel/PW deletion:
    a. Delete the LSP and PW through the control plane.
    b. Verify LSP and PW removal via CLI.
    c. Check that there is no more traffic with the tester.

6.9.4 Results

Topology discovery, topology updates, and LSP/PW creation and deletion were demonstrated successfully.

6.9.5 Test Conclusions

The control plane demonstration showed that MPLS-TP is able to operate with and without an NMS and that the use of the control plane is optional, complying with one of the main requirements of the MPLS-TP framework [RFC 5654]. The successful demonstration of topology discovery, topology updates, and LSP/PW creation and deletion proved that the implementation of OSPF, RSVP-TE and LDP protocols is mature enough for production environments. Due to time constraints, and to the fact that the test was a lab demo, there are no details on the protocol implementations and procedures.

6.10 Conclusions

The working sessions with the different vendors and the different demonstrations showed that MPLS-TP is a major focus in the industry. Not only Ciena, Alcatel-Lucent and NSN but many other vendors too are integrating MPLS-TP into their transport equipment. Since most of the equipment tested integrated both OTN and MPLS-TP in the same platform, it can be concluded that MPLS-TP is seen as the vendors' preferred technology to interface between OTN and the upper layers to deliver packet-based services providing transport features. Current vendor implementations are at different stages and it is apparent that each vendor has their own priorities. While some already have a control plane implementation, others have focused on OAM and survivability features.

OAM is currently one of the major topics, with vendors playing an important role in the discussion. Some vendors have decided to implement the ITU-T solution, while others advocate the IETF solution. The question about which one of these solutions will "win" still remains open. However, it seems likely that both solutions will coexist, at least for the time being. Despite the ITU-T/IETF issue, OAM is a very important requirement for transport networks and the test results showed that OAM adds value to the networks, providing the required transport grade functionality.

Another very important aspect, which JRA1 Task 1 has not been able to test, is MPLS/MPLS-TP interoperability. This is expected to be tested during GN3 Y4. In addition, the demonstrations showed that MPLS-TP is still under development; JRA1 Task 1 will therefore continue to monitor and investigate this technology during GN3 Y4. The Task also plans to repeat the services test with VLAN services, and to test the different protection schemes for every possible service, and to undertake a more detailed investigation of protocol implementations and procedures for control plane capabilities.

The use of MPLS-TP in the NREN community will depend on the type of NREN, the type of network and the kind of operational staff. For organisations with a tradition and experience of operating legacy transport networks like SDH, MPLS-TP will allow them to maintain transport capabilities such as OAM and survivability at the same time as transforming their networks into a more future-proof infrastructure. However, it is important to bear in mind that MPLS-TP is not the only alternative to SDH legacy networks. For organisations with a tradition of IP/MPLS networks, MPLS-TP will give them extra functionality in order to support and comply with the requirements for transport services at the same time as retaining their core functionality. In other words, MPLS-TP will allow them to add value to their networks without having to increase their operational staff.

7 OTN and GMPLS

7.1 Overview

The transport technology Optical Transport Network (OTN) was developed around ten years ago by the ITU-T and has been reviewed several times during recent years to adapt to new market needs, such as the ability to adapt all kinds of client signals, especially GE; more multiplexing granularity, with the addition of ODUflex; and higher bandwidths. In particular, equipment vendors have been developing new hardware known as OTN add-drop multiplexers or OTN switches. The main goals of this set of tests were to establish proof of concept, to learn as much as possible about the technology, to identify how NRENs can use it to advantage in their transport networks, and to trigger new and more specific discussions about OTN in the NREN community.

The following areas were covered during the testing:

- Optical Data Unit (ODU) switching: mapping of high-speed client signals; client signal mapping – ODU0 and ODUflex; and ODU switching.
- Survivability: protection and restoration.
- OAM: Tandem Connection Monitoring (TCM).
- Control plane (the scope was intentionally limited to Generalised Multi-Protocol Label Switching – GMPLS).

JRA1 Task 1 tested OTN technology in collaboration with ADVA Optical Networking, Ciena and Nokia Siemens Networks (NSN).

The main conclusions drawn from the ODU switching tests were as follows:

- The equipment was able to map high-speed client signals tested according to standard (G.709, G.Sup43) specifications. Options defined in the standard but not available at the time of the test are in the roadmap and will be available in the near future.
- The equipment was able to map 1 GE signals and signals with a bandwidth between 1 GE and 10 GE by using ODU0 and ODUflex.
- The equipment supports ODU switching functionality and multi-stage ODUk mapping, which offer NRENs a more flexible and integrated way to perform circuit switching in a large-scale, cross-domain,

IP-over-DWDM network environment, making it possible to deliver dedicated capacity with a high level of Quality of Service to the user.

- Taking into account that the minimum granularity is ODU0 and the possibility for ODUflex, the multiplexing possibilities of OTN give a lot of flexibility when providing connections in the network.

The conclusions drawn from the survivability tests were as follows:

- The equipment was able to support all the standard-defined (G.873.1) protection architectures tested.
- Mesh restoration, allowing 50 ms switching, could be of interest to NRENs carrying important traffic or networks with a complicated physical fibre infrastructure, because it allows extra protection compared to traditional SNCP. However, it requires up-front bandwidth reservation.
- In control plane restoration without mesh restoration, the new route can be chosen based on either the administrative cost or the total latency of the different links. The operator can control how the path is re-routed by manipulating these values per link. The latency can be measured using the new OTN features in the ODU overhead or can be manually configured, allowing the operator to control the traffic flow.
- Although the only way to observe dynamic protocol interaction in case of a link failure was via the management system logs, in the scenario tested (a dynamic wavelength restoration, supported by the GMPLS control plane) the GMPLS procedure in the restoration event was clear: the head end of the connection initiated a new RSVP session for establishing a new path.

The conclusions drawn from the TCM tests were as follows:

- The higher monitoring levels and more flexible surveillance possibilities offered by TCM make it potentially very beneficial in the NRENs' multi-domain, multi-vendor environment.
- The equipment behaved as expected in its defect detection and alarm handling in situations that are likely to occur in a multi-domain multi-vendor NREN environment.
- Utilising TCM functionality in a multi-vendor, multi-domain environment needs thorough planning and close collaboration between network operators.
- ITU-T G.709 introduces two additional potentially useful functions: the Fault Type, Fault Location (FTFL) channel, which helps to pinpoint the exact fault location when used in conjunction with the TCM functionalities, and delay measurement of ODUk path (DMp), which can be used to select the appropriate route in case of service restoration.

The conclusions drawn from the control plane (GMPLS) tests were as follows:

- The addition of a control plane to the OTN architecture provides new functionality – topology discovery and update, automated provisioning and decommissioning, and automated restoration – that covers the requirements for carrier-grade transport networks.
- The equipment tested supports topology discovery and update by means of the control plane routing protocol. The observed protocol interaction followed the standard. One instance of unexpected behaviour (no link information was exchanged on the failure of one of the routers) requires further investigation.

- The equipment supports signalling messaging during creation and tearing down of circuits. The observed protocol interaction followed the standard with one exception (Error code/value mismatch). During circuit decommissioning, the amount of unreserved bandwidth disseminated per link was smaller than when the service was created; this should be further analysed. However this could be related to a specific bug of this implementation.
- The equipment supports service restoration by means of GMPLS. The observed protocol interaction followed the standard.
- Automated restoration allows the reduction of pre-allocated bandwidth for protection and of manual intervention by a network operator to configure a backup path, thus reducing both CAPEX and OPEX in the network. However, the network design has to ensure that an alternative route with sufficient bandwidth is always available in case of a link failure, otherwise the connection cannot be re-established.
- Automated restoration is especially interesting used in combination with the recently added capability for measuring delay in the OAM bytes (please see Section 7.7.7.4).

OTN offers networks more powerful switching, mapping and survivability functionality in the digital domain compared to SDH. In addition, it brings seamless integration to the optical domain and provides a common vehicle for mapping, switching and transporting all types of client signals.

The testing of existing OTN platforms shows that the products are reaching market maturity. The most important functionalities, such as switching on different ODU levels and survivability based on different SNC parameters, including TCM, are already available. ODUflex is not fully implemented yet, but the tests confirmed that the basic functionalities to make ODUflex possible are working well. The integration of a control plane into the OTN technology adds important functionality and intelligence and opens up possibilities for dynamic provisioning tools integration, which is a major requirement for transport technologies in NRENs' transport networks.

OTN could provide major advantages to the NREN community. With capacity reaching – and soon to exceed – levels of 100G, a multiplexing and switching layer will be needed to gather the traffic into the core network in a flexible, scalable and standardised manner, at the same time being capable of mapping all types of client signals. Another advantage, given the multi-domain, multi-vendor NREN environment, is the fact that OTN is standardised and there is a common effort by all vendors to comply with the ITU-T recommendations. A standardised implementation of TCM across domains would provide NRENs deploying services across domains with full visibility and a common OAM infrastructure. A third advantage is that the integration of a control plane allows the nodes in the network to be aware of the total topology of the network and the capabilities of the different links, enabling automatic restoration of services in the best way possible by taking into account parameters such as cost and latency. Restoration by means of control plane capabilities avoids the need for immediate manual intervention in the case of a network outage.

7.2 Introduction

Optical Transport Network (OTN) technology is a transport technology that was developed around ten years ago by the ITU-T. However, OTN has been reviewed several times during the last few years to adapt to new market needs. Equipment vendors have been developing new hardware known as OTN add-drop multiplexers or OTN switches during the lifetime of JRA1 Task 1. For this reason, the Task has only had access to hardware at vendor premises, with the exception of a single vendor who loaned JRA1 Task 1 three Reconfigurable Optical Add-Drop Multiplexers (ROADMs) for a three-month period. However, the OTN functionality of these nodes was limited to OTN mapping of client signals. The testing carried out by JRA1 Task 1 has been done through close collaboration with vendors.

The testing and results described in this chapter are the continuation or follow-up of the work carried out during the first phase of the GN3 project. During that phase, OTN was identified as one of the most relevant and future-proof transport technologies. Most probably OTN will be – and in some cases already is – a technology that will play a very important role in next-generation NREN networks, providing multiplexing/switching, survivability, robust Operation, Administration and Maintenance (OAM) features and reliable transport of data traffic. This is the reason why JRA1 Task 1 has put a lot of effort into testing OTN during the second phase of the project. The main goal of this testing is to establish proof of concept and, primarily, to focus on the advantages of the technology itself. This chapter describes the test scenarios that were used during the testing and the results obtained.

The following areas were covered during the testing:

- Optical Data Unit (ODU) switching:
  - Mapping of high-speed client signals.
  - Client signal mapping: ODU0 and ODUflex.
  - ODU switching.
- Survivability: protection and restoration.
- OAM:
  - Tandem Connection Monitoring (TCM).
- Control plane (Generalised Multi-Protocol Label Switching – GMPLS).

JRA1 Task 1 was able to test OTN technology with the following vendors:

- ADVA Optical Networking. ADVA was tested at NORDUnet's lab premises and at the ADVA lab in Germany.
- Ciena. Ciena was tested at Ciena's lab premises in Atlanta, USA.
- Nokia Siemens Networks (NSN). NSN was tested at NSN's lab premises in Germany.

It is important to mention that the intention of this testing was to learn about the OTN technology and not to find weaknesses of the equipment under test or to compare the different hardware that was used during the testing.

7.3 Technology Briefing

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]. To provide the technological context, a short introduction to the aspects under test has been provided in each section.

7.4 Test Objective

As already mentioned in the *Introduction* to this chapter, OTN has been under discussion in recent years. Many commercial network providers are nowadays considering the use of OTN transport and OTN switching in their transport networks. AT&T in the USA, Deutsche Telecom, Telefónica and BT in Europe are examples of commercial operators that will replace their Synchronous Digital Hierarchy (SDH) legacy networks with OTN transport. Many NRENs in Europe are also in the process of planning, designing and procuring their next-generation transport networks; OTN is therefore being considered and is one of the main candidates. The primary purpose of this Task is to test OTN and its capabilities to learn as much as possible about it. It is also the intention to identify how NRENs can take advantage of the use of OTN in their transport networks. Moreover, it is the purpose of this chapter to trigger new and more specific discussions about OTN in the NREN community.

7.5 Test Infrastructure

7.5.1 Introduction

As mentioned in the *Introduction*, JRA1 Task 1 was able to test OTN technology with three different vendors. The test infrastructure used in each case was quite similar, and is described below. To comply with the Non-Disclosure Agreements (NDAs) signed with the vendors, the description of the test infrastructure has been kept as simple as possible.

7.5.2 ADVA Networks Test Infrastructure

NORDUnet was able to test ADVA equipment over a three–month period at NORDUnet's own premises. The test setup included three ADVA Fibre Service Platform (FSP) 3000 R7 ROADMs connected in a ring topology as shown in Figure 7.1. The nodes were managed by ADVA's NMS installed on a computer. All nodes were connected to a central Ethernet switch so it was possible to monitor GMPLS and the management traffic.

The ROADMs were equipped with 10 GE and 1 GE client interfaces:

- 4 x WCC-PCTN-10G

- 2 x 10TCC-PCTN-10G

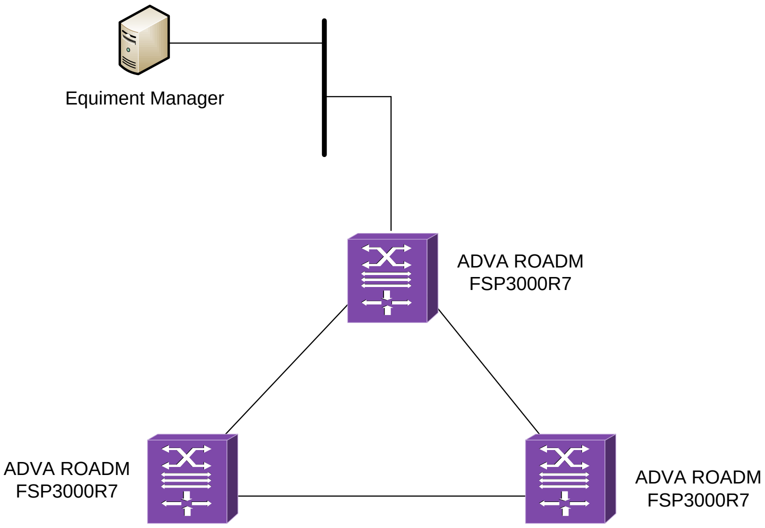


Figure 7.1: Test infrastructure used with ADVA

7.5.3 Ciena Test Infrastructure

The Ciena test was carried out at Ciena's premises in Atlanta, USA. The test setup was a mesh topology with four nodes from Ciena's 5400 family of Reconfigurable Switching Systems (RSSs). This equipment was equipped with the necessary hardware to provide OTN switching functionality. The equipment was connected as shown in Figure 7.2.

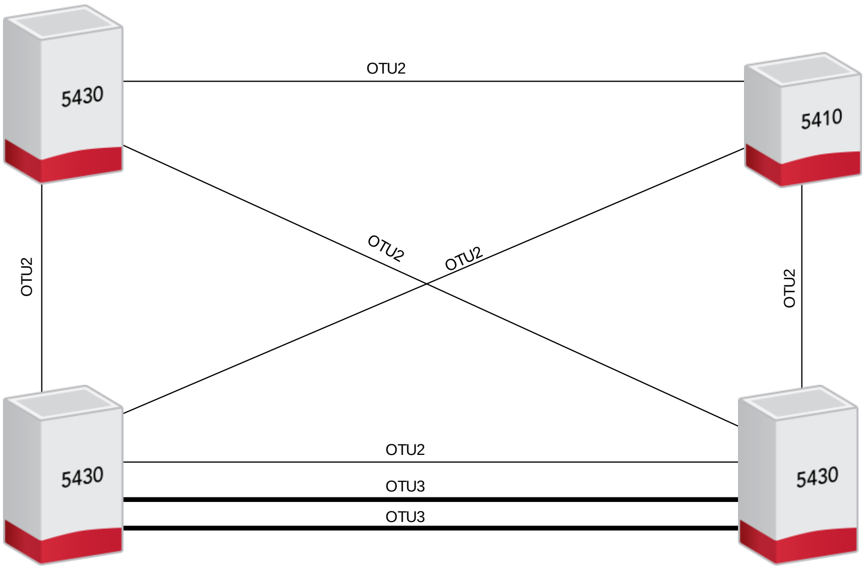


Figure 7.2: Test infrastructure used with Ciena

7.5.4 Nokia Siemens Networks (NSN) Test Infrastructure

The NSN test infrastructure used during the OTN testing is shown in Figure 7.3. The setup was formed by three OTN switches (Surpass hiT 7100 R1.0) and SDH add-drop multiplexers (Surpass hiT 7065). The OTN switches were connected in a ring topology.

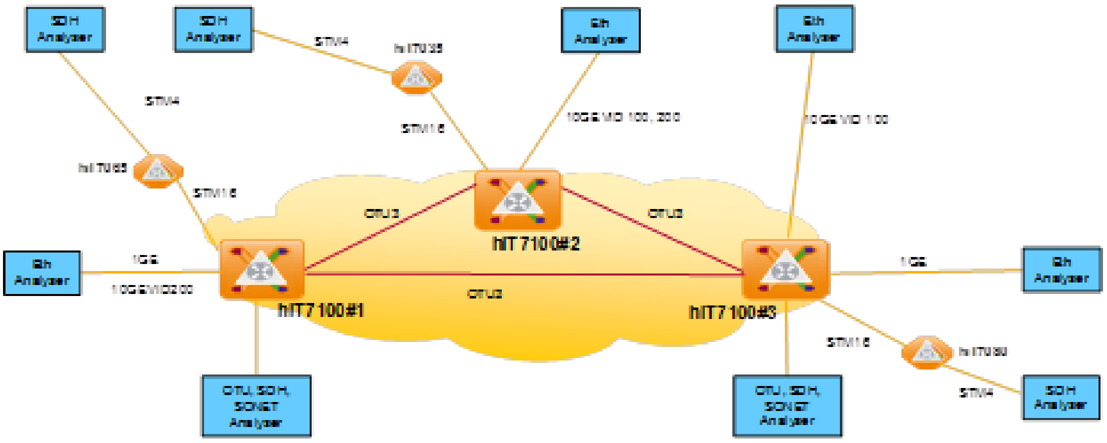


Figure 7.3: Test infrastructure used at NSN

7.6 ODU Switching

7.6.1 Mapping of High-Speed Client Signals

7.6.1.1 *Technology Briefing*

Initially the G.709 [ITU-T G.709] standard was developed to carry legacy Synchronous Optical Network / Synchronous Digital Hierarchy (SONET/SDH) traffic, but due to market demands and the evolution of transport networks it was decided that OTN had to be extended to support Ethernet. As shown in Figure 7.4, the G.709 multiplexing structure was extended with new formats in order to solve this challenge. The multiplexing hierarchy was extended to support the mapping of 1 GE, 10 GE, 40 GE and 100 GE into OTN.

Some of the Optical Transport Units (OTUs) had to be modified to support the mapping of certain interfaces, such as 10 GE-LAN. For this reason the ITU-T issued Supplement document G.Sup43 [ITU-T G.Sup43], which defined ODU2e/OTU2e and ODU3e/OTU3e (over clocked 44 Gbit to carry 4 x 10 GE-LAN). There are two solutions for ODU3e/OTU3e: the first one named ODU3e1, based on Asynchronous Mapping Procedure (AMP), and a second one named ODU3e2, based on Generic Mapping Procedure (GMP). For 100 Gbit OTN transport, the ITU-T has defined the ODU4/OTU4 and has chosen a bit rate of 111.809974 Gbit/s, in order to provide 80 tributary slots for up to 80 ODU0 and up to 10 ODU2e. Not only the hierarchy but also the mapping procedures were extended to include GMP, as this procedure had some advantages compared to existing ones due to its elasticity to bit-rate differences in the client signals. The evolution will not stop there, as the industry is already planning the extension of the structure by adding an OTU5. The question that still needs an answer is which signal should be mapped: 400 Gbit/s or 1 Tbit/s?

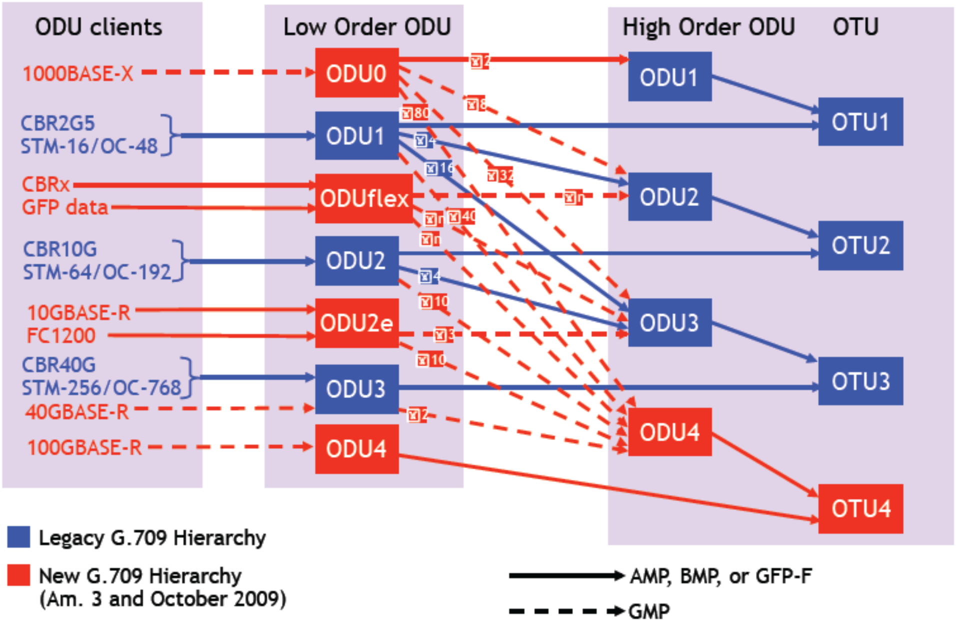
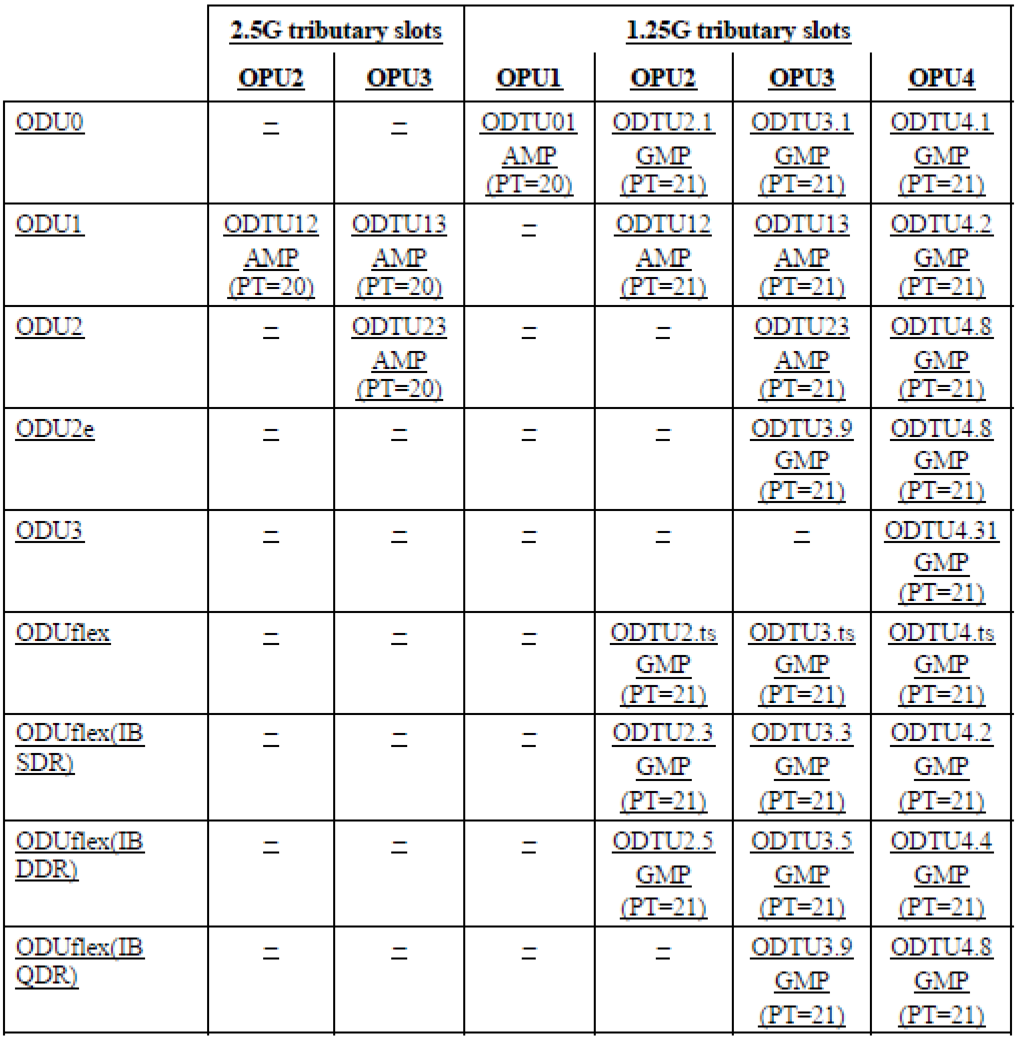


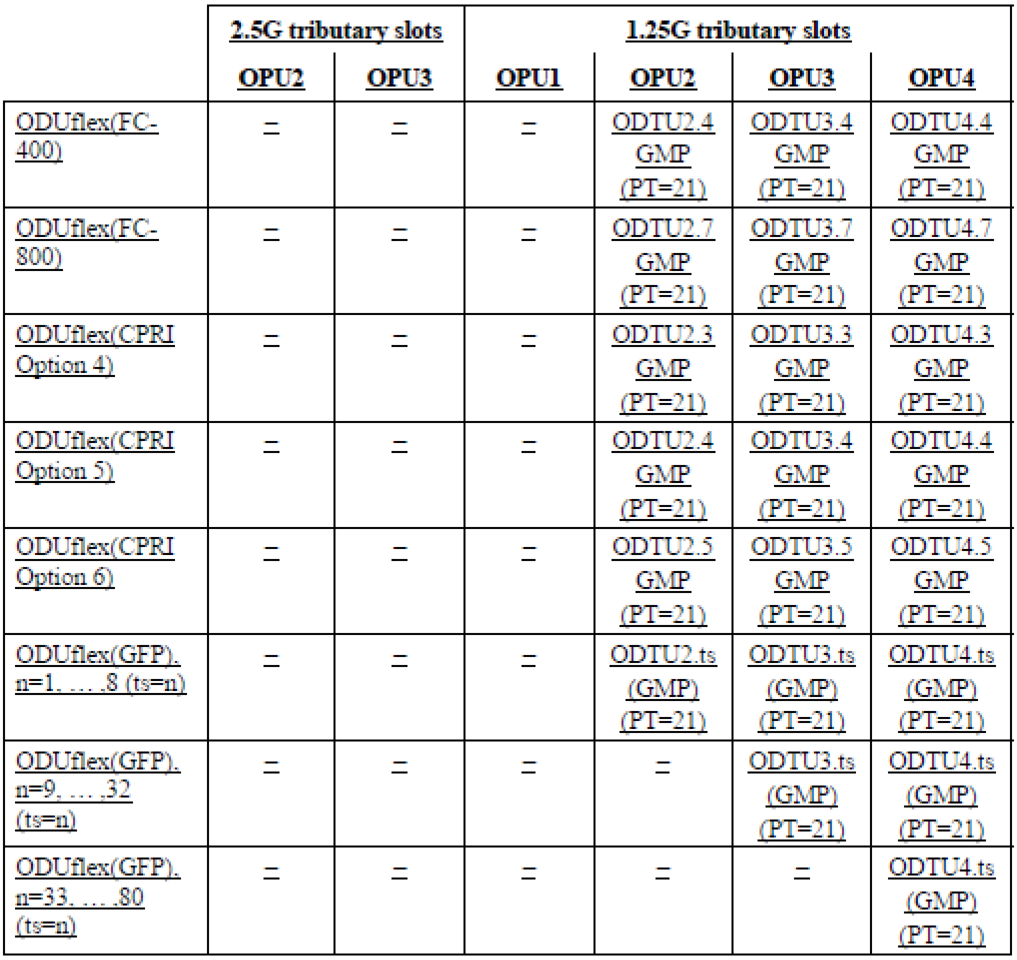
Figure 7.4: G.709 OTN multiplexing structure

The table shown in Figure 7.5 and Figure 7.6 below shows the mapping procedures used to map the different flavours of Low Order (LO) ODUk into High Order (HO) Optical Payload Unit (OPUk) and the different payload type (PT) used in each case.



	2.5G tributary slots		1.25G tributary slots			
	OPU2	OPU3	OPU1	OPU2	OPU3	OPU4
ODU0	–	–	ODTU01 AMP (PT=20)	ODTU2.1 GMP (PT=21)	ODTU3.1 GMP (PT=21)	ODTU4.1 GMP (PT=21)
ODU1	ODTU12 AMP (PT=20)	ODTU13 AMP (PT=20)	–	ODTU12 AMP (PT=21)	ODTU13 AMP (PT=21)	ODTU4.2 GMP (PT=21)
ODU2	–	ODTU23 AMP (PT=20)	–	–	ODTU23 AMP (PT=21)	ODTU4.8 GMP (PT=21)
ODU2e	–	–	–	–	ODTU3.9 GMP (PT=21)	ODTU4.8 GMP (PT=21)
ODU3	–	–	–	–	–	ODTU4.31 GMP (PT=21)
ODUflex	–	–	–	ODTU2.ts GMP (PT=21)	ODTU3.ts GMP (PT=21)	ODTU4.ts GMP (PT=21)
ODUflex(IB SDR)	–	–	–	ODTU2.3 GMP (PT=21)	ODTU3.3 GMP (PT=21)	ODTU4.2 GMP (PT=21)
ODUflex(IB DDR)	–	–	–	ODTU2.5 GMP (PT=21)	ODTU3.5 GMP (PT=21)	ODTU4.4 GMP (PT=21)
ODUflex(IB QDR)	–	–	–	–	ODTU3.9 GMP (PT=21)	ODTU4.8 GMP (PT=21)

Figure 7.5: Overview of LO ODU into HO OPU mapping types



	2.5G tributary slots		1.25G tributary slots			
	OPU2	OPU3	OPU1	OPU2	OPU3	OPU4
ODUflex(FC-400)	–	–	–	ODTU2.4 GMP (PT=21)	ODTU3.4 GMP (PT=21)	ODTU4.4 GMP (PT=21)
ODUflex(FC-800)	–	–	–	ODTU2.7 GMP (PT=21)	ODTU3.7 GMP (PT=21)	ODTU4.7 GMP (PT=21)
ODUflex(CPRI Option 4)	–	–	–	ODTU2.3 GMP (PT=21)	ODTU3.3 GMP (PT=21)	ODTU4.3 GMP (PT=21)
ODUflex(CPRI Option 5)	–	–	–	ODTU2.4 GMP (PT=21)	ODTU3.4 GMP (PT=21)	ODTU4.4 GMP (PT=21)
ODUflex(CPRI Option 6)	–	–	–	ODTU2.5 GMP (PT=21)	ODTU3.5 GMP (PT=21)	ODTU4.5 GMP (PT=21)
ODUflex(GFP), n=1, … ,8 (ts=n)	–	–	–	ODTU2.ts (GMP) (PT=21)	ODTU3.ts (GMP) (PT=21)	ODTU4.ts (GMP) (PT=21)
ODUflex(GFP), n=9, … ,32 (ts=n)	–	–	–	–	ODTU3.ts (GMP) (PT=21)	ODTU4.ts (GMP) (PT=21)
ODUflex(GFP), n=33, … ,80 (ts=n)	–	–	–	–	–	ODTU4.ts (GMP) (PT=21)

Figure 7.6: Overview of LO ODU into HO OPU mapping types (cont.)

7.6.1.2 *Test Setup*

The setup used for this test is described in Section 7.5 *Test Infrastructure* on page 178. The nodes were equipped with client interface cards capable of mapping different types of client signals.

7.6.1.3 *Test Objective*

The objective of the test was to verify that the equipment is able to map the following client signals according to standard specifications:

- OC-48/STM-16 into ODU1.
- 10 GE into ODU2e.
- OC-192/STM-64 into ODU2.
- 40 GE into ODU3.

- 100 GE into ODU4.

7.6.1.4 *Test Description*

The test consisted of the following steps:

1. Configure services to map client signals as shown in Figure 7.7.
2. Use a tester to prove that the traffic flows between the end points of the connection.

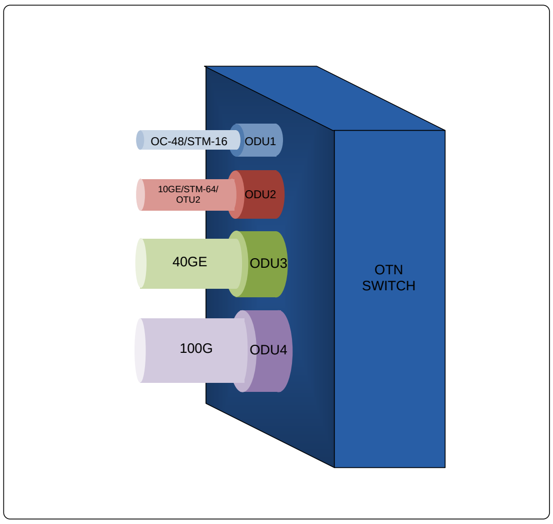


Figure 7.7: Mapping of high speed client signals

7.6.1.5 *Results*

The results of the test were positive. The client cards were able to map the different client signals according to the definitions in the standard for OTN. Although not all the options defined in the standard were available at the time of the test, they are in the roadmap and will be available in the near future.

The following mapping of client signals was verified:

- Mapping of Constant Bit Rate (CBR) signals STM-16/OC48 and STM-64/OC196 using AMP and BMP into ODU1 and ODU2 respectively.
- Mapping of 10 GE signal into ODU2 using Generic Framing Procedure (GFP) mapping.
- Mapping of 10 GE into ODU2e using BMP mapping procedure.
- Mapping of 40 GE into ODU3 using Generic Mapping Procedure (GMP).

Some currently available client signals, such as Fibre Channel, were not verified as they are not used in NREN networks and so were not considered relevant for the purpose of the test. The reason was because these type of client signals like for example, Fiber Channel are not used in NRENs networks.

7.6.2 ODU0 and ODUflex

7.6.2.1 *Technology Briefing*

As explained in Section 7.6.1.1 above, the OTN client-mapping capabilities were extended to support new requirements. ODU0 was added to support efficient transparent transport of GE signals over OTN. The reason for the addition was that if a GE were to be mapped into an ODU1, half of the bandwidth would be wasted. However, the ODU0 container is half the size of an ODU1. This allows the mapping of two ODU0s into an ODU1 (as shown in Figure 7.8), which is then mapped into an OTU1. Notice that OTU0 does not exist, as defined by the ITU-T G.709 standard.

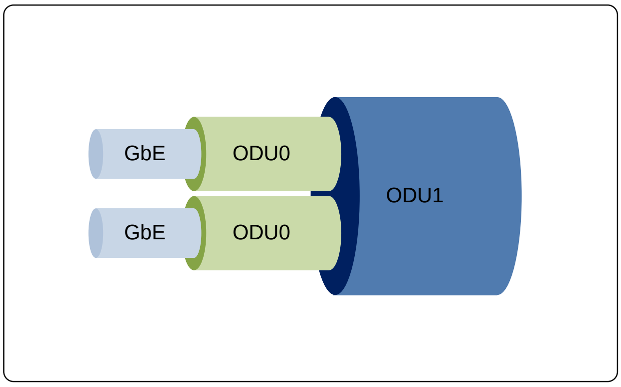


Figure 7.8: OTN ODU0

Additionally, ODUflex was developed to accommodate signal rates of different speeds so that it occupies the minimum number of time slots in a higher order ODUk. The client mapping is done so that the ODUflex container has the exact size of its client, leaving the remaining space for other client signals, as shown in Figure 7.9. ODUflex supports Constant Bit Rate (CBR) clients and packet-based clients. CBR clients are mapped by using Bit-Synchronous Mapping Procedure (BMP) and packet-based client signals are accommodated by using Frame-mapped Generic Framing Procedure (GFP-F). Then ODUflex is mapped into a number of time slots in a High Order ODU (HO ODU) by using Generic Mapping Procedure (GMP).

The introduction of ODUflex has also had a positive impact on the overall capacity of the network, as the resources are better utilised. For example, an OTU2 has 8 time slots of Payload Type 21 (1.25 Gbit/s). If a client signal of 4 Gbit/s were to be carried by the OTU2 without the use of ODUflex, only 4 time slots would be used and the rest would be wasted. ODUflex allows the remaining time slots to be used to carry other signals in the same OTU2.

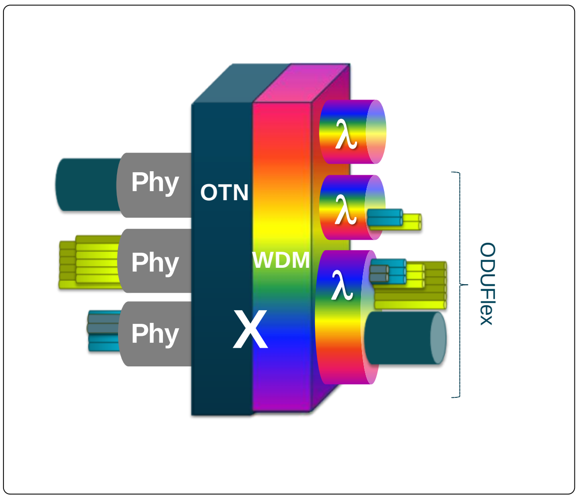


Figure 7.9: OTN ODUflex

7.6.2.2 *Test Setup*

The setup used during the testing is described in Section 7.5 *Test Infrastructure* on page 178.

7.6.2.3 *Test Objective*

The objective of the test was to verify the ability of the equipment to map 1 GE signals and signals with a bandwidth between 1 GE and 10 GE by using ODU0 and ODUflex.

7.6.2.4 *Test Description*

The test consisted of the following steps:

1. Configure a 1 GE service where the client signal is mapped into an ODU0.
2. Configure a second GE in the second available ODU0.
3. Map the ODU0s carrying the GE client signals into an ODU1.
4. Configure a client signal with a bandwidth between 1 GE and 10 GE and map the signal into a higher order ODU.
5. Configure a second client signal with a bandwidth equal to the remaining capacity in the ODU previously used.

7.6.2.5 *Results*

1 GE Mapping into ODU0

A 1 GE client signal was mapped into an ODU0 and transported through the network. The test was done by creating an OTN service with 1 GE as a client interface. A tester was connected at each end point to test the traffic flow between them. The 1 GE signal was mapped by using Timing Transparent Transcoding (TTT) and Generic Mapping Procedure (GMP).

ODUflex Mapping

At the time of the test, none of the vendors had ODUflex implementation in production equipment. However, the Task participants were able to see a lab demonstration. Figure 7.10 shows the setup used during the demonstration.

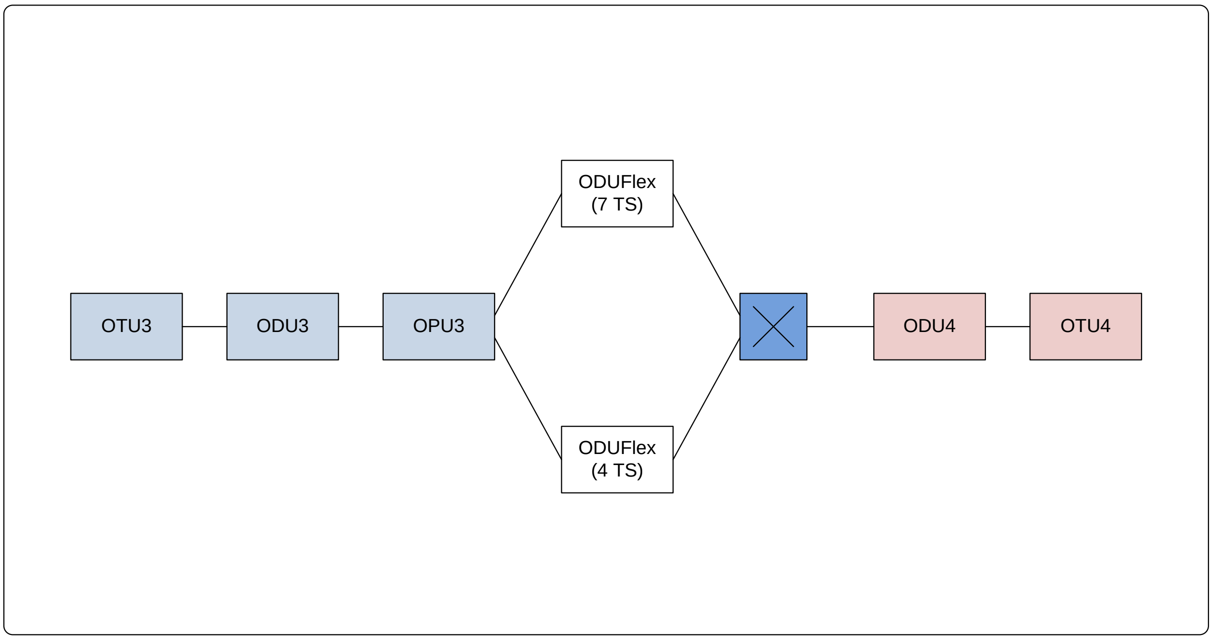


Figure 7.10: ODUflex demonstration setup

The demonstration verified the equipment's ability to split the signal into time slots of 1.25 Gbit/s and build a signal with bandwidth equal to 11 x Time Slot (TS) = 13.75 Gbit/s (the original signal bandwidths were 4 x TS = 5 Gbit/s and 7 TS = 8.75 Gbit/s). The new signal was then mapped into an ODU4 by using GMP.

ODUflex will be used, when available, to map CBR signals such as Fibre Channel (FC)-400, FC-800, InfiniBand Single Data Rate (IB SDR) or InfiniBand Double Data Rate (IB DDR) and packet-based signals such as Multi-Protocol Label Switching – Transport Profile (MPLS-TP) traffic flows. ODUflex is able to map any bandwidth; however, the recommendation is to map traffic bandwidths that are a multiple of the minimum time-slot size, which is approximately 1.25 Gbit/s.

7.6.3 ODU Switching and Multi-Stage ODUk Mapping

7.6.3.1 *Technology Briefing*

ODU switching means the ability to extract an arbitrary level of ODUk from any line interface and switch it to any direction over an ODUk of the same or higher level. Multi-stage mapping is the ability to map ODUk into a higher order ODUk several times. ODUk switching will enable a more flexible and integrated way to perform circuit switching in an IP-over-Dense Wavelength-Division Multiplexing (DWDM) network. It will make it possible to deliver dedicated capacity with a high level of Quality of Service (QoS) to the user. The ODU switching concept is illustrated in Figure 7.11.

Considering the large-scale cross-domain capabilities of the European NREN community, a scenario in which connectivity across network boundaries is established via an OTN switch is regarded as highly interesting and very advantageous.

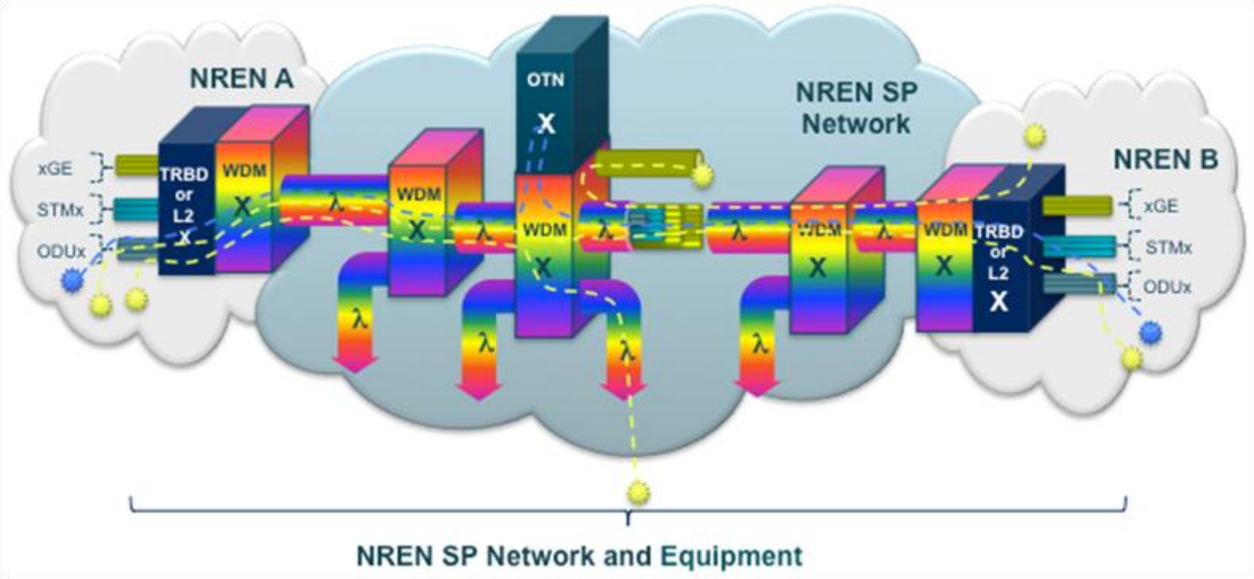


Figure 7.11 ODU switching concept

7.6.3.2 *Test Setup*

The setup used for this test is shown in Figure 7.12.

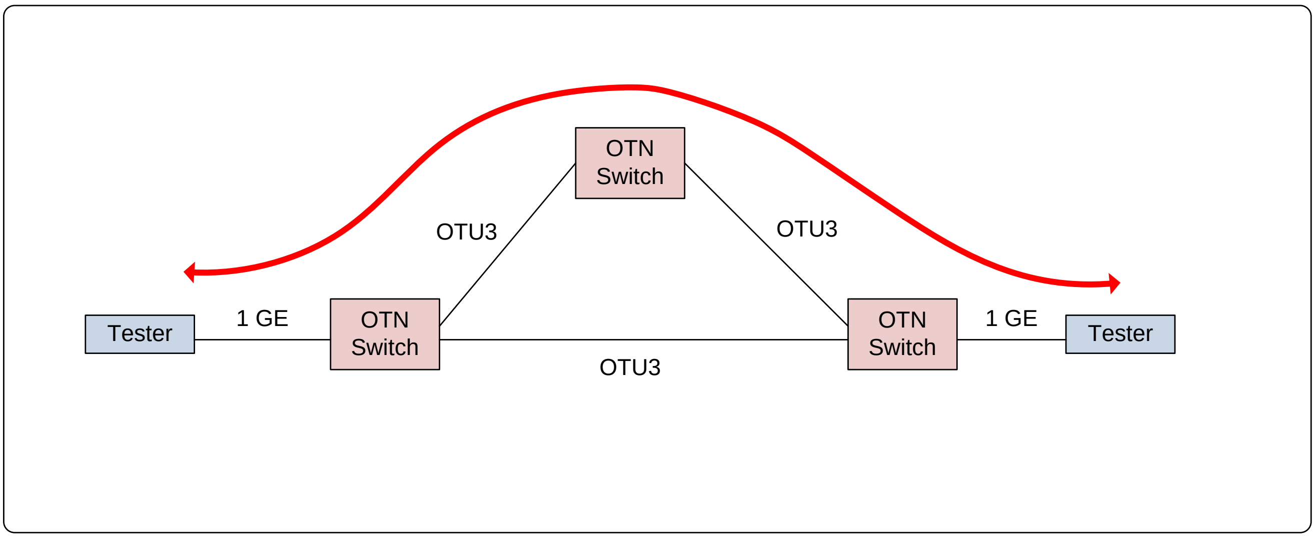


Figure 7.12: Test setup for OTN switching

7.6.3.3 *Test Objective*

The objective of this test was to verify ODU switching functionality and multi-stage mapping.

7.6.3.4 *Test Description*

The test consisted of the following steps:

1. Configure a 1 GE service.
2. Configure the following multiplexing path: ODU0-ODU1-ODU2-ODU3.
3. Send traffic from the tester and verify the traffic flow.

7.6.3.5 *Results*

Multi-stage mapping happens at the edge of the network where the client signal (1 GE in this case) is mapped in several stages into an ODU3. The mapping path chosen during the configuration was ODU0-ODU1-ODU2-ODU3. The ODU0 was switched in the node in the middle of the test setup shown in Figure 7.12 from an OTU3 between the first two switches into a second OTU3 between the switch in the middle and the last switch on the right-hand side of the figure. In this case, the second link was an OTU3, but it could have been an OTU2 or OTU1 link. At the end edge of the test network, the 1 GE signal was de-mapped from the ODU0. This process was verified by monitoring the signal in the tester.

The same procedure takes place in the case of a 10 GE client signal which is switched through an OTU3 link.

7.6.4 Test Conclusions

The conclusions drawn from the ODU switching tests were as follows:

- The equipment was able to map high-speed client signals tested according to standard specifications. Options defined in the standard but not available at the time of the test are in the roadmap and will be available in the near future.
- The equipment was able to map 1 GE signals and signals with a bandwidth between 1 GE and 10 GE by using ODU0 and ODUflex. When using ODUflex, the recommendation is to map traffic bandwidths that are a multiple of the minimum time-slot size, which is approximately 1.25 Gbit/s.
- ODU switching and multi-stage ODUk mapping offer NRENs a more flexible and integrated way to perform circuit switching in a large-scale, cross-domain, IP-over-DWDM network environment. It will make it possible to deliver dedicated capacity with a high level of Quality of Service (QoS) to the user.
- The equipment supports ODU switching functionality and multi-stage mapping.
- Taking into account that the minimum granularity is ODU0 and the possibility for ODUflex, the multiplexing possibilities of OTN give a lot of flexibility when providing connections in the network.

7.7 Survivability

7.7.1 Technology Briefing

G.709 [ITU-T G.709] defines two layers within the OTN technology: a digital layer (OPUk, ODUk) and an analogue layer (Optical Channel (OCh), Optical Multiplex Section (OMS) and Optical Transport Section (OTS)). The survivability (i.e., protection and restoration) procedures follow the already established standards for transport network resiliency defined in G.808.1 (“Generic protection switching – Linear trail and subnetwork protection”) [ITU-T G.808.1]. Currently, there are standards only for survivability provisioning at the ODUk layer (i.e., digital layer survivability): ITU-T G.873.1 (“Optical Transport Network (OTN) – Linear Protection”) [ITU-T G.873.1] and ITU-T G.873.2 (“Optical Transport Network (OTN) – Ring protection” (not published yet)) [ITU-T G.873.2]. Ongoing work is focusing on survivability at the OCh and OMS layers, but no consensus exists in the standardisation organisation on whether such methods should be standardised. ODUk survivability employs a different control mechanism: legacy operation, relying on the digital frame overhead bytes and GMPLS-based failure notification and recovery.

7.7.1.1 *OTN Survivability*

Inherited from the legacy networks, survivability in OTN relies on an Automatic Protection Switching (APS) protocol for bi-directional protection schemes. Upon the detection of a failure, the APS protocol coordinates the bridge and the selector nodes of the protected path to switch the traffic from the failed link/path/trail to a protection one. Whereas in legacy networks the APS protocol was implemented via the K1/L2 bytes, in OTN there are four bytes dedicated to APS/ Protection Communication Channel (PCC) protocol implementation. These carry information related to eight independent APS channels: for end-to-end path level protection, for the six Tandem Connection Monitoring (TCM) levels, and for one level ODUk server layer trail protection. Relying on the APS/PCC bytes from the ODUk overhead for survivability control is an in-band control method. Uni-directional protection schemes may or may not use APS. In this case, APS is not necessary as the head and the tail of the protection path do not need to be synchronised.

Control plane (GMPLS, Private Network-to-Network Interface (P-NNI)) can also be used for protection purposes; in this case, the protection path can be created on detection of a network failure. OTN uses the General Communication Channel (GCC) channels in the ODUk overhead to transport the signalling and routing information.

7.7.2 Test Setup

The setup used for the test is shown in Figure 7.13.

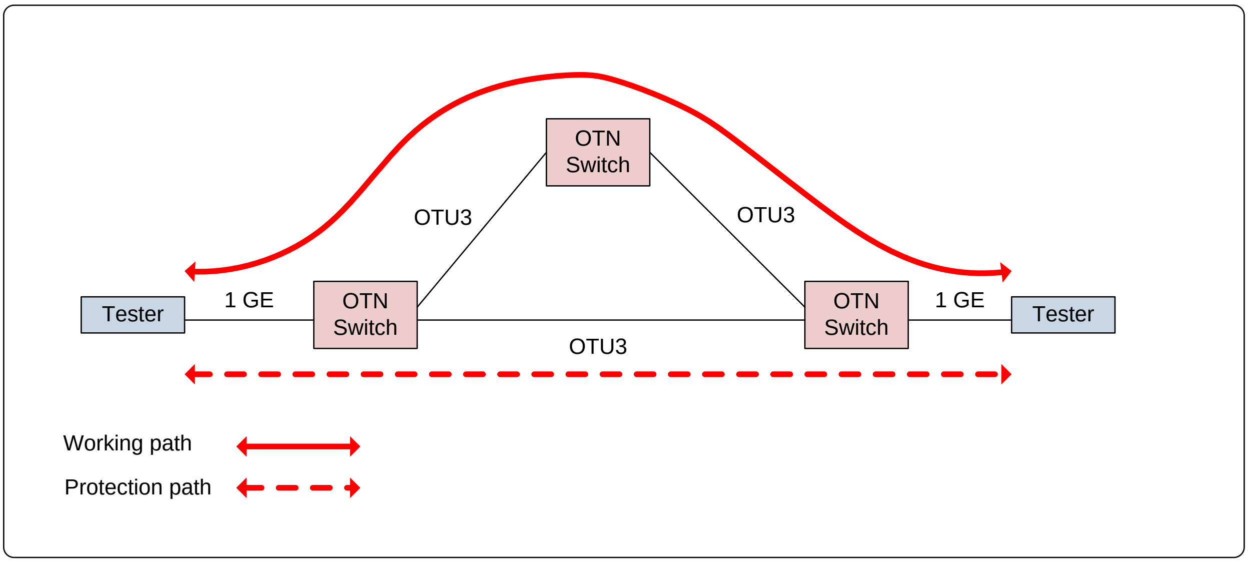


Figure 7.13: Test setup for OTN protection

7.7.3 Test Overview

The ITU-T G.873.1 [ITU-T G.873.1] linear protection standard defines different types of protection architectures. It is important to mention that the equipment used during the testing was early release and not all the possible protection schemes were supported. The following schemes were tested:

- ODUk 1+1 SNC/I protection (uni-directional and bi-directional).
- ODUk 1+1 SNC/N protection (uni-directional).
- ODUk 1+1 SNC/S protection (uni-directional and bi-directional).
- ODUk control plane-assisted restoration.

- ODUk 1+1 SNC GMPLS-assisted protection.

Notice that 1:1 schemes and group protection were not tested, as they were not available at the time of testing.

7.7.4 ODUk 1+1 SNC/I Protection

7.7.4.1 *Technology Briefing*

Sub-Network Connection with Inherent monitoring (SNC/I) protection is a protection scheme that is triggered by defects (Signal Degrade (SD) or Signal Failure (SF)) detected at the ODUk link connection (OTU overhead). No defect detection is performed at the ODUk layer itself. The conditions that trigger a SNC/I protection are shown in Figure 7.14.

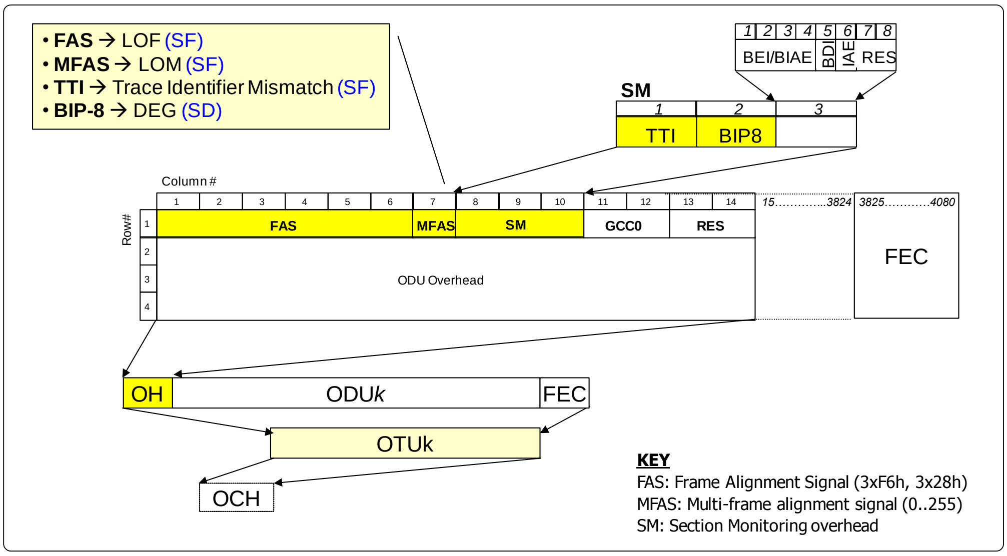


Figure 7.14: ODUk SNC/I protection SD/SF conditions

The protection can be defined to be uni-directional or bi-directional (Figure 7.15). In the case of a bi-directional protection scheme, the APS protocol needs to be used to be able to synchronise the switching of the head and tail ends of the protection path. In the case of a uni-directional scheme, APS is optional. More details about the APS protocol can be found in the ITU-T G.873.1 standard [ITU-T G.873.1].

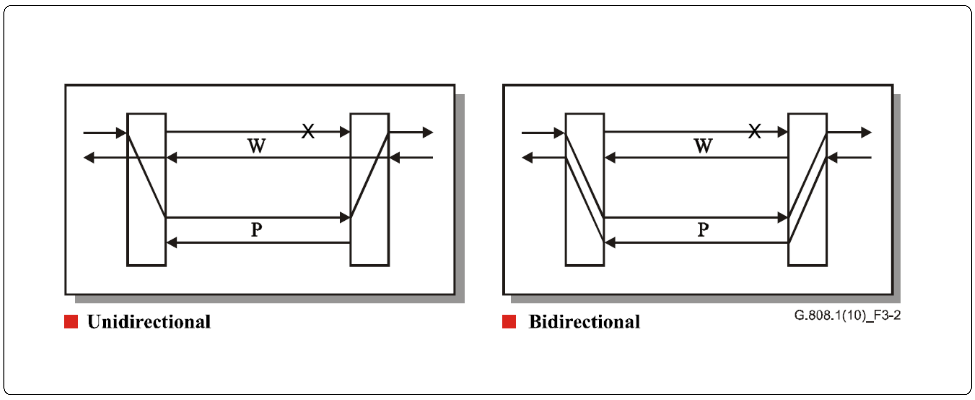


Figure 7.15: Switching types

7.7.4.2 *Test Objective*

The objective of the test was to verify 1+1 SNC/I protection functionality at the OTN layer.

7.7.4.3 *Test Description*

The test consisted of the following steps:

1. Configure a service with 1+1 SNC/I protection.
2. Simulate a network failure in the active circuit. The failure was simulated by introducing bit errors with an attenuator in the fibre.
3. Verify traffic flow between end points.
4. Verify the new route of the traffic flow.

7.7.4.4 *Results*

SNC/I protection was verified. After the condition detection, the traffic was switched to the protection path as expected. The test was performed for both uni-directional and bi-directional cases. In the case of uni-directional switching, only the affected direction was switched to the protection path, while for the bi-directional case both directions were switched to the protection path (APS was used for this purpose).

7.7.5 ODUk 1:1 SNC/N Protection

7.7.5.1 *Technology Briefing*

Sub-Network Connection protection with non-intrusive monitoring (SNC/N) uses non-intrusive monitoring of the ODUk trail at the tail end of the protection path. Only uni-directional switching is defined in the ITU-T G.873.1

standard [ITU-T G.873.1]. Figure 7.16 shows the conditions that trigger the SNC/N protection. The monitoring is done by inspecting the Path Monitoring (PM) bytes in the ODUk overhead.

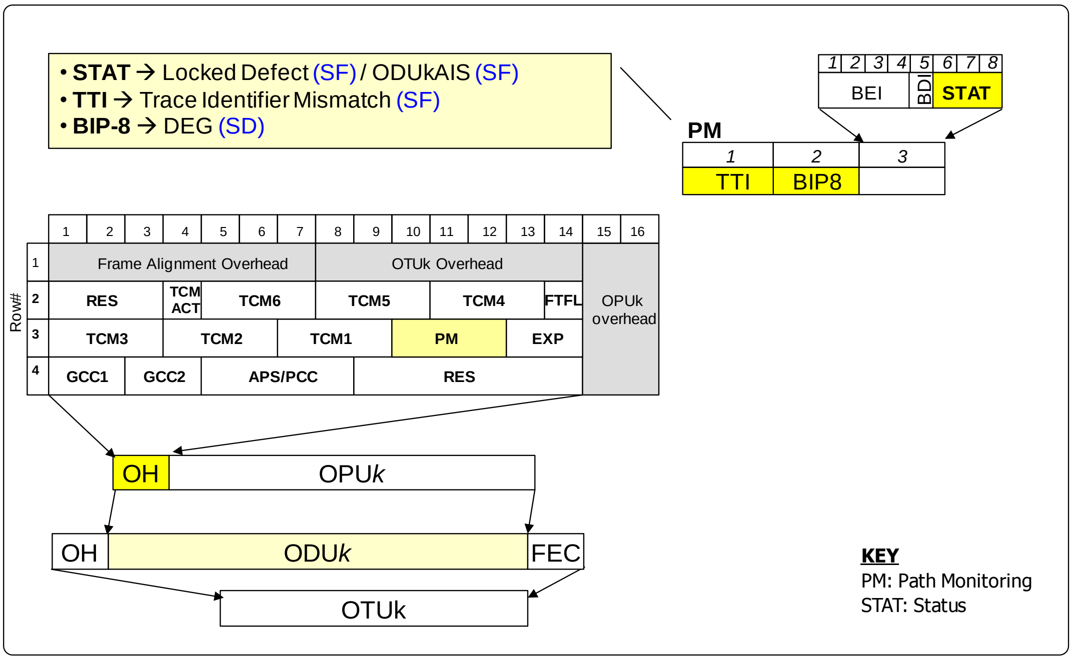


Figure 7.16: SNC/N protection SD and SF conditions

7.7.5.2 *Test Objective*

The objective of the test was to verify 1+1 SNC/N protection.

7.7.5.3 *Test Description*

The test consisted of the following steps:

1. Configure a service with 1+1 SNC/N protection.
2. Simulate a link failure in the active circuit. The failure was simulated by placing an attenuator in the fibre.
3. Verify traffic flow between the end points over the protection path.
4. Verify the new route of the traffic flow.

7.7.5.4 *Results*

SNC/N protection was verified. After the condition detection, the traffic was switched to the protection path as expected. The test was performed only for uni-directional switching. In this case, only the affected direction was switched to the protection path while the other direction was kept in the working path. APS in this case was optional.

When using uni-directional switching, the head end bridges the traffic in both the working and protection path while the tail end is in charge of switching to the protection path in case of a failure (Figure 7.17).

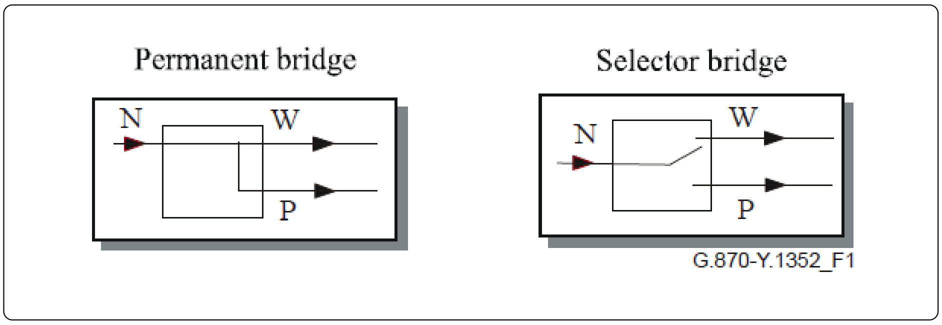


Figure 7.17 Protection bridges

7.7.6 ODUk 1:1 SNC/S Protection

7.7.6.1 *Technology Briefing*

SNC/S is a sub-network connection protection with sub-layer monitoring and is related to Tandem Connection Monitoring (TCM). Given the interest in the extended TCM functionality in OTN, this test scenario is particularly interesting. Protection switching is triggered by defects detected at the ODUkT sub-layer trail. Figure 7.18 shows the conditions that could trigger protection.

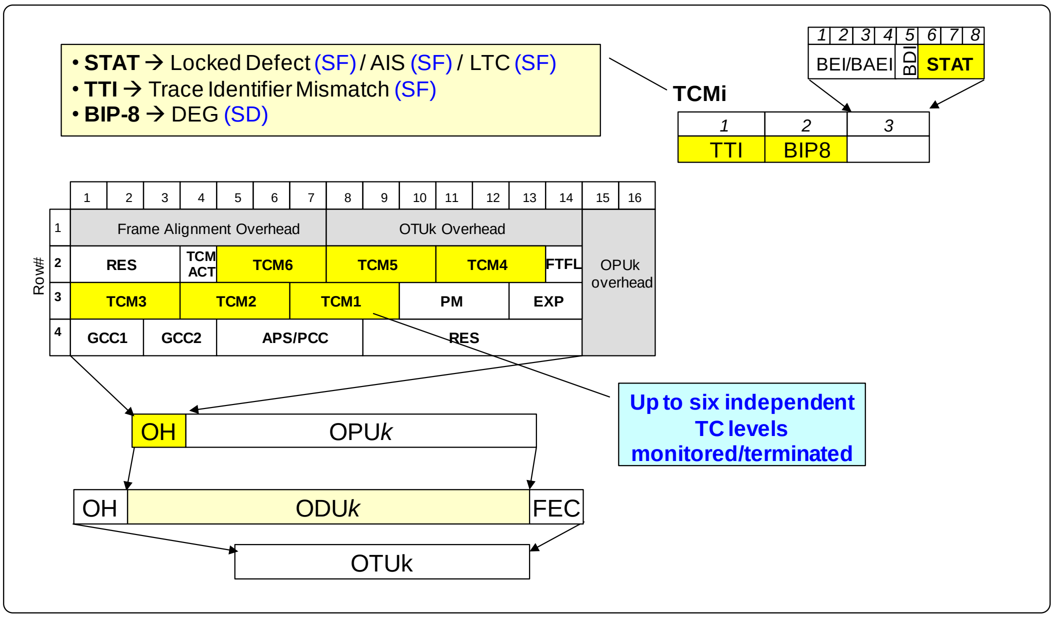


Figure 7.18: SNC/S protection SD and SF conditions

7.7.6.2 *Test Objective*

The objective of the test was to verify 1+1 SNC/S functionality based on TCM at the ODUk layer.

7.7.6.3 *Test Description*

The test consisted of the following steps:

1. Configure a service with 1:1 SNC/S protection.
2. Configure TCM1 between the two end points, as shown in Figure 7.19.
3. Simulate a link failure in the active circuit covered by a TCM trail. The alarm was simulated in this case by manipulating the Trail Trace Identifier (TTI) value in the TCM1 field.
4. Observe and document TCM alarms for SF/SD conditions.
5. Verify traffic flow between the end points over the backup path.

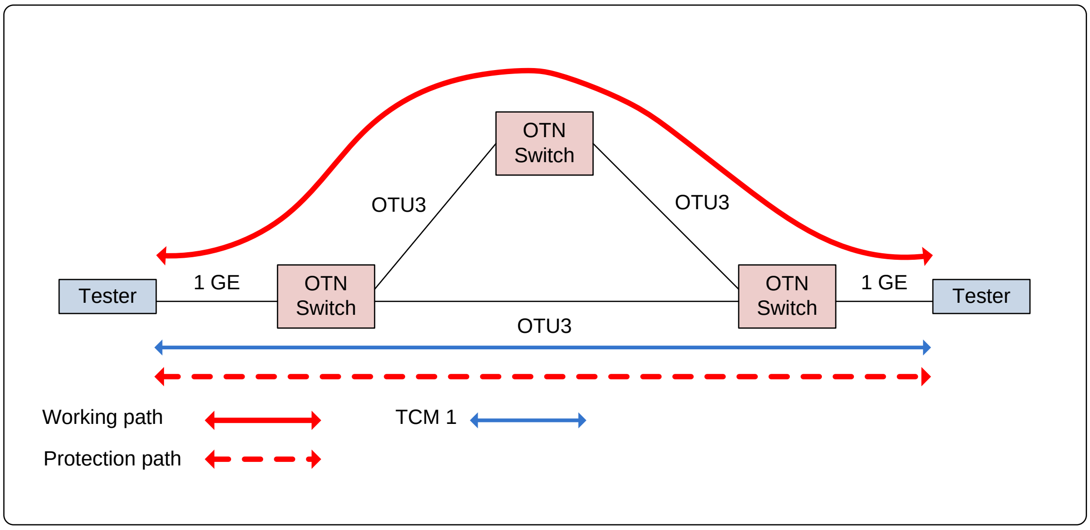


Figure 7.19 SNC/S test setup with TCM1 configured

7.7.6.4 *Results*

SNC/S protection worked as expected. The traffic switched from the working path to the protection path when the defect was detected. In this case, the raised alarm was a Trail Trace Identifier Mismatch (TIM) at TCM1 as expected. SNC/S supports both uni-directional and bi-directional switching. In the case of bi-directional switching, APS protocol is used and both directions are switched to the protection path. Not all the different possibilities were tested due to time constraints.

7.7.7 ODUk Control Plane-Assisted Restoration

7.7.7.1 *Technology Briefing*

OTN supports protection and restoration by means of the control plane. In case of a failure in the network, the control plane mechanism is able to find a new route (routing) and to establish (signalling) the new path. Different implementations and functionalities are available. An option that is quite attractive is the so-called “mesh restoration”, where a protection path is pre-provisioned by the control plane at the same time as the configuration of the primary path, but is only activated (i.e. traffic is switched from the active path to the protection path) in the case of a network failure. After the traffic has been restored on the protection path, a new protection path is calculated and pre-provisioned. This option allows 50 ms switching despite the use of control plane restoration. However, it requires bandwidth reservation.

7.7.7.2 *Test Overview and Setup*

The first test describes control plane restoration without mesh restoration; the test with mesh restoration is described in Section 7.7.7.5. Figure 7.20 shows the setup used for this test.

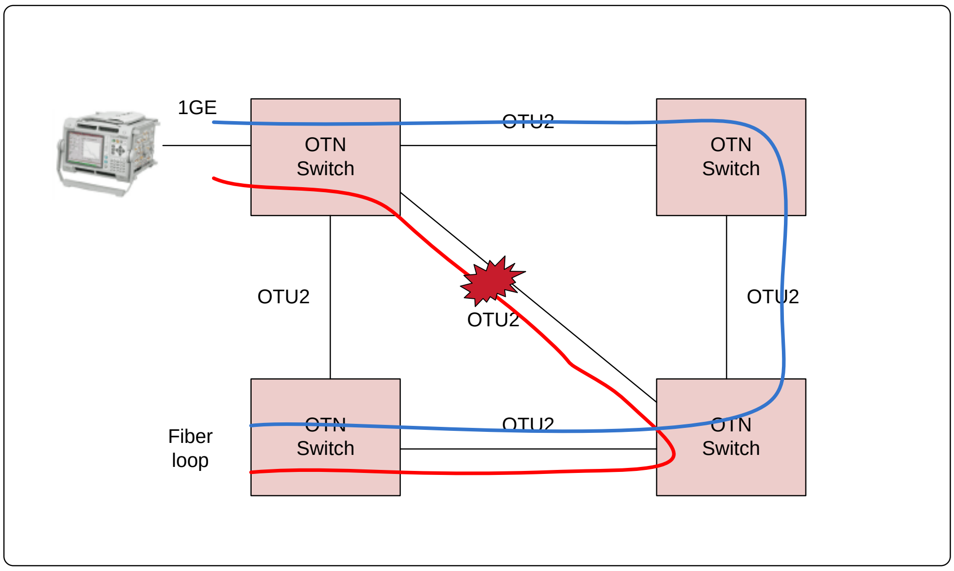


Figure 7.20: Control plane-assisted restoration test setup

7.7.7.3 *Control Plane Restoration without Mesh Restoration*

Test Objective

The objective of the test was to verify ODUk 1+1 protection and restoration of a circuit by means of the control plane.

Test Description

The test consisted of the following steps:

1. Configure a 1+1 Sub-Network Connection Protection (SNCP) protected service.
2. Verify the current route of the service.
3. Simulate a link failure in the circuit.
4. Verify that traffic is still flowing between end points.
5. Verify that the traffic is taking a different route.

Results

Protection and restoration based on control plane was verified. After the detection of the failure, the path was restored by using a new available route in the network. The new route is chosen based on the administrative cost of the different links. The operator can control how the path is re-routed by manipulating the administrative cost values.

7.7.7.4 *Route Cost Modification (Latency-Based Routing)*

Test Overview and Setup

This test is a variation of the previous one. The test had two parts. First, rerouting of a path was forced by manipulating the administrative cost of a link; second, latency was used as the criterion instead of using the administrative cost. In this case, the routing was based on the total latency of an individual path. The route with the lowest latency should be selected. The latency is measured by using a bit in the PM overhead, as shown in Figure 7.21 below. This measurement can be done automatically or on demand. The latency values per link can also be configured manually by the operator. The setup used for this test is described in Figure 7.20 above.

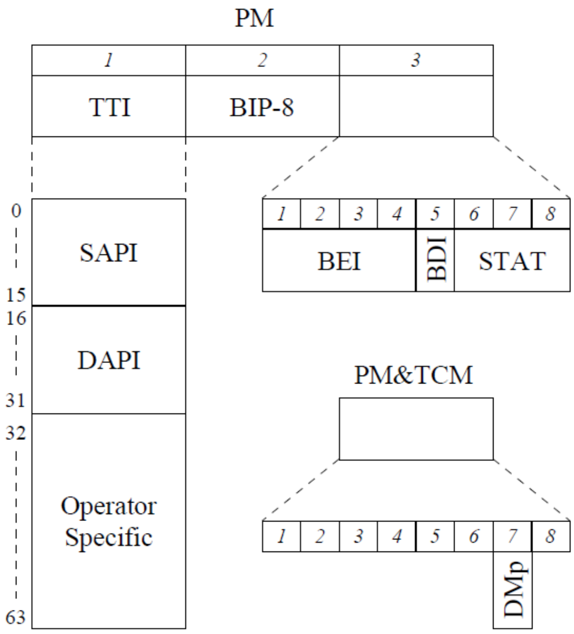


Figure 7.21: ODUk path monitoring overhead

Test Objective

The objective of the test was to demonstrate rerouting of circuits based on manipulation of route cost or by means of delay (latency) measurements.

Test Description

The test consisted of the following steps:

1. Provision a path between two nodes.
2. Change the administrative weight of another link.
3. Check whether the path is rerouted via the link with the lowest administrative weight.
4. Select latency-based routing.
5. Measure the latency between nodes.
6. Simulate a network failure.
7. Check whether the path is rerouted based on the lowest latency route.

Results

The test verified that ODUk restoration was possible based on route cost modification or latency measurements. During the test it was also verified that the operator was able to control the route of the traffic by manipulating the cost and latency values. The latency could be measured using the new OTN features in the ODU overhead or could be manually configured by the operator. This last option gives the operator the possibility of controlling the traffic flow.

7.7.7.5 *Hybrid-Restoration Protection*

Test Overview and Setup

This type of protection uses the control plane to pre-compute and establish the protection path at the time of provisioning the primary connection, in comparison to standard operation where the protection path is calculated and pre-established only when the outage occurs. Hybrid-restoration protection allows 50 ms switching times as for normal SNCP without the control plane. In case the working path fails, the traffic is switched to the protection path, a second protection path is calculated and established, and in the event of a second failure a new protection path is available. The bandwidth in the protection path in this case is reserved. The concept is shown in Figure 7.22.

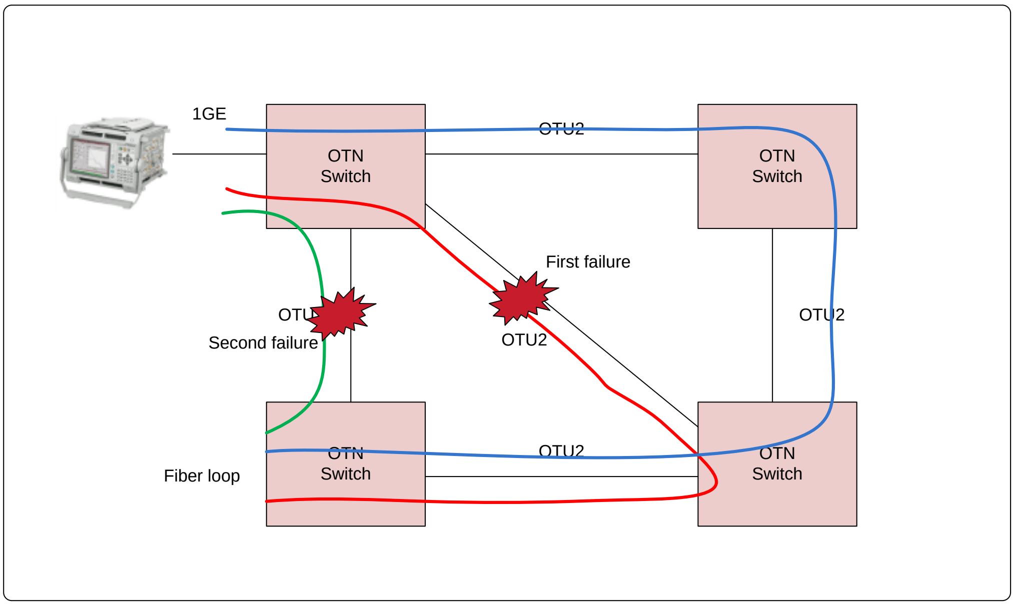


Figure 7.22: Hybrid-restoration protection example

Test Objective

The objective of the test was to verify SNCP hybrid-restoration protection.

Test Description

The test consisted of the following steps:

1. Configure a path with hybrid-restoration protection (the red path in Figure 7.22).
2. Verify the traffic is flowing with the test equipment.
3. Verify the current route of the traffic.
4. Verify that a protection path has been calculated and established (the green path in Figure 7.22).
5. Simulate a failure in the current path (the red path in Figure 7.22).
6. Verify that the traffic is switched to the protection path (the green path in Figure 7.22).
7. Verify that a second protection path has been calculated and established (the blue path in Figure 7.22).

8. Simulate a second failure in the network (the green path in Figure 7.22).
9. Verify that the traffic is switched to the new route (the blue path in Figure 7.22).

Results

Hybrid-restoration protection was verified and worked as expected. It was also possible to verify in the tester that the switching time was less than 50 ms.

7.7.8 GMPLS-Based Fast Wavelength Restoration (FWR)

7.7.8.1 *Test Overview*

Due to limitations in the simulation setup available at the time, the only possible way to observe dynamic protocol interaction in case of a link failure was via the management system logs. The scenario tested was a dynamic wavelength restoration, supported by the GMPLS control plane.

7.7.8.2 *Test Objective*

The objective of the test was to observe the GMPLS protocol interactions during service restoration.

7.7.8.3 *Test Description*

The test consisted of the following steps:

1. Configure an unprotected OTU2 service.
2. Introduce a link failure on the working path.
3. Observe the control plane log for GMPLS-related messages via the management system.
4. Observe Open Shortest Path First with Traffic Engineering (OSPF-TE) for updating the change in topology.
5. Observe Resource Reservation Protocol with Traffic Engineering (RSVP-TE) for failure notification and Label Switched Path (LSP) reestablishment (backup path establishment).
6. Verify traffic flow between the end points.

7.7.8.4 *Expected Results*

Since the objective of the test was to observe the GMPLS-assisted restoration of a service, it was expected that GMPLS-related activity would be seen in all nodes along the ring. In particular, Node A and Node C should exchange GMPLS messages at service creation (to set up the ODU2 service). At failure, Node A should initiate restoration procedure; Node B and Node C should receive GMPLS setup messages for configuring the nodes for the restoration path. Meanwhile, Node A should send a resource release message for the failed working path and Node C should release the resources previously occupied by the working path. OSPF messages should be exchanged once a change in the topology has occurred.

7.7.8.5 *Results*

The control plane message interaction was observed via the management system. The traces shown in Figure 7.23, Figure 7.24 and Figure 7.25 were obtained. The observed processes were as follows:

Node A

- The service creation was indicated: 15:22:20.48.
- An RSVP message was sent and an RSVP state was created for the initialised LSP.
- About 20 ms later, the GMPLS indicated that the connection had been established.
- At 15:25:10.06, the GMPLS control plane requested dynamic restoration, a new RSVP session was initiated, a new LSP was established and a GMPLS request to tear down the failed working path was sent (15:25:27.35).

Node B

Node B was involved only in the setup of the restoration path. At 15:25:27.26 an RSVP request was received. Shortly after, GMPLS indicated that the connection had been established.

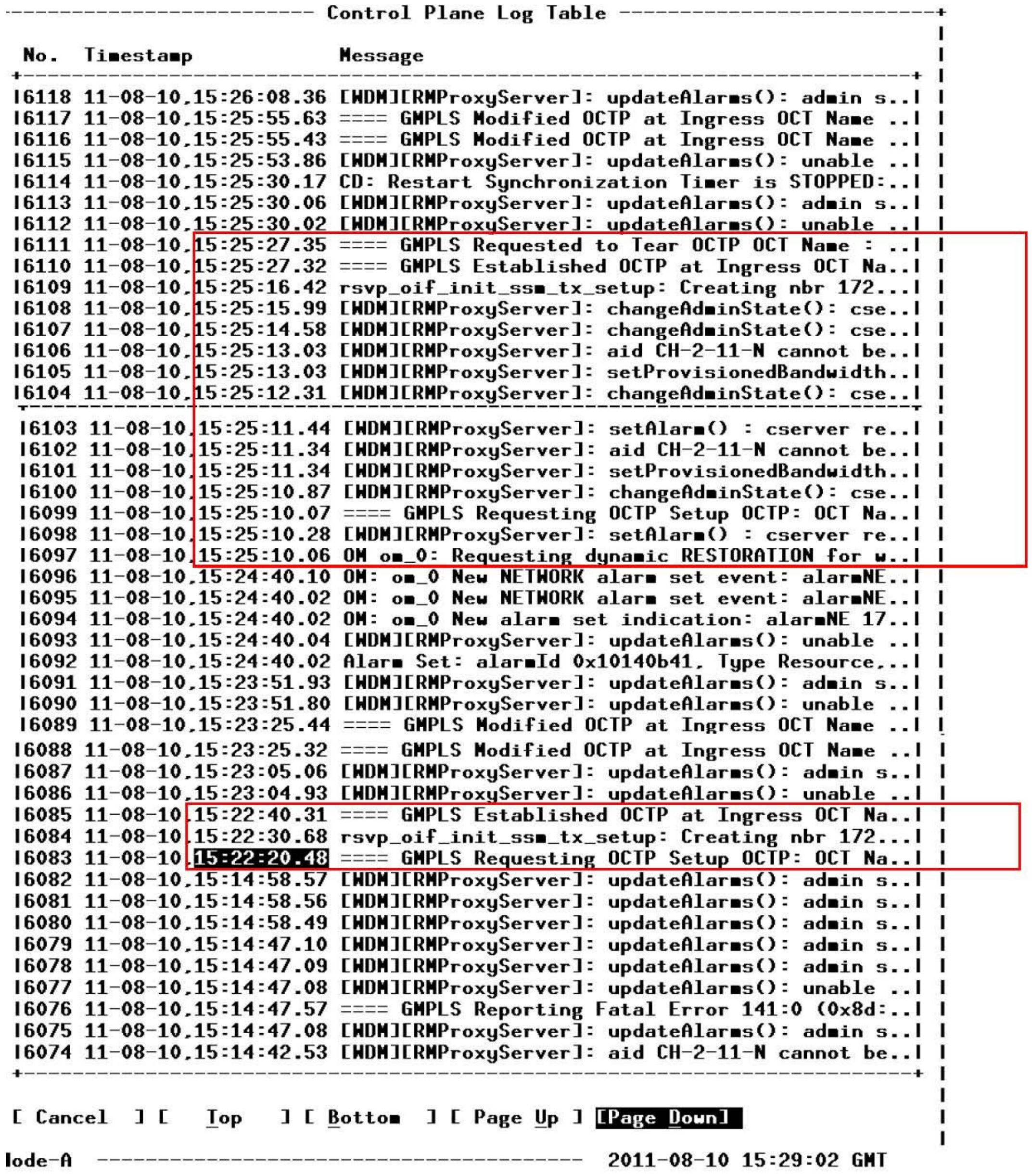
Node C

Node C was the destination of the created service and was involved in all processes related to the restoration.

- At 15:23:07.56, the working service was established.
- After the failure, the restoration connection was established (15:25:54.54).
- The request for Path Tear of the working path was received (15:25:57.45).

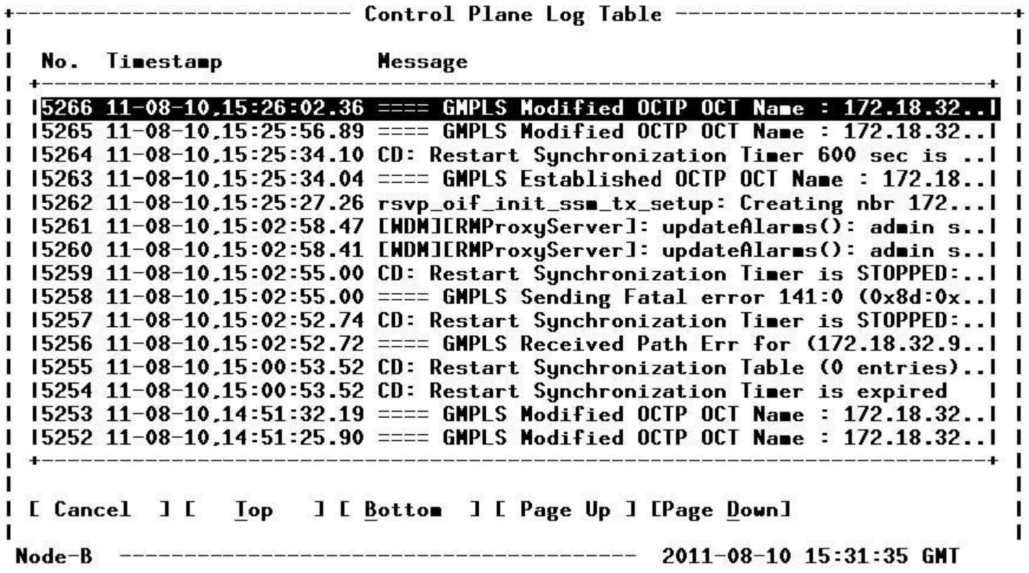
Analysis

A clear picture of the restoration procedure was obtained by following the control plane log. All nodes participated in the procedures as expected. Due to the crudeness of the log system, it was not possible to obtain a more detailed view of the exact messages exchanged between the nodes. Also, due to the specifics of the implementation, additional log entries were observed; it is unclear (even after asking the vendor) whether these are related to the GMPLS protocol exchange or are internal vendor implementation-related messages. Furthermore, only a general idea about the speed of the restoration could be obtained since the exact time of failure was not indicated (Node C logged an alarm event at 15:25:07, whereas Node A logged an alarm event at 25:24:40). Whether these were related to the exact failure moment or to internal system notifications and refreshing alarm indicators is unclear (again, even after asking the vendor). Furthermore, no OSPF messages were observed. This is due to the fact that under the vendor control plane implementation, a link failure is not treated as a change in the bandwidth availability on a link, and thus no OSPF re-convergence is initiated.



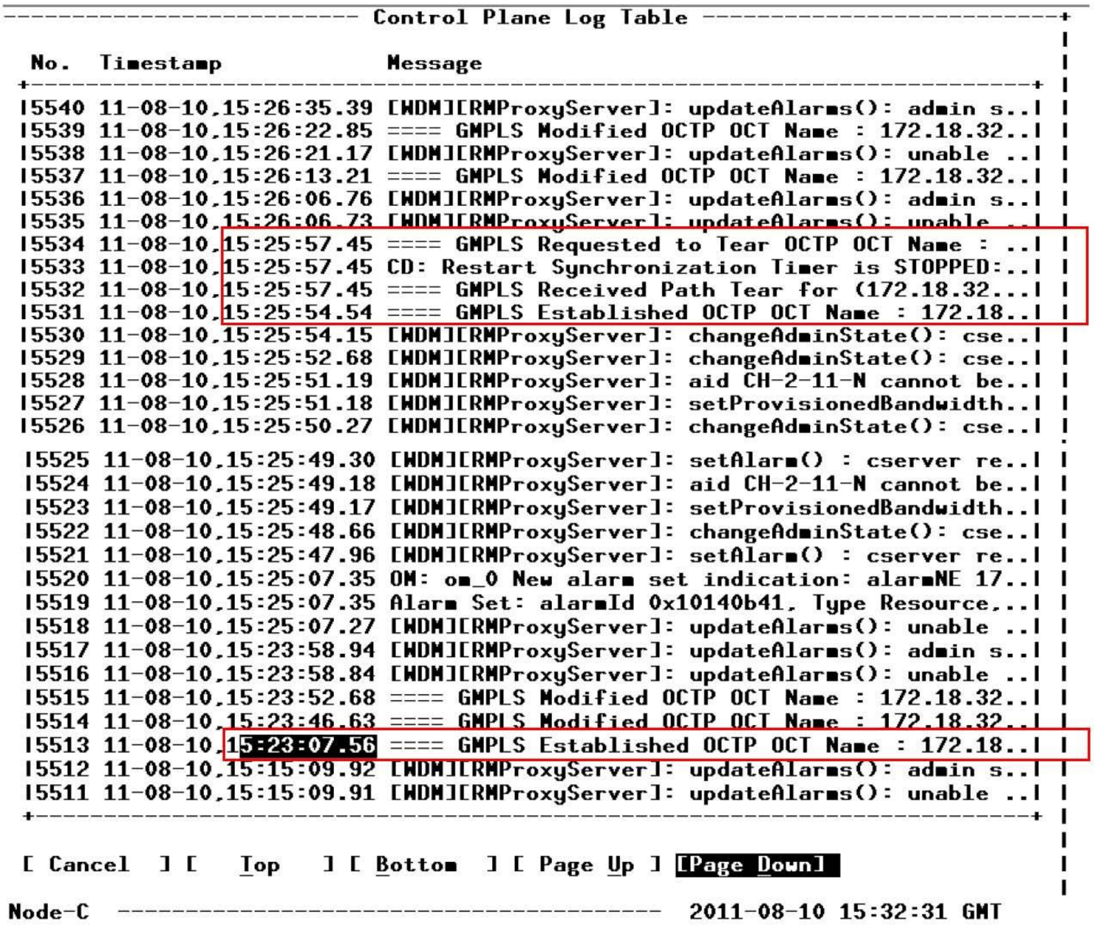
```
-------------------------- Control Plane Log Table --------------------------+
                                                                             |
 No.  Timestamp              Message                                         |
+--------------------------------------------------------------------------+ |
|6118 11-08-10,15:26:08.36 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6117 11-08-10,15:25:55.63 ==== GMPLS Modified OCTP at Ingress OCT Name ..| |
|6116 11-08-10,15:25:55.43 ==== GMPLS Modified OCTP at Ingress OCT Name ..| |
|6115 11-08-10,15:25:53.86 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
|6114 11-08-10,15:25:30.17 CD: Restart Synchronization Timer is STOPPED:..| |
|6113 11-08-10,15:25:30.06 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6112 11-08-10,15:25:30.02 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
|6111 11-08-10,15:25:27.35 ==== GMPLS Requested to Tear OCTP OCT Name : ..| |
|6110 11-08-10,15:25:27.32 ==== GMPLS Established OCTP at Ingress OCT Na..| |
|6109 11-08-10,15:25:16.42 rsvp_oif_init_ssm_tx_setup: Creating nbr 172...| |
|6108 11-08-10,15:25:15.99 [WDM][RMProxyServer]: changeAdminState(): cse..| |
|6107 11-08-10,15:25:14.58 [WDM][RMProxyServer]: changeAdminState(): cse..| |
|6106 11-08-10,15:25:13.03 [WDM][RMProxyServer]: aid CH-2-11-N cannot be..| |
|6105 11-08-10,15:25:13.03 [WDM][RMProxyServer]: setProvisionedBandwidth..| |
|6104 11-08-10,15:25:12.31 [WDM][RMProxyServer]: changeAdminState(): cse..| |
+--------------------------------------------------------------------------+ |
 |6103 11-08-10,15:25:11.44 [WDM][RMProxyServer]: setAlarm() : cserver re..| |
 |6102 11-08-10,15:25:11.34 [WDM][RMProxyServer]: aid CH-2-11-N cannot be..| |
 |6101 11-08-10,15:25:11.34 [WDM][RMProxyServer]: setProvisionedBandwidth..| |
 |6100 11-08-10,15:25:10.87 [WDM][RMProxyServer]: changeAdminState(): cse..| |
 |6099 11-08-10,15:25:10.07 ==== GMPLS Requesting OCTP Setup OCTP: OCT Na..| |
 |6098 11-08-10,15:25:10.28 [WDM][RMProxyServer]: setAlarm() : cserver re..| |
 |6097 11-08-10,15:25:10.06 OM om_0: Requesting dynamic RESTORATION for w..| |
 |6096 11-08-10,15:24:40.10 OM: om_0 New NETWORK alarm set event: alarmNE..| |
 |6095 11-08-10,15:24:40.02 OM: om_0 New NETWORK alarm set event: alarmNE..| |
 |6094 11-08-10,15:24:40.02 OM: om_0 New alarm set indication: alarmNE 17..| |
 |6093 11-08-10,15:24:40.04 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
 |6092 11-08-10,15:24:40.02 Alarm Set: alarmId 0x10140b41, Type Resource,..| |
 |6091 11-08-10,15:23:51.93 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
 |6090 11-08-10,15:23:51.80 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
 |6089 11-08-10,15:23:25.44 ==== GMPLS Modified OCTP at Ingress OCT Name ..| |
|6088 11-08-10,15:23:25.32 ==== GMPLS Modified OCTP at Ingress OCT Name ..| |
|6087 11-08-10,15:23:05.06 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6086 11-08-10,15:23:04.93 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
|6085 11-08-10,15:22:40.31 ==== GMPLS Established OCTP at Ingress OCT Na..| |
|6084 11-08-10,15:22:30.68 rsvp_oif_init_ssm_tx_setup: Creating nbr 172...| |
|6083 11-08-10,15:22:20.48 ==== GMPLS Requesting OCTP Setup OCTP: OCT Na..| |
|6082 11-08-10,15:14:58.57 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6081 11-08-10,15:14:58.56 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6080 11-08-10,15:14:58.49 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6079 11-08-10,15:14:47.10 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6078 11-08-10,15:14:47.09 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6077 11-08-10,15:14:47.08 [WDM][RMProxyServer]: updateAlarms(): unable ..| |
|6076 11-08-10,15:14:47.57 ==== GMPLS Reporting Fatal Error 141:0 (0x8d:..| |
|6075 11-08-10,15:14:47.08 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
|6074 11-08-10,15:14:42.53 [WDM][RMProxyServer]: aid CH-2-11-N cannot be..| |
+--------------------------------------------------------------------------+ |
                                                                             |
[ Cancel  ] [   Top    ] [ Bottom   ] [ Page Up ] [Page Down]                |
                                                                             |
lode-A  ---------------------------------------   2011-08-10 15:29:02 GMT
```

Figure 7.23: Control plane log table Node A



```
+-------------------------- Control Plane Log Table --------------------------+
|                                                                             |
|  No.  Timestamp              Message                                        |
| +-------------------------------------------------------------------------+ |
| |15266 11-08-10,15:26:02.36 ==== GMPLS Modified OCTP OCT Name : 172.18.32..| |
| |15265 11-08-10,15:25:56.89 ==== GMPLS Modified OCTP OCT Name : 172.18.32..| |
| |15264 11-08-10,15:25:34.10 CD: Restart Synchronization Timer 600 sec is ..| |
| |15263 11-08-10,15:25:34.04 ==== GMPLS Established OCTP OCT Name : 172.18..| |
| |15262 11-08-10,15:25:27.26 rsvp_oif_init_ssm_tx_setup: Creating nbr 172...| |
| |15261 11-08-10,15:02:58.47 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
| |15260 11-08-10,15:02:58.41 [WDM][RMProxyServer]: updateAlarms(): admin s..| |
| |15259 11-08-10,15:02:55.00 CD: Restart Synchronization Timer is STOPPED:..| |
| |15258 11-08-10,15:02:55.00 ==== GMPLS Sending Fatal error 141:0 (0x8d:0x..| |
| |15257 11-08-10,15:02:52.74 CD: Restart Synchronization Timer is STOPPED:..| |
| |15256 11-08-10,15:02:52.72 ==== GMPLS Received Path Err for (172.18.32.9..| |
| |15255 11-08-10,15:00:53.52 CD: Restart Synchronization Table (0 entries)..| |
| |15254 11-08-10,15:00:53.52 CD: Restart Synchronization Timer is expired   | |
| |15253 11-08-10,14:51:32.19 ==== GMPLS Modified OCTP OCT Name : 172.18.32..| |
| |15252 11-08-10,14:51:25.90 ==== GMPLS Modified OCTP OCT Name : 172.18.32..| |
| +-------------------------------------------------------------------------+ |
|                                                                             |
| [ Cancel  ] [   Top    ] [ Bottom   ] [ Page Up ] [Page Down]               |
|                                                                             |
 Node-B  ----------------------------------------  2011-08-10 15:31:35 GMT
```

Figure 7.24: Control plane log table Node B



```
--------------------------- Control Plane Log Table ---------------------------+
                                                                              |
  No.  Timestamp              Message                                         |
+-------------------------------------------------------------------------+   |
|15540 11-08-10,15:26:35.39 [WDM][RMProxyServer]: updateAlarms(): admin s..|   |
|15539 11-08-10,15:26:22.85 ==== GMPLS Modified OCTP OCT Name : 172.18.32..|   |
|15538 11-08-10,15:26:21.17 [WDM][RMProxyServer]: updateAlarms(): unable ..|   |
|15537 11-08-10,15:26:13.21 ==== GMPLS Modified OCTP OCT Name : 172.18.32..|   |
|15536 11-08-10,15:26:06.76 [WDM][RMProxyServer]: updateAlarms(): admin s..|   |
|15535 11-08-10,15:26:06.73 [WDM][RMProxyServer]: updateAlarms(): unable ..|   |
|15534 11-08-10,15:25:57.45 ==== GMPLS Requested to Tear OCTP OCT Name : ..|   |
|15533 11-08-10,15:25:57.45 CD: Restart Synchronization Timer is STOPPED:..|   |
|15532 11-08-10,15:25:57.45 ==== GMPLS Received Path Tear for (172.18.32...|   |
|15531 11-08-10,15:25:54.54 ==== GMPLS Established OCTP OCT Name : 172.18..|   |
|15530 11-08-10,15:25:54.15 [WDM][RMProxyServer]: changeAdminState(): cse..|   |
|15529 11-08-10,15:25:52.68 [WDM][RMProxyServer]: changeAdminState(): cse..|   |
|15528 11-08-10,15:25:51.19 [WDM][RMProxyServer]: aid CH-2-11-N cannot be..|   |
|15527 11-08-10,15:25:51.18 [WDM][RMProxyServer]: setProvisionedBandwidth..|   |
|15526 11-08-10,15:25:50.27 [WDM][RMProxyServer]: changeAdminState(): cse..|   |
 |15525 11-08-10,15:25:49.30 [WDM][RMProxyServer]: setAlarm() : cserver re..|  |
 |15524 11-08-10,15:25:49.18 [WDM][RMProxyServer]: aid CH-2-11-N cannot be..|  |
 |15523 11-08-10,15:25:49.17 [WDM][RMProxyServer]: setProvisionedBandwidth..|  |
 |15522 11-08-10,15:25:48.66 [WDM][RMProxyServer]: changeAdminState(): cse..|  |
 |15521 11-08-10,15:25:47.96 [WDM][RMProxyServer]: setAlarm() : cserver re..|  |
 |15520 11-08-10,15:25:07.35 OM: om_0 New alarm set indication: alarmNE 17..|  |
 |15519 11-08-10,15:25:07.35 Alarm Set: alarmId 0x10140b41, Type Resource,..|  |
 |15518 11-08-10,15:25:07.27 [WDM][RMProxyServer]: updateAlarms(): unable ..|  |
 |15517 11-08-10,15:23:58.94 [WDM][RMProxyServer]: updateAlarms(): admin s..|  |
 |15516 11-08-10,15:23:58.84 [WDM][RMProxyServer]: updateAlarms(): unable ..|  |
 |15515 11-08-10,15:23:52.68 ==== GMPLS Modified OCTP OCT Name : 172.18.32..|  |
 |15514 11-08-10,15:23:46.63 ==== GMPLS Modified OCTP OCT Name : 172.18.32..|  |
 |15513 11-08-10,15:23:07.56 ==== GMPLS Established OCTP OCT Name : 172.18..|  |
 |15512 11-08-10,15:15:09.92 [WDM][RMProxyServer]: updateAlarms(): admin s..|  |
 |15511 11-08-10,15:15:09.91 [WDM][RMProxyServer]: updateAlarms(): unable ..|  |
 +-------------------------------------------------------------------------+  |
                                                                              |
 [ Cancel  ] [   Top    ] [ Bottom   ] [ Page Up ] [Page Down]                |
                                                                              |
Node-C  ------------------------------------------  2011-08-10 15:32:31 GMT
```

Figure 7.25: Control plane log table Node C

7.7.9 Test Conclusions

The conclusions drawn from the survivability tests were as follows:

- The equipment was able to support all the standard-defined protection architectures tested.
- Hybrid restoration, where an SNCP consisting of a protection path and a primary path is provisioned by the control plane in combination with the restoration function, allows 50 ms switching in case of a failure, but restores the redundant path after the switchover to cater for further faults. This could be of interest to NRENs carrying very important traffic or networks with a very complicated physical fibre infrastructure, because it allows extra protection compared to traditional SNCP and faster service re-establishment compared to control plane restoration. However, it requires more hardware and additional bandwidth compared to dynamic restoration of a path without SNCP.
- In control plane restoration without mesh restoration, the new route can be chosen based on either the administrative cost or the total latency of the different links. The operator can control how the path is re-routed by manipulating the administrative cost and latency values per link. The latency can be measured using the new OTN features in the ODU overhead or can be manually configured by the operator, giving the operator the possibility of controlling the traffic flow.
- Although the only way to observe dynamic protocol interaction in case of a link failure was via the management system logs, in the scenario tested (a dynamic wavelength restoration, supported by the GMPLS control plane) the GMPLS procedure in the restoration event was clear: the head end of the connection initiated a new RSVP session for establishing a new path.

7.8 Tandem Connection Monitoring

7.8.1 Technology Briefing

A major advantage, for the NREN community, of OTN technology is the advance in Tandem Connection Monitoring (TCM). Research and Education networks tend to operate in a multi-domain, multi-vendor environment, often facing the challenge of providing end-to-end monitoring and performance guarantees for the provided services. In the G.709 OTN specification [ITU-T G.709], Tandem Connection Monitoring tackles these problems by specifying higher monitoring levels and more flexible surveillance possibilities.

A tandem connection is a bi-directional connection between two tandem connection terminating elements (TCTEs).

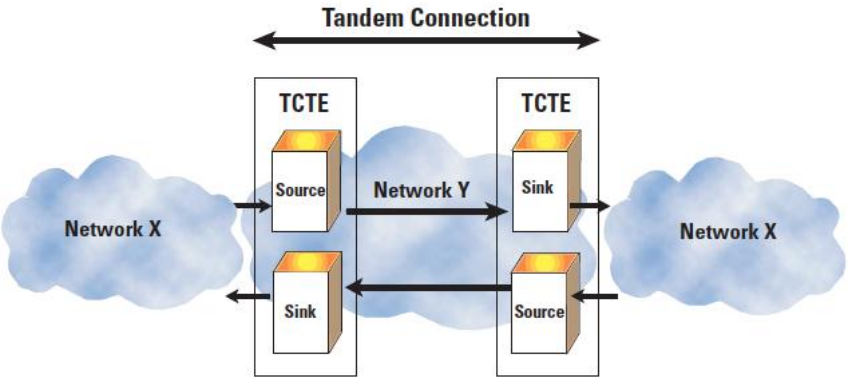


Figure 7.26: Tandem connection monitoring

In the legacy systems, errors are detected and recorded when entering the tandem connection source point (TCM source) and then checked again when exiting at the sink point (TCM sink), as shown in Figure 7.26. By simple comparison of the two error checks, the TC sections can be monitored for fast fault localisation.

OTN technology offers from zero to six levels of connections of TCM. These levels can be configured for each ODU trail and in a flexible manner, allowing nested and overlapping TCM sections (Figure 7.27). Compared to legacy systems, this method provides an indication of the quality of service offered at each segment of the network, which makes it easier for the user and carrier to isolate faulty sections [ITU-T G.709].

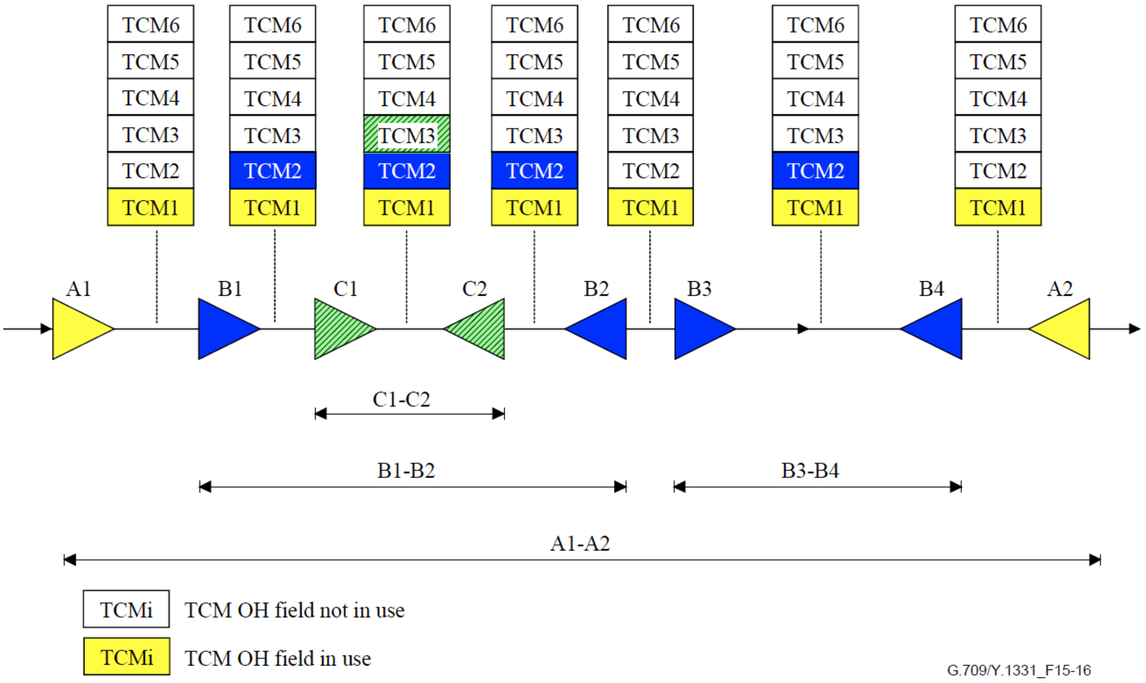

TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
TCM6 TCM5 TCM4 TCM3 TCM2 TCM1
A1
B1
C1
C2
B2
B3
B4
A2
C1-C2
B1-B2
B3-B4
A1-A2
TCMi
TCM OH field not in use
TCMi
TCM OH field in use
G.709/Y.1331_F15-16

Figure 7.27: Tandem Connection Monitoring example

7.8.2 Test Objective

The objective of the test was to verify defect detection and alarm handling in situations that are likely to occur in a multi-domain multi-vendor NREN environment.

7.8.3 Test Setup

As explained in Section 7.8.1, the monitoring advantages offered by OTN are highly relevant for the NREN community, as providing services quite often involves several research networks, thereby complicating the provisioning and OAM processes. The aim of the test setups was therefore to reflect real-life situations that are likely to occur in a multi-domain environment. The three test scenarios are shown in Figure 7.28, Figure 7.29 and Figure 7.30.

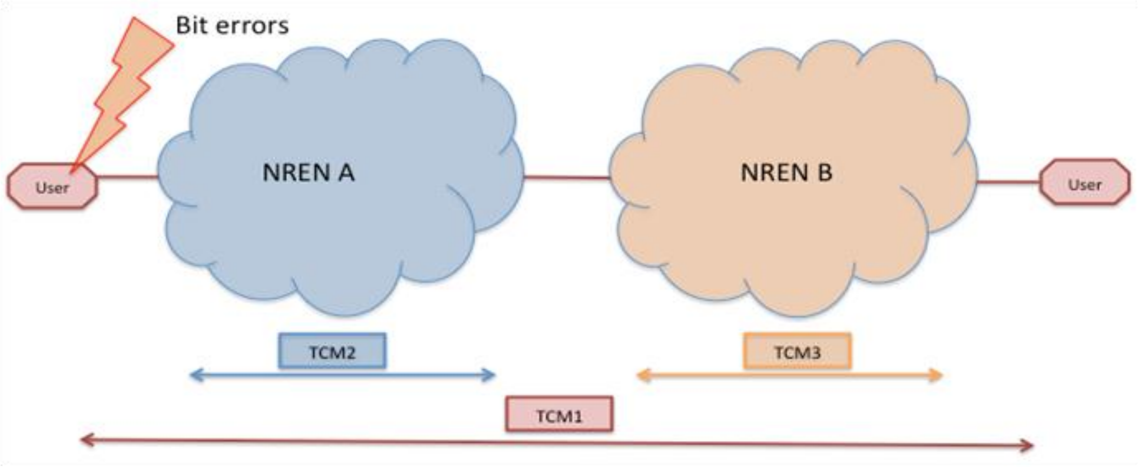


Figure 7.28: Scenario 1: Bit errors from client

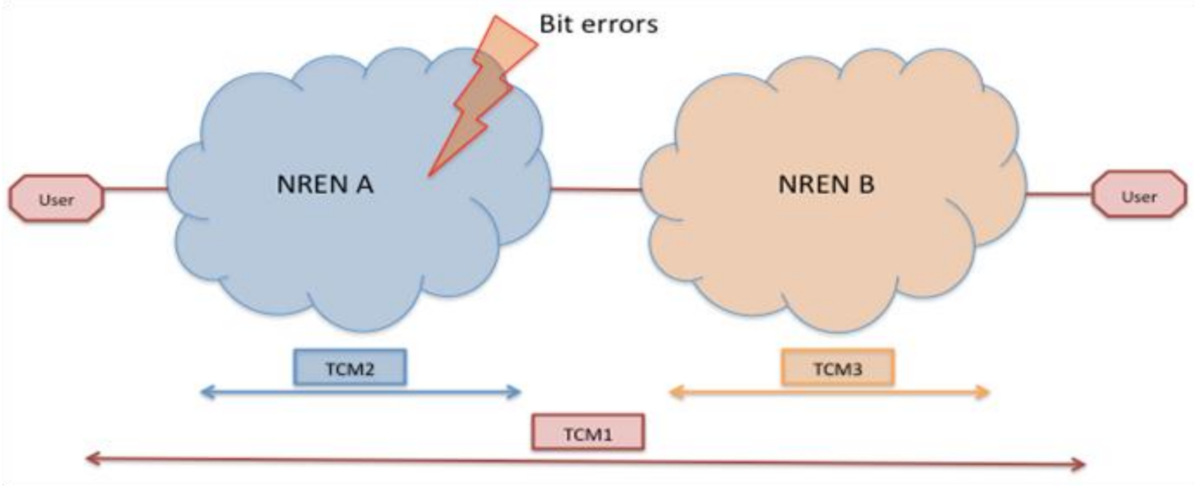


Figure 7.29: Scenario 2: Bit errors inside the domains

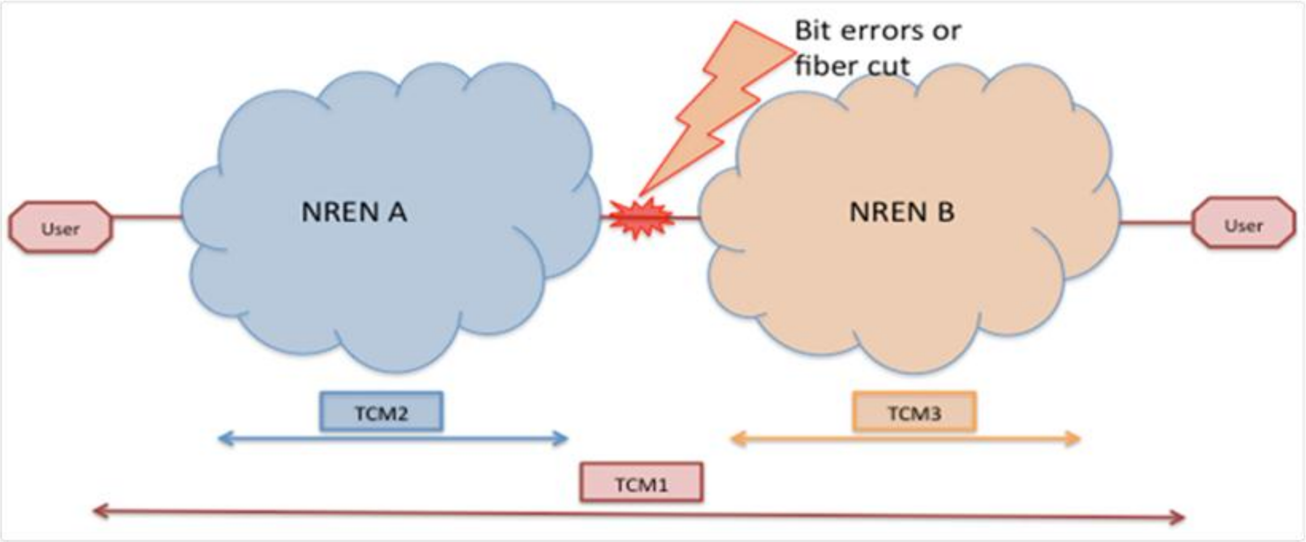


Figure 7.30: Scenario 3: Bit errors between the domains

Although the testing was done on different vendors' setups, the principles of the suggested scenarios were demonstrated.

7.8.3.1 *General Remarks on the Different Configurations*

Considering the different testbed configurations that were investigated during the testing period, it has become evident that utilising the Tandem Connection Monitoring functionality in a multi-vendor, multi-domain environment needs thorough planning and close collaboration between network operators. Although vendors are clearly following the relevant standards, the differences in their actual implementations complicate the task of stitching the TCM domains together. Nevertheless, it is definitely possible, and successful TCM implementations could provide the expected OAM advantages for multi-vendor and multi-domain services.

7.8.4 Test Description

The test consisted of the following steps:

1. Configure the three setups shown in Section 7.8.3.
2. Connect a tester and simulate the error scenarios.
3. Verify and analyse defects and alarm handling in TCM1 (User).
4. Verify and analyse defects and alarm handling in TCM2 (NREN A).
5. Verify and analyse problems between the two TCM sections.

7.8.5 Bit Errors Introduced in TCM1 (User)

The first test scenario covered the user perspective. Errors were introduced in the User TCM section, TCM 1. The alarms in all the TCM sections and measurement points were logged and the behaviour was observed.

7.8.5.1 *Expected Results*

ITU-T recommendation G.798.1 [ITU-T G.798.1] specifies the maintenance signal interaction by OTN OAM (summarised in Figure 7.31 and Figure 7.32). The relationships between detected and generated alarms can be found in Section 7.1.3 of the standard [ITU-T G.798.1].

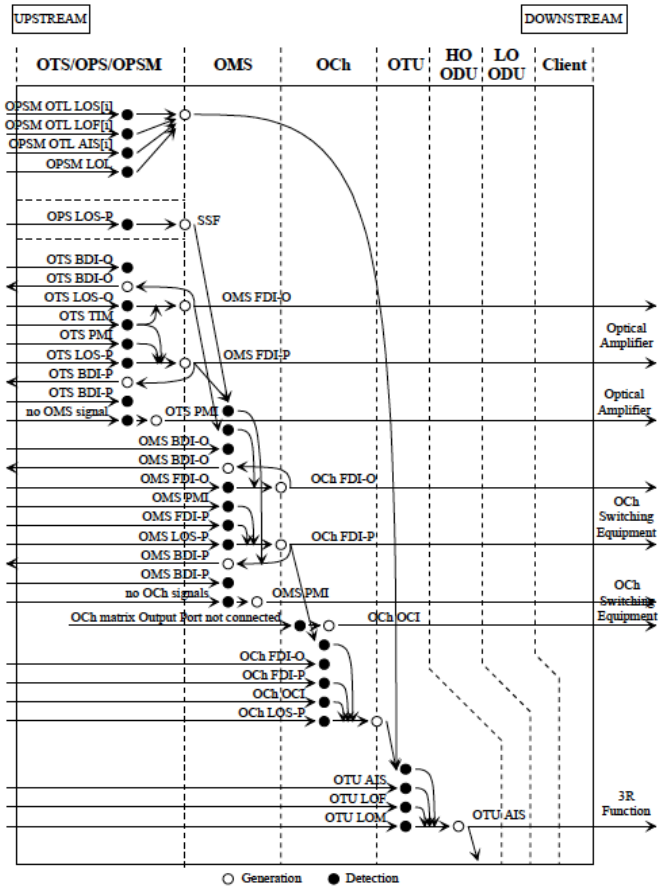


Figure 7-1/G.798.1 – OTN maintenance signal interaction (part I)

Figure 7.31: Maintenance signal interaction

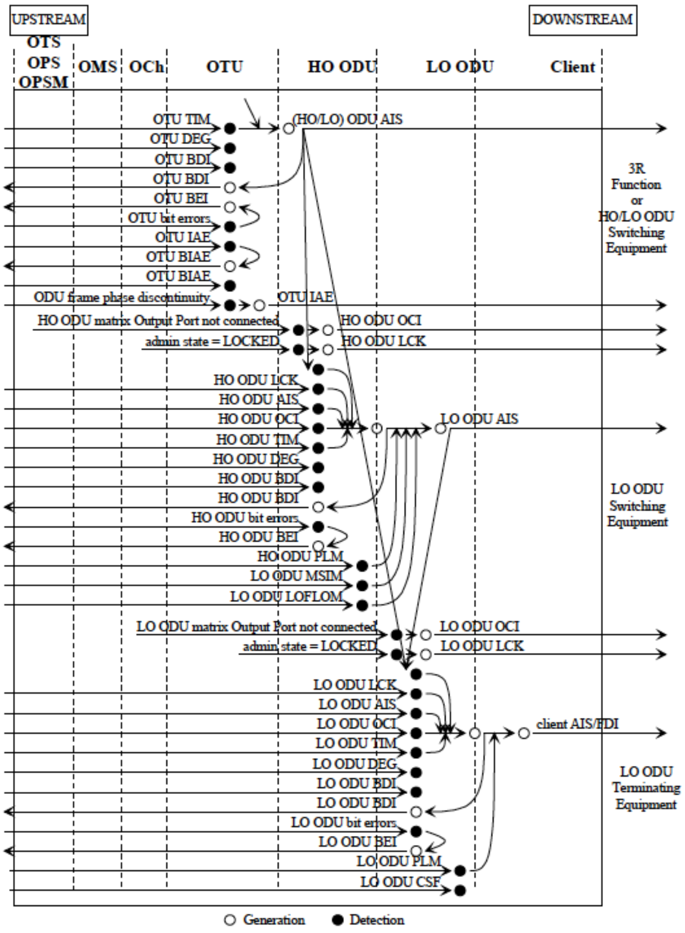


Figure 7-2/G.798.1 – OTN maintenance signal interaction (part II)

Figure 7.32: Maintenance signal interaction (cont.)

As described in this section (i.e. G.798.1 Section 7.1.3), in the event of a Signal Failure (SF) or Signal Degrade (SD), the consequent action will generate a Backward Defect Indication (BDI) or Backward Error Indication (BEI) in the reverse direction of the detection. Based on the above information, the alarm interactions presented in Figure 7.33 were expected upon the failure conditions in the User TCM section.

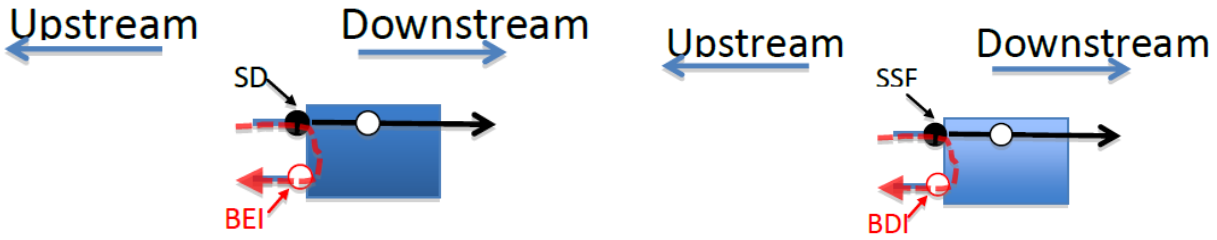


Figure 7.33: Detection of SD and SF

Detection of one Server Signal Failure (SSF) or SD should lead to distribution of the event downstream. Additionally, the event should trigger alarm indications in the opposite (upstream) direction.

In order to introduce bit errors and thereby simulate “signal degrade” conditions, a light source can be connected at the point in the network where the bit errors should be injected, as shown in the figures in Section 7.8.3 above. Furthermore, when simulating the user circuits, it is not recommended to loop the client end point as it disorders the alarm propagation. The problem can be solved by, for example, feeding the TX of the far-end client port to the RX of the test set.

The Signal Fail condition was simulated with variable optical attenuators or simple fibre removal.

7.8.5.2 *Results*

The alarm propagations listed below reflect only the alarms related to the tandem connection sub-layer monitoring. During the simulations, other alarms were propagated in the network, but these were considered irrelevant with respect to understanding the TCM alarm propagations.

In the first test, TCM1 was set up as the User monitoring section and the two other TCM domains were cascaded utilising TCM2 and TCM3. TCM4 was used in between the two domains.

The scenario gave the following results, also shown in Figure 7.34.

• NREN A:	TCM1 alarms (SSF/SD and BDI/BEI).
• NREN B:	TCM1 alarms (SSF/SD).
• Between NRENs:	No alarms.

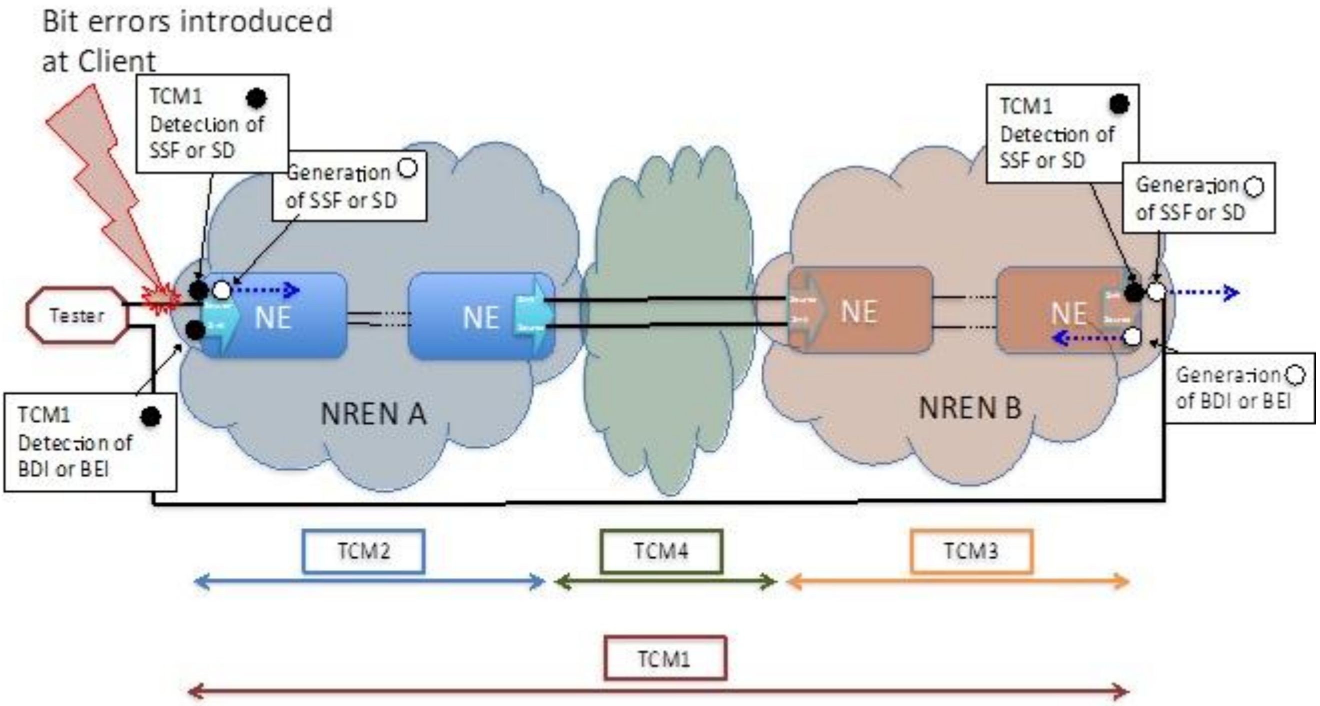

Bit errors introduced at Client
TCM1 ● Detection of SSF or SD
Generation ○ of SSF or SD
Tester
TCM1 ● Detection of BDI or BEI
NE
NE
NE
NE
NREN A
NREN B
TCM1 ● Detection of SSF or SD
Generation ○ of SSF or SD
Generation ○ of BDI or BEI
TCM2
TCM4
TCM3
TCM1

Figure 7.34: Results when bit errors were introduced at the client side

The alarm indications in this scenario quite clearly point to the user/client, as TCM1 alarms were the only alarms seen in the domains.

7.8.6 Bit Errors Introduced in TCM2 (NREN A)

The second scenario presented the case where NREN A is somehow introducing bit errors somewhere in their network. Alarms in all the TCM sections and measurement points were logged and the behaviour was observed.

7.8.6.1 *Results*

In line with expectations, the following alarms, also shown in Figure 7.35, were observed:

• NREN A:	TCM1 alarms (BDI/BEI), TCM2 alarms (SSF/SD).
• NREN B:	TCM1 alarms (SSF/SD).
• Between NRENs:	No alarms.

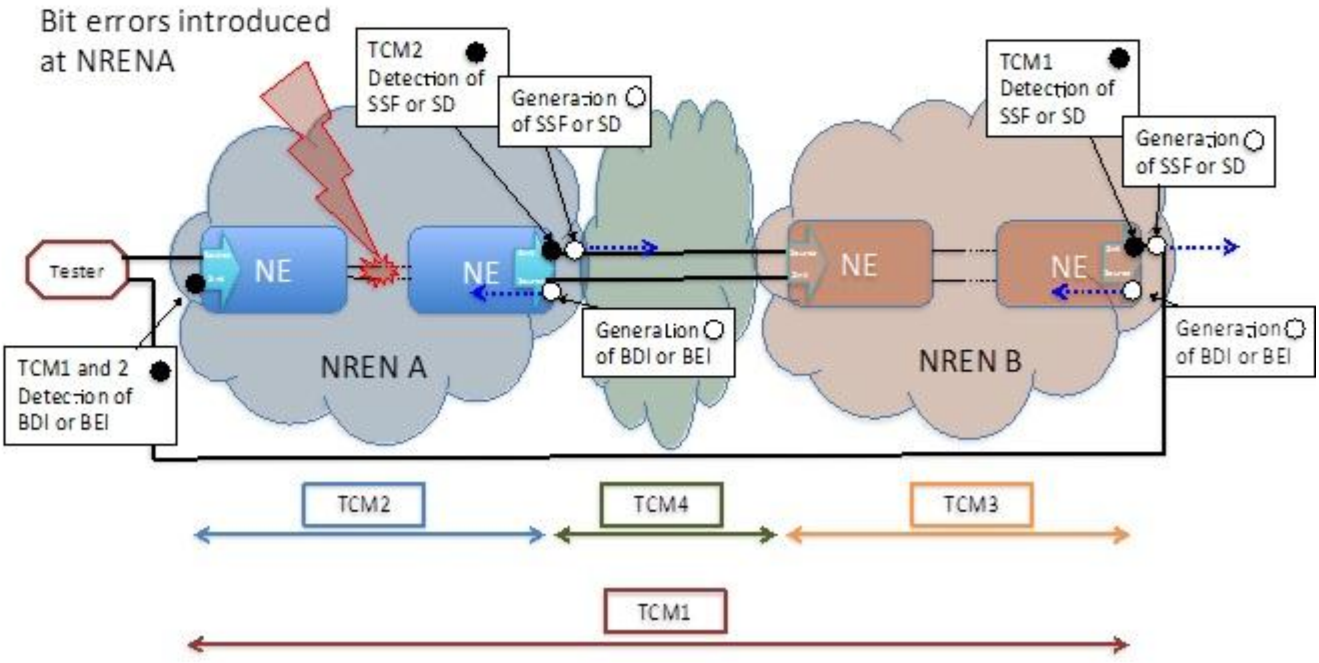


Figure 7.35: Results when bit errors were introduced at NREN A

The alarms indicate that the errors are coming from NREN A, as only TCM2 and TCM1 alarms are detected. The TCM2 alarm in NREN A is a SF/SD alarm, as opposed to the TCM1 alarm, which is BDI/BEI.

7.8.7 Bit Errors between TCM Sections

This scenario describes the situation where an event/alarm arises between two domains.

7.8.7.1 *Results*

The scenario gave the following results, also shown in Figure 7.36.

• NREN A:	TCM1 alarms (BDI/BEI), TCM4 alarms (BDI/BEI).
• NREN B:	TCM1 alarms (SSF/SD), TCM4 alarms (SSF/SD).
• Between NRENs:	TCM4 alarms (SSF/SD and BDI/BEI).

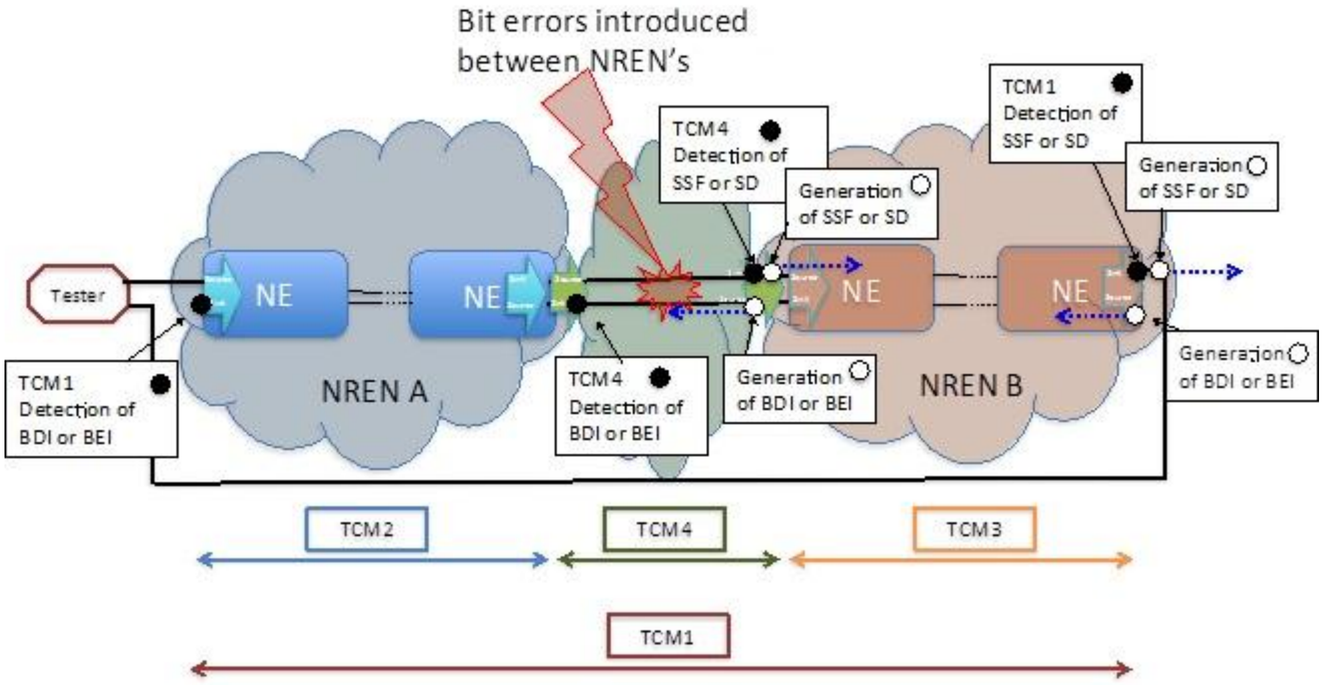

Bit errors introduced between NREN's
Tester
NE
NE
NE
NE
NREN A
NREN B
TCM 1 ● Detection of BDI or BEI
TCM 4 ● Detection of SSF or SD
Generation ○ of SSF or SD
TCM 1 ● Detection of SSF or SD
Generation ○ of SSF or SD
TCM 4 ● Detection of BDI or BEI
Generation ○ of BDI or BEI
Generation ○ of BDI or BEI
TCM 2
TCM 4
TCM 3
TCM1

Figure 7.36: Results when bit errors were introduced between NRENs

Although this scenario generated the most alarms, the indication is quite clearly that there's a problem between the two NRENs. This is due to the fact that no TCM2 or TCM3 alarms have arrived, only Client, TCM1 and TCM4 alarms.

Note that when adding the TCM4 between the NRENs, the TCM2 Sink point also becomes a TCM4 source point of NREN A. Similarly, the TCM3 Source point also becomes the TCM4 Sink point of NREN B.

The configuration of TCM in this type of situation needs to be coordinated between all the domains involved. It is recommended that the configuration be planned and analysed in advance.

7.8.8 Bit Errors Introduced in NREN B

This scenario was also demonstrated (although not in the plans).

7.8.8.1 *Results*

The scenario produced the results listed below and shown in Figure 7.37:

NREN A:	TCM1 alarms (BDI/BEI).
NREN B:	TCM1 alarms (SSF/SD), TCM3 alarms (SSF/SD).
Between NRENS:	No alarms.

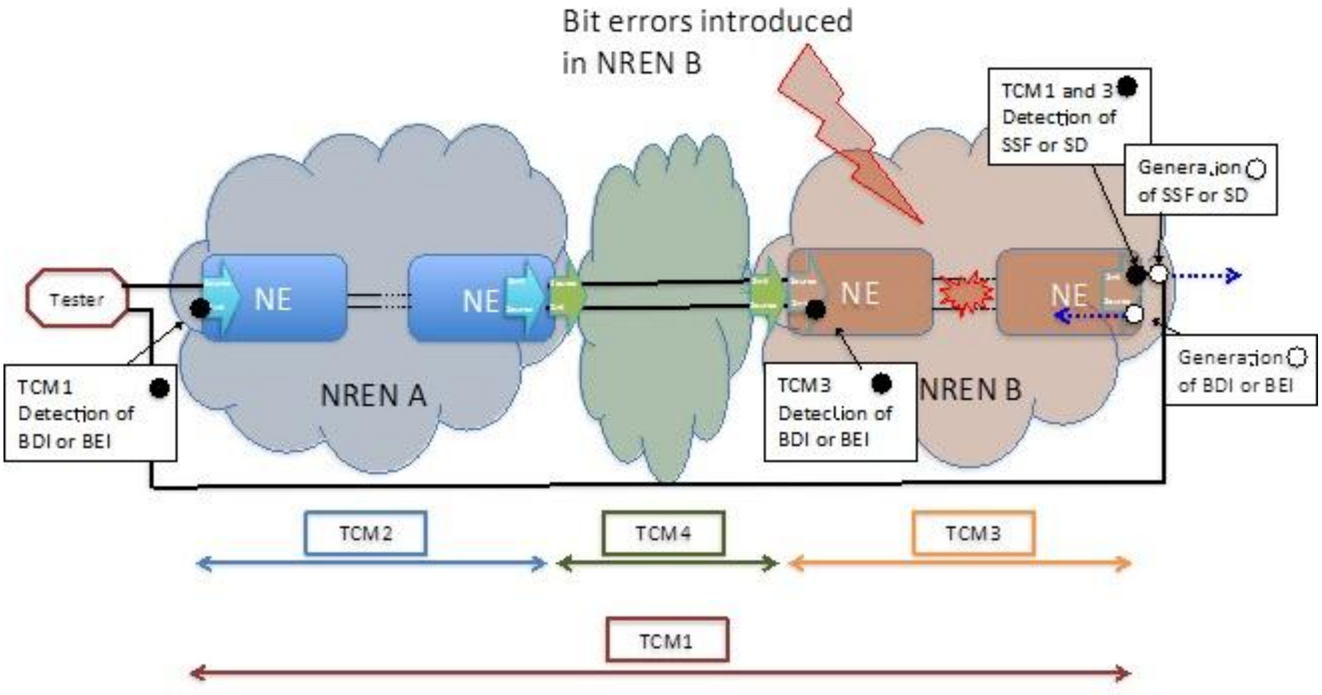

Bit errors introduced in NREN B
Tester
NE
NE
NE
NE
NREN A
NREN B
TCM1 Detection of BDI or BEI
TCM3 Detection of BDI or BEI
TCM1 and 3 Detection of SSF or SD
Generation of SSF or SD
Generation of BDI or BEI
TCM2
TCM4
TCM3
TCM1

Figure 7.37: Results when bit errors were introduced in NREN B

As only TCM1 and TCM3 alarms were detected in the network, the client or NREN B had the problem. Furthermore, as the SSF/SD alarms were only detected on the TCM3 level, the indication is that the problem originated in NREN B.

7.8.9 Fault Type, Fault Location (FTFL) Byte

The scenarios in Sections 7.8.5 to 7.8.8 above reflect possible real-life situations, but in reality the domains naturally consist of more than two network elements per domain. From a collaborative service point of view, the TCM alarms give clear indications as to which operator/domain is having problems. However, to pinpoint the exact location of the fault, another mechanism is introduced in the OTN G.709 specification [ITU-T G.709], the Fault Type, Fault Location (FTFL) channel.

The FTFL channel can be utilised to give operators an indication of the fault origin. FTFL is a 256 byte multi-frame signal providing fault status and fault location information. The concept is shown in Figure 7.38.

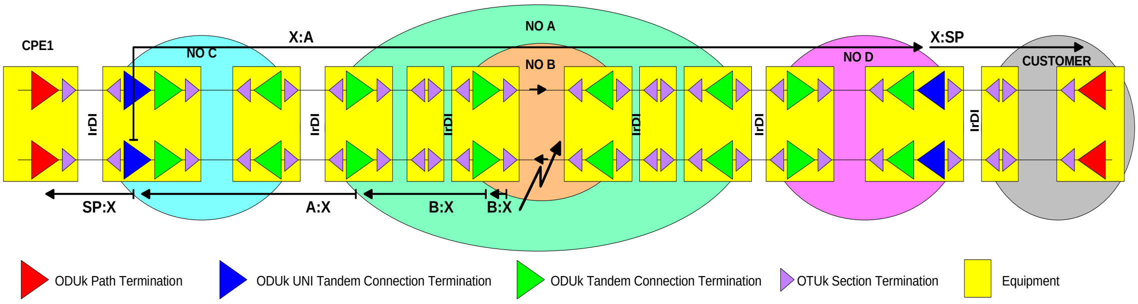


Figure 7.38: FTFL concept

At Section and Tandem Connection end points the FTFL status information is inserted in the forward direction upon detection of a SF or SD condition. At the User Network Interface (UNI), this information is extracted and sent in the opposite direction as backward information to locate the type and fault. The forward and backward identifier fields contain operator-specific identifiers/fields.

Although the FTFL feature was not available from any of the vendors at the time of demonstration, it could be a helpful capability used in conjunction with the TCM functionalities.

7.8.10 TCM Delay Measurements

ITU-T recommendation G.709 [ITU-T G.709] specifies a path monitoring function called delay measurement of ODUk path (DMp). The idea is to create a one-bit path delay signal defined to provide a momentary two-way transfer delay status, or a proactive, 15-minute and 24-hour transfer delay performance. This delay measurement can be used to select the appropriate route in case of service restoration, as explained in Section 7.7.7.4.

7.8.10.1 *Test Description*

The test consisted of the following steps:

1. Configure the largest possible network propagation.
2. Set up an OTN path through this network.
3. Configure the timers involved with the delay measurements on each end point.
4. Measure the embedded delay data.

7.8.10.2 *Results*

Despite the fact that this feature is quite new, the TCM delay measurement function was successfully demonstrated. It is important to mention that the test setup cannot be compared to real network situations where the fibre spans are of several kilometres. For this reason the delay measurements obtained in the lab were quite low.

7.8.11 Test Conclusions

The conclusions drawn from the TCM tests were as follows:

- The higher monitoring levels and more flexible surveillance possibilities offered by TCM make it potentially very beneficial in the NRENs' multi-domain, multi-vendor environment.
- The equipment behaved as expected in its defect detection and alarm handling in situations that are likely to occur in a multi-domain multi-vendor NREN environment.
- Utilising TCM functionality in a multi-vendor, multi-domain environment needs thorough planning and close collaboration between network operators, for example, planning the number of TCM levels and the end points of the TCM.
- The FTFL feature could be helpful in pinpointing the exact fault location when used in conjunction with the TCM functionalities. FTFL is especially relevant in multi-domain environments.
- The DMp path monitoring function can be used to select the appropriate route in case of service restoration.

7.9 Control Plane (GMPLS)

7.9.1 Technology Briefing

The addition of a control plane to the OTN architecture allows new functionality that covers the requirements for carrier-grade transport networks. GMPLS allows traditional Time-Division Multiplexing (TDM) networks to behave like IP networks, introducing the advantages of having routing and signalling while still being able to keep those capabilities that defined OTN as a carrier-grade technology. The control plane provides the following functionality:

- Topology discovery.
- Automated provisioning and decommissioning.
- Automated restoration.

The control plane (GMPLS) uses the Open Shortest Path First with Traffic Engineering (OSPF-TE) protocol to build topology information about the network. This information is present on every node in the network and Resource Reservation Protocol with Traffic Engineering (RSVP-TE) is used to signal the paths along the network during the building or tearing down process of the path.

Automated restoration allows the reduction of pre-allocated bandwidth for protection and of manual intervention by a network operator to configure a backup path, thus reducing both CAPEX and OPEX in the network. On the other hand, network planners would have to design the network so that there is always an alternative route with sufficient bandwidth available in case of a link failure.

Automated restoration is especially interesting used in combination with the recently added capability for measuring delay in the OAM bytes, as shown in Figure 7.21 on page 199. In very large networks with mesh topologies, it is necessary to be able to find the most suitable route across the network so that delay-sensitive applications are not affected by the selection of a long path.

7.9.2 Test Setup

The setup used to demonstrate GMPLS operation is described in Section 7.5.

7.9.3 Topology Discovery

7.9.3.1 *Test Objective*

The objective of the test was to verify topology discovery and topology update by means of the control plane routing protocol.

7.9.3.2 *Test Description*

The test consisted of the following steps:

1. Connect a new node to the pre-established network of two nodes as shown in Figure 7.39.
2. Monitor routing messages between nodes.
3. Verify that all nodes are aware of the complete topology.

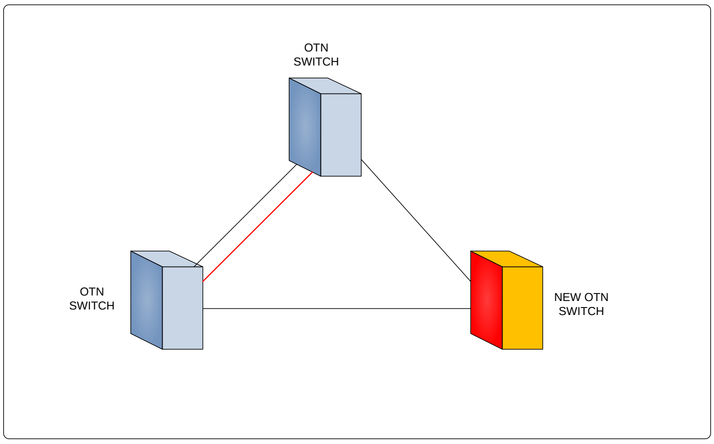


Figure 7.39: Topology discovery test setup

7.9.3.3 *Expected Results*

It was expected that only OSPF protocol interaction would be observed. When an OSPF router is brought up, a process of database synchronisation is initiated with the router's neighbours. At the end of the procedure, all routers should have the same view of the network topology. The synchronisation process is not trivial and involves multiple exchanges of information.

7.9.3.4 *Results*

This section presents a summary of the observed OSPF protocol interactions.

A protocol exchange diagram was generated, summarising the main points in the process. A network topology was generated, based on the observed OSPF Update messages (see Figure 7.40).

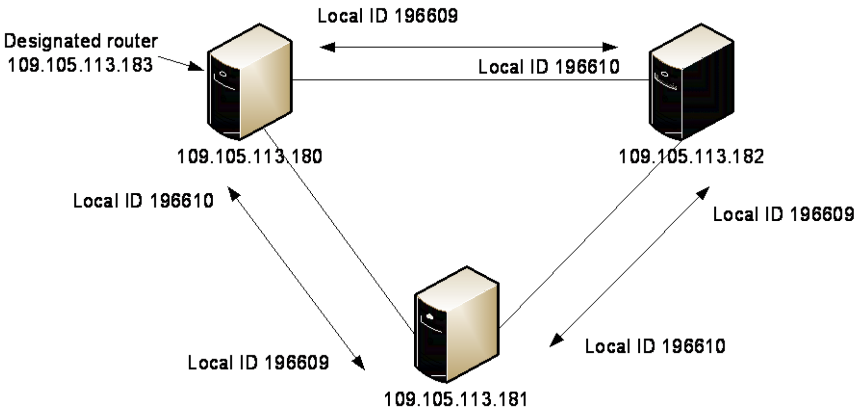


Figure 7.40: Network topology, router IP addresses and link local identifiers as advertised in the network

Node .180 is also the Designated Router for the network. Each link was advertised as a Traffic Engineering (TE) link – lambda switching type (Lambda Switch Capable – LSC), with lambda encoding type. All links had available bandwidth only at priorities 0 and 7. Multiple unknown Link State Advertisements (LSAs) and sub-Type Length Values (TLVs) were also advertised (possibly due to the vendor's proprietary implementation). The local interfaces for each node, related to each link, are indicated in Figure 7.40; in addition, Figure 7.41 presents the observed protocol interaction. Router .180 was the first one up, router .181 was second and, after synchronisation, router .182 was brought up.

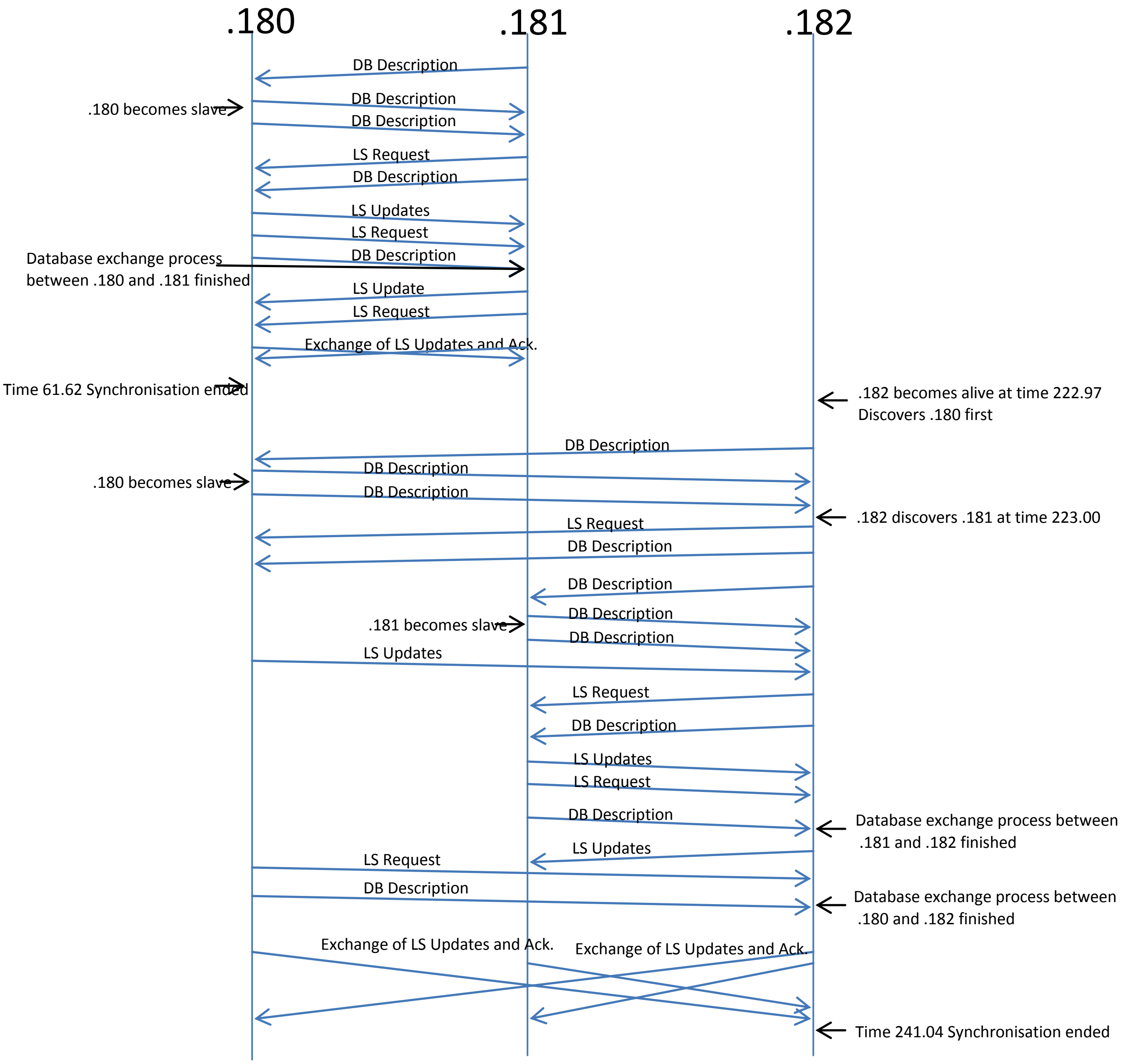

.180
.181
.182
DB Description
.180 becomes slave
DB Description
DB Description
LS Request
DB Description
LS Updates
LS Request
Database exchange process between .180 and .181 finished
DB Description
LS Update
LS Request
Exchange of LS Updates and Ack.
Time 61.62 Synchronisation ended
.182 becomes alive at time 222.97 Discovers .180 first
DB Description
.180 becomes slave
DB Description
DB Description
.182 discovers .181 at time 223.00
LS Request
DB Description
DB Description
.181 becomes slave
DB Description
DB Description
LS Updates
LS Request
DB Description
LS Updates
LS Request
DB Description
Database exchange process between .181 and .182 finished
LS Updates
LS Request
DB Description
Database exchange process between .180 and .182 finished
Exchange of LS Updates and Ack.
Exchange of LS Updates and Ack.
Time 241.04 Synchronisation ended

Figure 7.41: Protocol interaction diagram

During the test, the following procedure was implemented after final synchronisation: one of the routers was brought down, in order to observe the dynamic reaction of the control plane and the other routers. The resulting protocol interaction is shown in Figure 7.42. Router .181 was the one being disconnected.

The observed protocol interaction followed the standard. The nodes discovered each other and managed to create a consistent topology view. Under network topology change they managed dynamically to discover the changes and to disseminate the new network topology. A question remains as to why no link information was exchanged on the failure of the .181 router. This might have been due to the way the router was brought down (i.e. only in the control plane and not at the data plane).

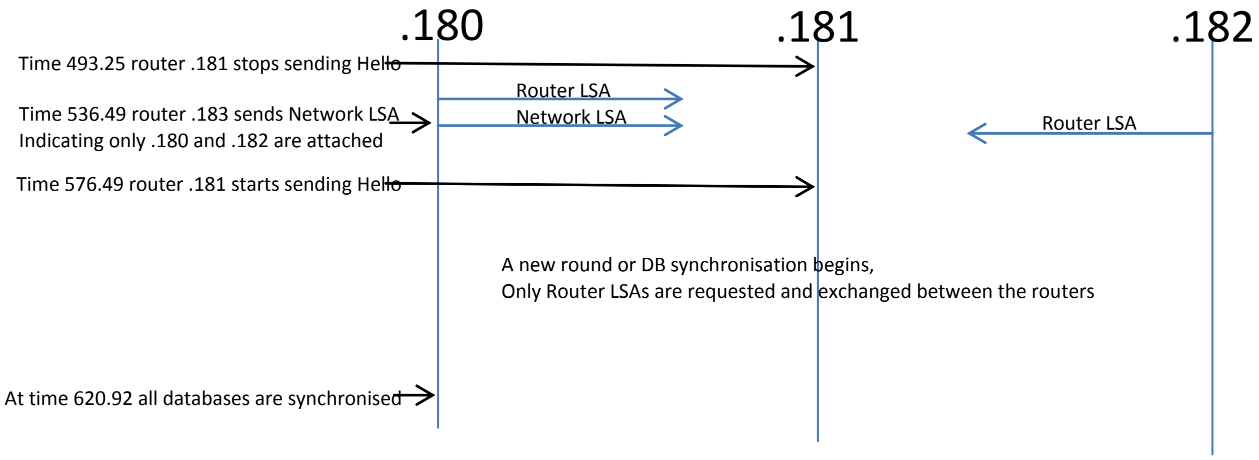

.180
.181
.182
Time 493.25 router .181 stops sending Hello
Router LSA
Network LSA
Time 536.49 router .183 sends Network LSA
Indicating only .180 and .182 are attached
Router LSA
Time 576.49 router .181 starts sending Hello
A new round or DB synchronisation begins,
Only Router LSAs are requested and exchanged between the routers
At time 620.92 all databases are synchronised

Figure 7.42: Protocol interactions with router .181 going down

7.9.4 Circuit Provisioning and Decommissioning

7.9.4.1 *Test Objective*

The objective of the test was to verify signalling messages during creation and tearing down of circuits.

7.9.4.2 *Test Description*

Using the same or similar network topology to that shown in Figure 7.40, the test consisted of the following steps:

1. Configure a circuit between two of the nodes.
2. Verify signalling messages between nodes during the creation of the circuit.
3. Decommission the recently created circuit.
4. Verify signalling messages between the involved nodes.

7.9.4.3 *Expected Results*

In the process of service creation and decommissioning, both GMPLS protocols have an active role. RSVP is responsible for reserving the resources for the new service or for releasing the resources of a decommissioned service, whereas OSPF must update the resource availability in the network by updating the routing tables in the routers. It was expected that OSPF activity would be seen straight after the nodes reserved resources for the upcoming connection (in the tested system, this happened after the first PATH message had been received in each node – i.e., in the downstream direction) and when resources were released (in the case of decommissioning).

7.9.4.4 *Results*

The processes of circuit provisioning and decommissioning are presented in two separate sections. Each section presents the protocol interaction diagram for both protocols as they appear in the obtained traces (i.e., the protocols are not separated, so that it is clear when OSPF is initiated with respect to which process in RSVP). A short discussion is presented for each section.

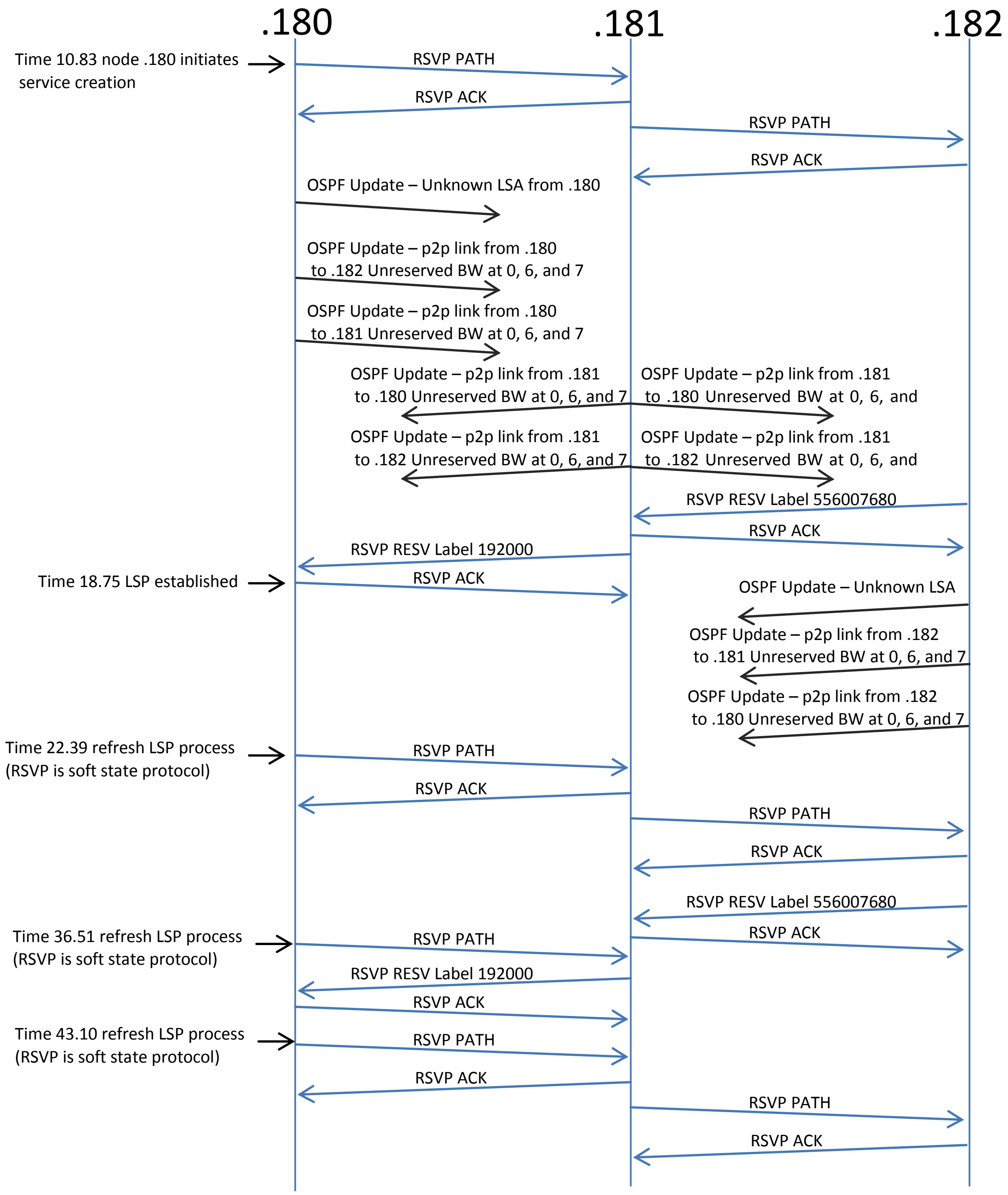
Circuit Provisioning

Figure 7.43 presents a combined protocol interaction diagram. Each protocol is represented with a different colour arrow: blue for RSVP messages, black for OSPF messages. Hello messages are not considered.

The observed protocol interaction followed the standard. RSVP-TE refreshed the state of the created LSP periodically. Since the initial OSPF-TE messages were not captured, it is difficult to say whether the updates noted the change in topology or not. However, considering the timing in the messages, it can be concluded that they did not. No OSPF messages were observed after the first RESV message was received in the source (by which point the LSP should be established).

A NOTIFY message was sent at Time 45.50 (there is a Notify Request object in the PATH messages). The indicated C-Type 3 of the ERROR_SPEC object indicates this is RSVP-TE ERROR_SPEC (RSVP has defined only Types 1 and 2). However, the included Error code and Error value are non-standard. According to the standard, an Error code of 0 (Confirmation) is combined with an Error value of 0, but here the Error value is 255 [RFC 2205].

The initial PATH messages show the administrative state of the LSP to be ADMIN-DOWN. This means that the LSP is down even though it is established – i.e. there is no traffic on it. Only after a few PATH/RESV exchange sessions was the ADMIN-DOWN status set to FALSE. This indicates that the path is operational and client traffic can be put on it.


.180
.181
.182
Time 10.83 node .180 initiates service creation
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK
OSPF Update – Unknown LSA from .180
OSPF Update – p2p link from .180 to .182 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .180 to .181 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0, 6, and
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0, 6, and
RSVP RESV Label 556007680
RSVP ACK
RSVP RESV Label 192000
Time 18.75 LSP established
RSVP ACK
OSPF Update – Unknown LSA
OSPF Update – p2p link from .182 to .181 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .182 to .180 Unreserved BW at 0, 6, and 7
Time 22.39 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK
RSVP RESV Label 556007680
RSVP ACK
Time 36.51 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP RESV Label 192000
RSVP ACK
Time 43.10 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK

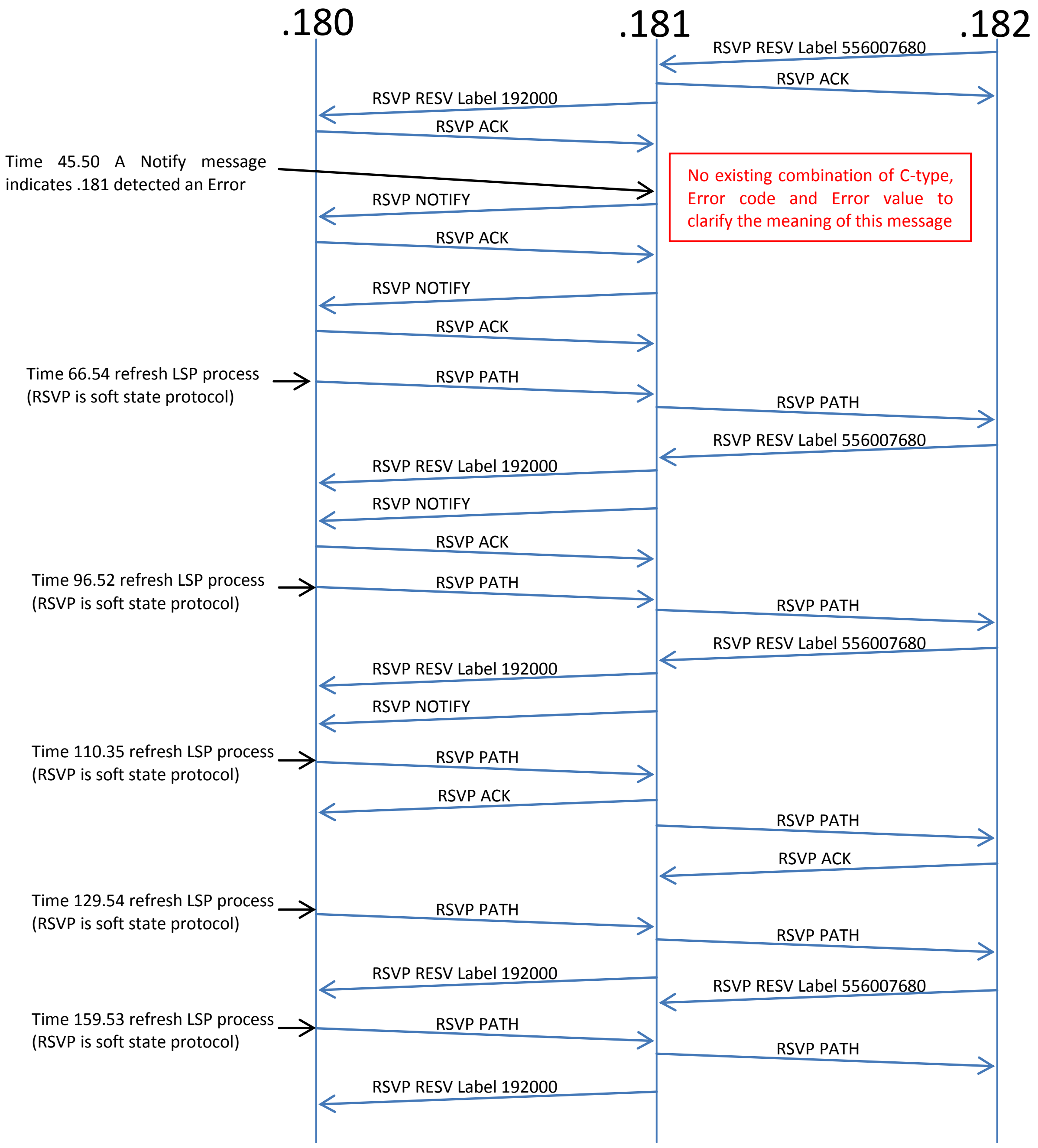

.180
.181
.182
RSVP RESV Label 556007680
RSVP ACK
RSVP RESV Label 192000
RSVP ACK
Time 45.50 A Notify message indicates .181 detected an Error
No existing combination of C-type, Error code and Error value to clarify the meaning of this message
RSVP NOTIFY
RSVP ACK
RSVP NOTIFY
RSVP ACK
Time 66.54 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP PATH
RSVP RESV Label 556007680
RSVP RESV Label 192000
RSVP NOTIFY
RSVP ACK
Time 96.52 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP PATH
RSVP RESV Label 556007680
RSVP RESV Label 192000
RSVP NOTIFY
Time 110.35 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK
Time 129.54 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP PATH
RSVP RESV Label 192000
RSVP RESV Label 556007680
Time 159.53 refresh LSP process (RSVP is soft state protocol)
RSVP PATH
RSVP PATH
RSVP RESV Label 192000

Figure 7.43: Combined RSVP-TE and OSPF-TE protocol interaction diagram on service creation

Circuit Decommissioning

As with circuit provisioning, Figure 7.44 presents a combined protocol interaction diagram; RSVP messages are represented with a blue arrow, OSPF messages with a black arrow. Hello messages are not considered.

The service deletion was initiated not by a PATH TEAR message but with a PATH message for updating the state of the LSP with the ADMIN STATUS object set to Delete-in-progress = TRUE. This is the so-called “graceful-delete” procedure [RFC 3473]. Instead of answering with a RESV message, the destination answered with a PATH ERROR message. The OSPF-TE protocol updated the states of the links after the service had been deleted. What is intriguing is that the amount of unreserved bandwidth disseminated per link is smaller than when the service was created (see the Circuit Provisioning test above). This should be further analysed.

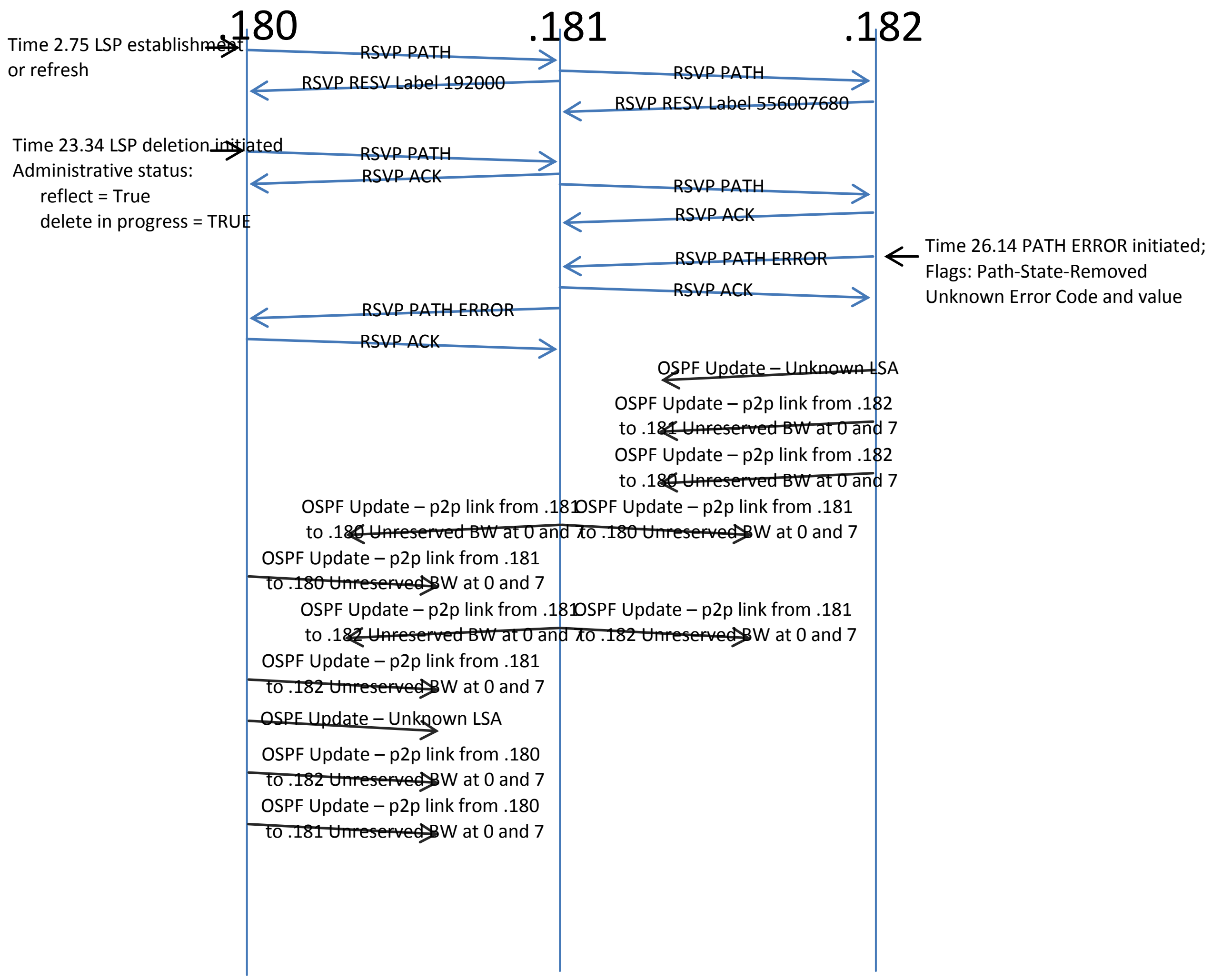

.180
.181
.182
Time 2.75 LSP establishment or refresh
RSVP PATH
RSVP RESV Label 192000
RSVP PATH
RSVP RESV Label 556007680
Time 23.34 LSP deletion initiated
Administrative status:
reflect = True
delete in progress = TRUE
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK
Time 26.14 PATH ERROR initiated;
Flags: Path-State-Removed
Unknown Error Code and value
RSVP PATH ERROR
RSVP ACK
RSVP PATH ERROR
RSVP ACK
OSPF Update – Unknown LSA
OSPF Update – p2p link from .182 to .181 Unreserved BW at 0 and 7
OSPF Update – p2p link from .182 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0 and 7
OSPF Update – Unknown LSA
OSPF Update – p2p link from .180 to .182 Unreserved BW at 0 and 7
OSPF Update – p2p link from .180 to .181 Unreserved BW at 0 and 7

Figure 7.44: Combined RSVP-TE and OSPF-TE protocol interaction diagram on service deletion

7.9.5 Service Restoration

7.9.5.1 *Test Objective*

The objective of the test was to verify service restoration by means of GMPLS.

7.9.5.2 *Test Description*

The test consisted of the following steps:

1. Configure a service between two end points.
2. Simulate a network failure in the link where the traffic is flowing.
3. Verify that the service has been restored.

7.9.5.3 *Expected Results*

It was expected that both protocols would be active during the restoration process. RSVP must be initiated at time of failure to establish a restoration path between the source and the destination of the failed connection. Furthermore, RSVP is responsible for releasing the resources occupied by the failed working connection. OSPF should be active when the new path is established, for updating the change in resource availability. Since the control plane implementation under test treats a link failure and a change in the link capacity as different types of events, the OSPF activity in this case is related not to the fact that there is a missing link in the topology, but to the fact that a new connection is established, i.e., for updating the resource availability.

7.9.5.4 *Results*

As with the circuit provisioning and decommissioning tests, Figure 7.45 and Figure 7.46 present combined protocol interaction diagrams; RSVP messages are represented with a blue arrow, OSPF messages with a black arrow. Hello messages are not considered.

At the beginning of the trace a series of LSP state refresh procedures was observed. At Time 99.04 a new LSP setup procedure was initiated. This is the restoration LSP request. It carries the same tunnel ID, but the LSP ID at the sender is different. The LSP was created under “Administratively down” status, which is typical for optical connections. Only when a PATH/RESV exchange with “Administratively down” status False is exchanged is the LSP up and ready to be used. At Time 118.52 the restoration path was active. After the first RESV message related to the restoration path was received at the source, the source released the resources for the working LSP by sending a RSVP PATH TEAR message.

OSPF updates started after the new LSP establishment had begun, by advertising the change in link capacity on link .180 – .182.

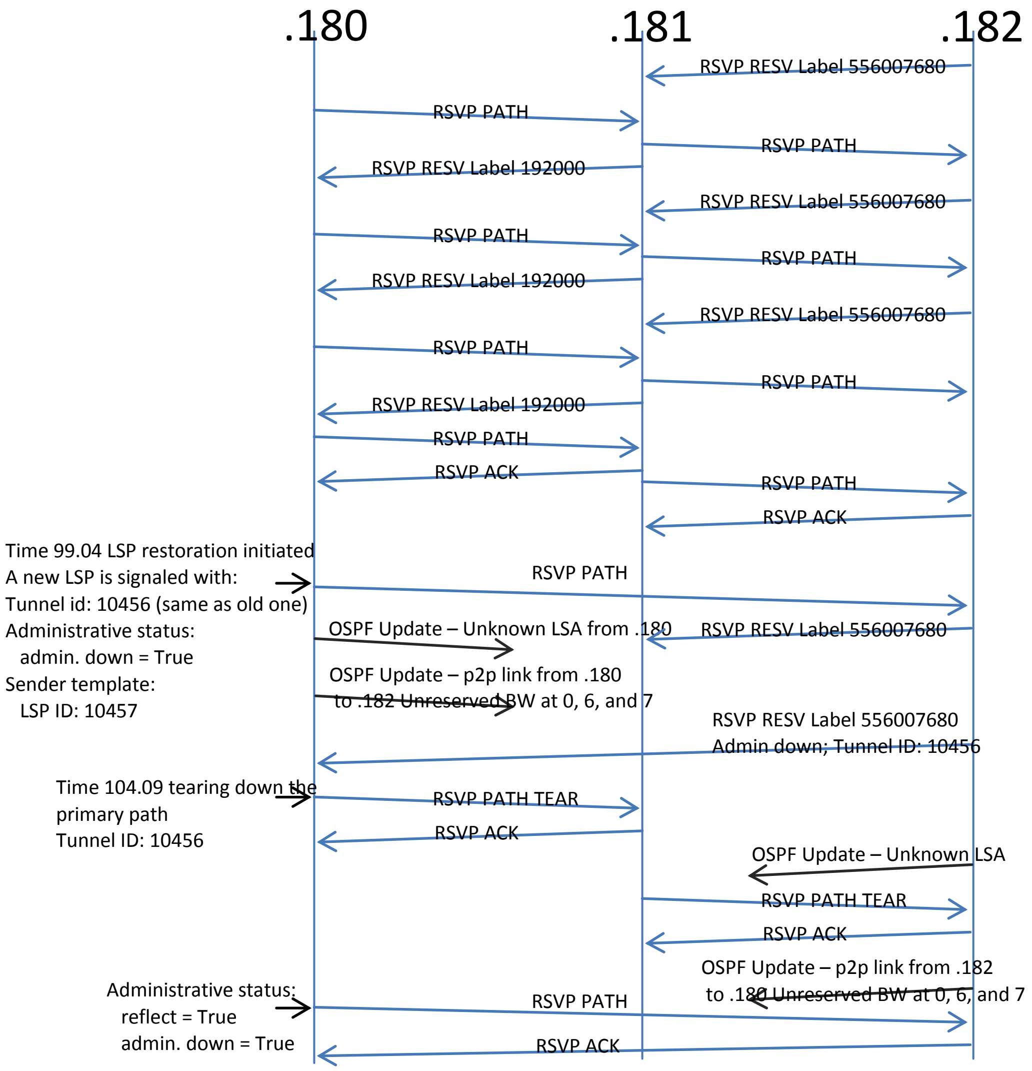

.180
.181
.182
RSVP RESV Label 556007680
RSVP PATH
RSVP PATH
RSVP RESV Label 192000
RSVP RESV Label 556007680
RSVP PATH
RSVP PATH
RSVP RESV Label 192000
RSVP RESV Label 556007680
RSVP PATH
RSVP PATH
RSVP RESV Label 192000
RSVP PATH
RSVP ACK
RSVP PATH
RSVP ACK
Time 99.04 LSP restoration initiated
A new LSP is signaled with:
Tunnel id: 10456 (same as old one)
Administrative status:
admin. down = True
Sender template:
LSP ID: 10457
RSVP PATH
OSPF Update – Unknown LSA from .180
RSVP RESV Label 556007680
OSPF Update – p2p link from .180
to .182 Unreserved BW at 0, 6, and 7
RSVP RESV Label 556007680
Admin down; Tunnel ID: 10456
Time 104.09 tearing down the
primary path
Tunnel ID: 10456
RSVP PATH TEAR
RSVP ACK
OSPF Update – Unknown LSA
RSVP PATH TEAR
RSVP ACK
OSPF Update – p2p link from .182
to .180 Unreserved BW at 0, 6, and 7
Administrative status:
reflect = True
admin. down = True
RSVP PATH
RSVP ACK

Figure 7.45: Combined RSVP-TE and OSPF-TE protocol interaction diagram on service restoration

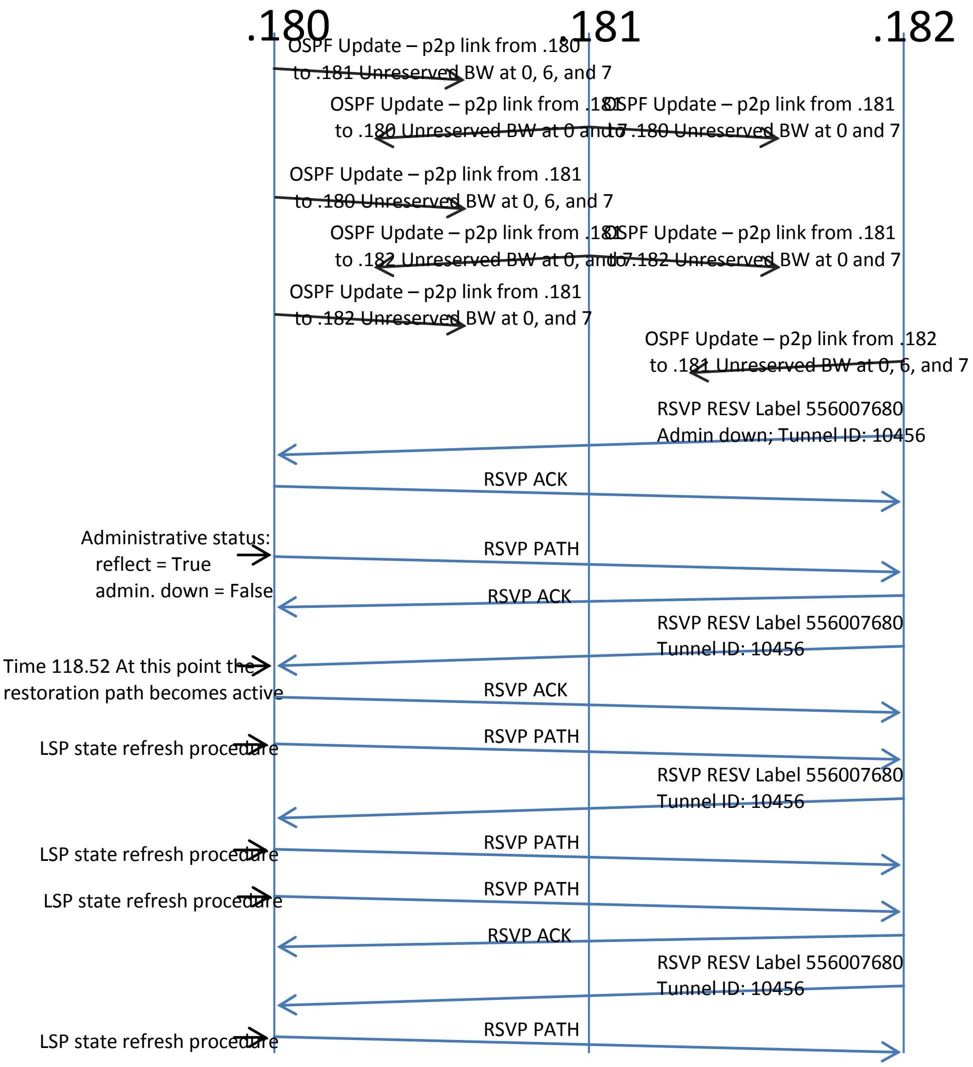

.180
.181
.182
OSPF Update – p2p link from .180 to .181 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .180 Unreserved BW at 0, 6, and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0, and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0 and 7
OSPF Update – p2p link from .181 to .182 Unreserved BW at 0, and 7
OSPF Update – p2p link from .182 to .181 Unreserved BW at 0, 6, and 7
RSVP RESV Label 556007680 Admin down; Tunnel ID: 10456
RSVP ACK
Administrative status: reflect = True admin. down = False
RSVP PATH
RSVP ACK
RSVP RESV Label 556007680 Tunnel ID: 10456
Time 118.52 At this point the restoration path becomes active
RSVP ACK
LSP state refresh procedure
RSVP PATH
RSVP RESV Label 556007680 Tunnel ID: 10456
LSP state refresh procedure
RSVP PATH
LSP state refresh procedure
RSVP PATH
RSVP ACK
RSVP RESV Label 556007680 Tunnel ID: 10456
LSP state refresh procedure
RSVP PATH

Figure 7.46: Combined RSVP-TE and OSPF-TE protocol interaction diagram on service restoration (cont.)

7.9.6 Test Conclusions

The conclusions drawn from the control plane (GMPLS) tests were as follows:

- The addition of a control plane to the OTN architecture provides new functionality – topology discovery and update, automated provisioning and decommissioning, and automated restoration – that covers the requirements for carrier-grade transport networks.
- The equipment supports topology discovery and update by means of the control plane routing protocol. The observed protocol interaction followed the standard. Unexpected behaviour (no link information was exchanged on the failure of the .181 router) requires further investigation.
- The equipment supports signalling messaging during creation and tearing down of circuits. The observed protocol interaction followed the standard with one exception (Error code/value mismatch).

During circuit decommissioning, the amount of unreserved bandwidth disseminated per link was smaller than when the service was created; this should be further analysed. However, it could be related to a bug in this specific implementation.

- The equipment supports service restoration by means of GMPLS. The observed protocol interaction followed the standard.
- Automated restoration allows the reduction of pre-allocated bandwidth for protection and of manual intervention by a network operator to configure a backup path, thus reducing both CAPEX and OPEX in the network. However, network planners would have to design the network so that there is always an alternative route with sufficient bandwidth available in case of a link failure.
- Automated restoration is especially interesting used in combination with the recently added capability for measuring delay in the OAM bytes, as shown in Figure 7.21 on page 199. In very large networks with mesh topologies, it is necessary to be able to find the most suitable route across the network so that delay-sensitive applications are not affected by the selection of a long path.

7.10 Conclusions

OTN is considered to be a promising technology to take over the established SDH networking layer and boost the network with more powerful switching, mapping and survivability functionality in the digital domain. In addition, it brings seamless integration to the optical domain and provides a common vehicle for mapping, switching and transporting all types of client signals. These are some of the main reasons why this technology was selected to be part of JRA1 Task 1 scope.

The initial intention of the Task was to test OTN switches in NRENs’ test facilities in close cooperation with the equipment manufacturers. It was thought that a joint effort between the NRENs participating in the Task and vendors could bring major benefits for both the GN3 project and the vendors. To achieve this, the Task approached vendors to try to secure their cooperation and involvement in the GN3 project. Quite early in the process it became clear that getting access to OTN equipment was a complicated task for various reasons. It therefore took considerable time and effort to achieve the primary objective of this chapter, which was not to benchmark or disclose what features are available at which vendor, but instead to demonstrate the technology and visualise whether and how it can be of benefit to the NREN community. The vendors that contributed to this study have recognised the importance of disseminating knowledge to the NREN community and it is evident that the benefit for NRENs is the ability and willingness to be early adopters of the technology and research.

The testing of existing OTN platforms shows that the products are reaching market maturity. The most important functionalities, such as switching on different ODU levels and survivability based on different SNC parameters, including TCM, are already available. ODUflex is not fully implemented yet, but the tests confirmed that the basic functionalities to make ODUflex possible are working well. The integration of a control plane into the OTN technology adds important functionality and intelligence and opens up possibilities for dynamic provisioning tools integration, which is a major requirement for transport technologies in NRENs’ transport networks.

The testing showed that OTN is a transport technology that could provide major advantages to the NREN community. It is obvious that with the latest developments in optical networking, where the capacity is reaching levels of 100G and soon 400G and 1T, a multiplexing and switching layer will be needed to gather the traffic into the core network in a flexible, scalable and standardised manner, at the same time being capable of mapping all types of client signals, from Constant Bit Rate such as STM-64 to packet-based signals such as MPLS-TP traffic.

Another important aspect to consider, given the multi-domain, multi-vendor NREN environment, is the fact that OTN is standardised by the ITU-T and there is a common effort by all vendors to comply with the ITU-T recommendations. A standardised implementation of Tandem Connection Monitoring across domains would provide NRENs deploying services across domains with full visibility and a common OAM infrastructure. This is one of the major advantages of OTN compared to its main competitors.

The integration of a control plane allows the nodes in the network to be aware of the total topology of the network and the capabilities of the different links. This allows automatic restoration of services in the best way possible by taking into account parameters such as cost and latency. Restoration by means of control plane capabilities avoids the need for immediate manual intervention in the case of a network outage.

Finally, it is the intention of this chapter to help NRENs to understand OTN and help them with the process of planning and designing their next-generation core networks. In addition, it is intended to trigger a discussion about OTN and other technologies or to add more relevant, specific information to the ongoing debate. The information in this chapter will be used in future technology workshops.

8 Cross-Activity Work

8.1 Overview

This chapter describes the tests conducted by JRA1 Task 1 participants on the prototype Carrier Grade Ethernet (CGE) technology proxy (TP) designed and implemented for AutoBAHN (JRA2 Multi-Domain Network Service Research, Task 2 Hybrid Network Provisioning, where “hybrid” refers to provisioning over different technologies in a multi-domain environment). The TP is a module that is able to communicate with the AutoBAHN software (see Figure 8.1 and [AutoBAHN]) and reserve resources for the Bandwidth on Demand (BoD) service. It was developed to support CGE technology and was based on the implementations of the relevant standards and technologies by Extreme Networks, in particular on the BlackDiamond® 12804 switches, running ExtremeXOS Network Operating System version 12. The work was supported by Essex University, and used the Essex University testbed.

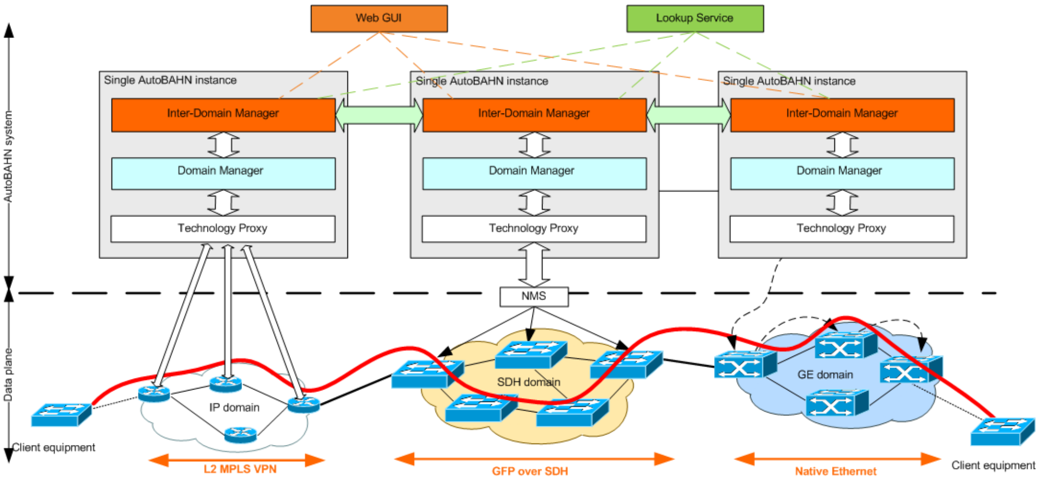


Figure 8.1: AutoBAHN concept

Although the CGE TP testing relates to a specific use case rather than to the more generic CCTNT implementations described in the previous chapters of this document, it is included here as it nonetheless represents a technology implementation trial, and, in addition, is an example of the cross-Activity work that features in the Task's objectives and success criteria, illustrating how the research Activities can contribute directly to GÉANT service support and development.

The chapter begins with a description of Carrier Grade Ethernet, focusing on the supporting technologies and service types that are of relevance to AutoBAHN, namely Provider Bridges (PB) / Q-in-Q, Provider Backbone Bridge Traffic Engineering (PBB-TE) and E-Line service types.

It then describes the test, stating the purpose, describing the testbed infrastructure (including detailed configuration commands), and outlining the stages, expected results and results before drawing conclusions.

The TP was successfully able to configure the testbed switches, set up a PBB-TE tunnel and enable Layer 2 connectivity between the desired end points. The compatibility of CGE technology with the BoD service and with the AutoBAHN tool in particular was therefore verified. However, extensive configuration needs to be performed on each underlying device, as there is currently no Network Management System that could abstract these operations. In addition, not all of the BoD parameters were available on the selected equipment, notably bandwidth limits for the created paths. As a result, support for BoD in a CGE domain has to be combined with over-provisioning / dimensioning of the service in terms of capacity.

Some test stages, including circuit creation between testbed edge clients, are still ongoing; further results will be available in due course.

8.2 Technology Briefing

The information in this section is taken from "Deliverable DJ2.2.1: Specification of enhancements and developments for the AutoBAHN system" [DJ2.2.1]. Further information about PBB-TE is provided in Chapter 5 of the current document.

Ethernet technology provides high bandwidth at a relatively low cost, easy installation, and the capability for point-to-point and point-to-multipoint operation. These properties caused Ethernet to emerge as a WAN protocol even in provider backbones.

Carrier Grade Ethernet is defined by the Metro Ethernet Forum (MEF) as "a ubiquitous, standardised, carrier-class Service and Network defined by five attributes that distinguish it from familiar LAN based Ethernet" [MEF-CGE]. These attributes are:

- Standardised Services.
- Scalability.
- Reliability.
- Service Management.
- Quality of Service (QoS).

The basic technologies defined by the Institute of Electrical and Electronics Engineers (IEEE) that provide essential support for backbone deployment of Ethernet are:

- Link Layer Discovery Protocol (LLDP) [IEEE 802.1ab].
- Provider Bridges (PB) / Q-in-Q [IEEE 802.1ad].
- Provider Backbone Bridges (PBB) / MAC-in-MAC [IEEE 802.1ah].
- Provider Backbone Bridge Traffic Engineering (PBB-TE) [IEEE 802.1Qay].
- Ethernet Operation, Administration and Maintenance (OAM) [IEEE 802.1ag] (also published as ITU-T Y.1731).

LLDP has little effect on AutoBAHN, as its main purpose is to discover the elements and the topology of an Ethernet-based network. The discovered topology will, however, have to be maintained by AutoBAHN, managed and abstracted.

Q-in-Q and MAC-in-MAC are essentially encapsulating technologies that enable customer VLANs (or entire MAC addresses) to be transparently handled by the backbone network, achieving scalability and manageability. This is discussed further in Section 8.2.1 below.

802.1Qay is the amendment of 802.1. This is probably most relevant to the AutoBAHN service model as it allows explicitly selected traffic-engineered paths within Provider Backbone Bridge Networks (PBBNs). It is expected that functionality adhering to 802.1Qay will be offering explicitly routed, bandwidth-guaranteed paths over a PBB core, so the corresponding AutoBAHN technology proxy will have to be developed accordingly. This is discussed further in Section 8.2.2 below.

802.1ag provides Operation, Administration and Maintenance (OAM) functions. These are not directly relevant to the dynamic circuit reservation and provisioning functionality required by AutoBAHN. However, it is expected that such functions will be essential for monitoring the performance of and troubleshooting AutoBAHN-provisioned circuits, and thus exploited accordingly by advanced releases of the AutoBAHN CGE technology proxy.

The MEF’s definition of Carrier Ethernet specifies two service types (E-Line and E-LAN), which can be supported in port mode (as Ethernet Private Line (EPL) and EP-LAN (Ethernet Private LAN)) or VLAN-multiplexed mode at the User Network Interface (UNI) (Ethernet Virtual Private Line (EVPL) and EVP-LAN (Ethernet Virtual Private LAN)). In MEF 6.1 [MEF 6.1] an additional service type (E-Tree) has been defined, with corresponding Ethernet Private Tree (EP-Tree) and Ethernet Virtual Private Tree (EVP-Tree) modes. Services are implemented using Ethernet Virtual Connections (EVCs).

Since AutoBAHN deals with point-to-point connections, the focus is on the support of E-Line service types: EPL (Ethernet Wire Service (EWS) in Cisco terminology) and EVPL (Ethernet Relay Service (ERS) in Cisco terminology) services.

The level of implementation of the AutoBAHN Carrier Ethernet technology proxy depends on the underlying implementations of the aforementioned standards. PBB/PBB-TE functionality, at least in its early phases, is primarily supported at the management plane. Therefore the specification of a proxy is heavily reliant on individual implementations rather than a standardised control plane.

8.2.1 Provider Bridges / Q-in-Q – IEEE802.1ad

IEEE 802.1ad (Provider Bridges) [IEEE 802.1ad] is an amendment to IEEE standard IEEE 802.1Q-1998 (also known as Q-in-Q or Stacked VLANs) [IEEE 802.1Q], intended to develop an architecture and bridge protocols to provide separate instances of the MAC services to multiple independent users of a Bridged Local Area Network (BLAN) in a manner that does not require cooperation among the users, and requires a minimum of cooperation between the users and the provider of the MAC service. IEEE 802.1ad was approved on 8 December 2005 and published on 26 May 2006.

The idea is to provide, for example, the possibility for customers to run their own VLANs inside the VLAN provided by the service provider. This way the service provider can just configure one VLAN for the customer, which the customer can then treat as if it were a trunk.

8.2.2 Provider Backbone Bridge Traffic Engineering (PBB-TE) – IEEE802.1Qay

Provider Backbone Bridge Traffic Engineering (PBB-TE) is an approved networking standard, IEEE 802.1Qay-2009 [IEEE 802.1Qay]. PBB-TE adapts Ethernet technology to carrier class transport networks. It is based on the layered VLAN tags and MAC-in-MAC encapsulation defined in IEEE 802.1ah (Provider Backbone Bridges (PBB) [IEEE 802.1ah], but it differs from PBB in eliminating flooding, dynamically created forwarding tables, and spanning tree protocols. Compared to PBB and its predecessors, PBB-TE behaves more predictably and its behaviour can be controlled more easily by the network operator, albeit at the expense of requiring up-front connection configuration at each bridge along a forwarding path. PBB-TE OAM is usually based on IEEE 802.1ag. It was initially based on Nortel’s proprietary Provider Backbone Transport (PBT).

PBB-TE’s connection-oriented features and behaviour, as well as its OAM approach, are inspired by Synchronous Digital Hierarchy / Synchronous Optical Network (SDH/SONET). PBB-TE can also provide path protection levels similar to the Uni-directional Path Switched Ring (UPSR) protection in SDH/SONET networks.

Further information about PBB-TE is provided in Chapter 5 on page 101.

8.3 Test Objective

The objective of the test was to verify the implementation and operation of the CGE TP developed to enable AutoBAHN to reserve resources for the Bandwidth on Demand (BoD) service supported by CGE.

8.4 Test Infrastructure

The Essex University testbed infrastructure is shown in Figure 8.2 below; the resources are described in Table 8.1.

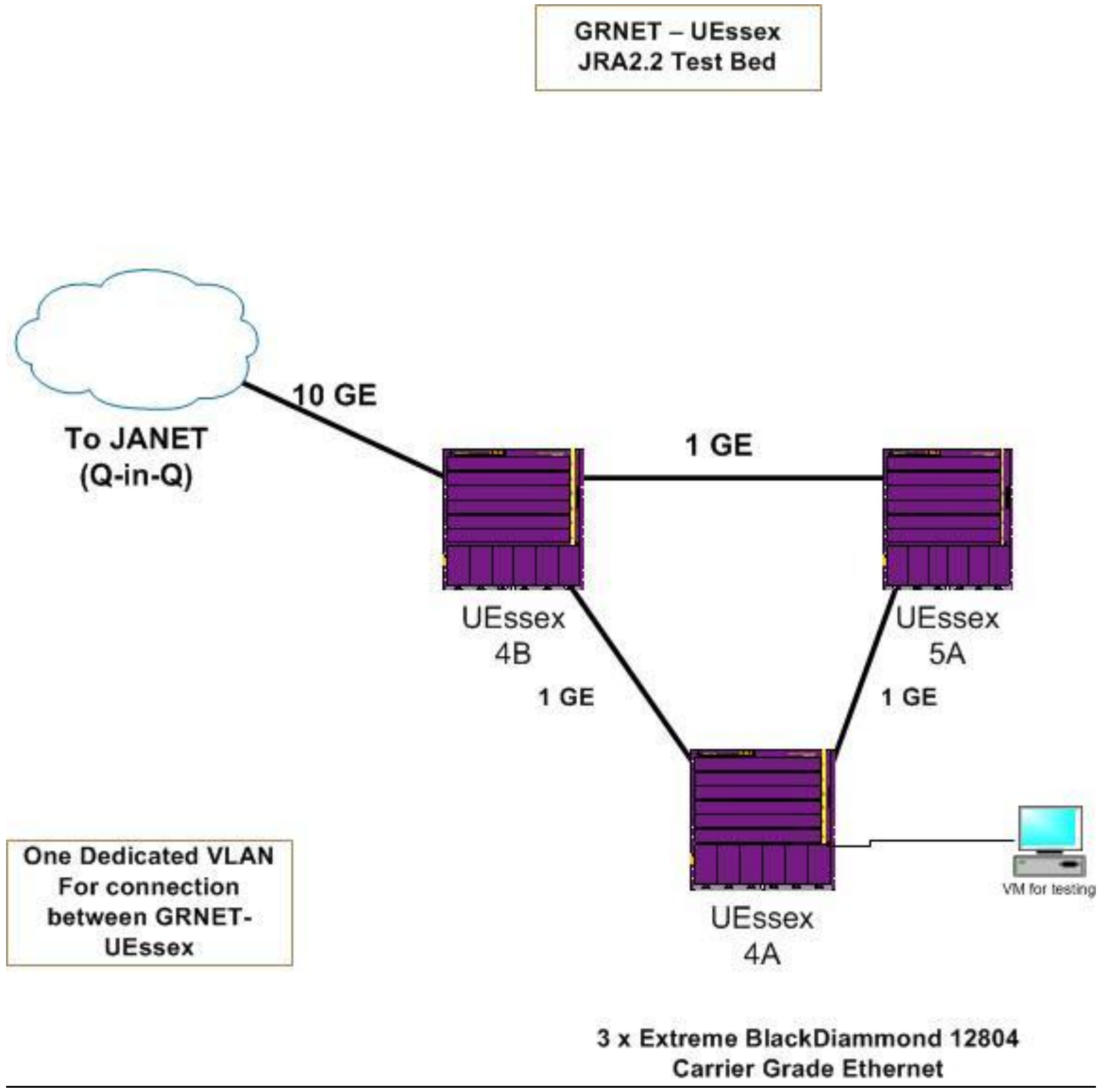


Figure 8.2: Essex University testbed

Resource Name	Resource Type	MAC Address	Description
UEssex_4A	Extreme switch	00:04:96:1E:FD:60	3 Extreme BlackDiamond® 12804 carrier-grade Ethernet switches running ExtremeXOS Network Operating System version 12 [ExtremeXOS_CmdRef, ExtremeXOS_Concepts] supporting OAM, connection to JANET, and configuration and creation of the PBB-TE paths used to carry the dynamically provisioned VLAN traffic on well-defined paths through the testbed network. Each BD switch was connected to the other two BD switches, forming a triangle. Switch 4B was
UEssex_4B	Extreme switch	00:04:96:3B:23:10	
UEssex_5A	Extreme switch	00:04:96:3B:09:60	

Resource Name	Resource Type	MAC Address	Description
			also connected to the external network (JANET) and switch 4A was also connected to a Virtual Machine end host.
-	Overture switch	-	Overture ISG24 switch to support OAM.
-	Servers	-	Terminal Access Controller Access-Control System (TACACS) and OpenVPN servers for access to and authorisation of the resources.
-	Virtual Machine end host	08:00:27:53:5e:5b	VM as traffic source, sink and tester. Connected to switch 4A.

Table 8.1: Testbed resources

The CGE TP developed for AutoBAHN was tested using a simplified client application that provided the TP with incoming requests.

8.4.1 PBB-TE Tunnel

For the purpose of the PBB-TE tunnel function, two Bridged VLANs (B-VLANs) were initially defined. They were created statically on the interface, so that the TP did not have to recreate them every time it processed a reservation request from AutoBAHN. The following additional setup measures were also taken:

- Address learning on the switches was disabled to gain complete control over the PBBN path, since each path is a static route. On a PBB-TE link, all broadcast, multicast, and unicast packets with an unknown destination MAC address are discarded. The PBB-TE trade-off is that it takes away the Ethernet self-configure and self-healing mechanisms that are supported with the help of Multiple Spanning Tree Protocol (MSTP), Spanning Tree Protocol (STP) or Shortest Path Bridging (SPB). AutoBAHN was relied on for selecting the desired route.
- Flooding was disabled to ensure that all path traffic was limited to the configured path.
- Forwarding Database (FDB) entries were configured on the egress port of each switch along the route, to define the possible paths.

The configuration commands for these measures are shown in the next section.

8.5 Configuration

This section presents the following configuration commands:

- Static B-VLAN – on switches 4A, 4B, 5A.

- TP Framework:
  - Add method:
    - Create S-VLAN – on switches 4A and 4B.
    - Create FDB Entry – on switches 4A and 4B.
  - Remove method – on switches 4A and 4B.

8.5.1 Static B-VLAN Configuration Commands

The first B-VLAN (b1) included ports that belonged only to switches UEssex_4A and UEssex_4B, whereas the second B-VLAN (b2) included all the ports of switch 5A and appropriate ports of switches UEssex_4A and UEssex_4B. The configuration commands applied on each Black Diamond 12804 switch regarding the B-VLANs were as follows:

Switch UEssex_4A

```
create bvlan b1
create bvlan b2

configure bvlan b1 tag 100
configure bvlan b2 tag 200

configure bvlan b1 add port *4A_4B* tagged
configure bvlan b2 add port *4A_5A* tagged

disable learning bvlan b1
disable learning bvlan b2

disable flooding bvlan b1
disable flooding bvlan b2

create fdbentry *MAC_4B* vlan b1 *port_4A_4B*
create fdbentry *MAC_4B* vlan b2 *port_4A_5A*
```

Switch UEssex_4B

```
create bvlan b1
create bvlan b2

configure bvlan b1 tag 100
configure bvlan b2 tag 200

configure bvlan b1 add port *4B_4A* tagged
configure bvlan b2 add port *4B_5A* tagged
```

```
disable learning bvlan b1
disable learning bvlan b2

disable flooding bvlan b1
disable flooding bvlan b2

create fdbentry *MAC_4A* vlan b1 *port_4B_4A*
create fdbentry *MAC_4A* vlan b2 *port_4B_5A*
```

Switch UEssex_5A:

```
create bvlan b2

configure bvlan b2 tag 200

configure bvlan b2 add port *5A_4B*,*5A_4A* tagged

disable learning bvlan b2

disable flooding bvlan b2

create fdbentry *MAC_4A* vlan b2 *port_5A_4A*
create fdbentry *MAC_4B* vlan b2 *port_5A_4B*
```

8.5.2 TP Framework Configuration Commands

8.5.2.1 *Add Method*

Create S-VLAN

On BlackDiamond 12804 switches, Service VLANs (S-VLANs) support customer VLAN traffic. An S-VLAN was therefore created to encapsulate VLAN traffic from external devices and forward it to the PBB-TE tunnel created on the testbed. Supposing the TP accepts a request from AutoBAHN, defining an S-VLAN named s1 and tagged 1000, the network reservation is implemented by applying the following commands on the ingress and egress devices:

```
create svlan s1

configure svlan s1 tag 1000

configure svlan add port *

configure bvlan ** add svlan s1
```

Notes:

> * Ingress and egress ports of edge interfaces (switches UEssex_4A, UEssex_4B) defined in the request.
>
> ** Using a switch-case function, TP chooses the appropriate B-VLAN, according to the initial static definition of the B-VLANs and the number of nodes given in AutoBAHN's request. If the number is three, that means that the defined route includes all switches and, as a result, B-VLAN b2 is chosen. Otherwise, when only edge switches UEssex_4A and UEssex_4B participate in the given path, B-VLAN b1 is chosen.

Create FDB Entry

Using an if-condition indicating each switch, static FDB entries were created on the switch's egress port facing the outer network, together with the appropriate mac-bindings, based on the defined B-VLAN and using once again the same switch-case as in the initial choice of the B-VLAN:

Switch UEssex_4A

```
create fdbentry *MAC_PC* vlan s1 *

create mac-binding bvlan ** *MAC_4B* svlan s1 *MAC_PC*
```

Switch UEssex_4B

```
create fdbentry *MAC_janet* vlan s1 *

create mac-binding bvlan ** *MAC_4B* svlan s1 *MAC_janet*
```

Switch UEssex_5A

No further configuration on the core switch UEssex_5A was applied, since it participated to a minor degree in the creation of the PBB-TE tunnel.

8.5.2.2 *Remove Method*

Since no configuration was applied on core switch UEssex_5A, the configuration commands below were applied only on the edge switches UEssex_4A and UEssex_4B:

```
delete svlan s1

configure svlan s1 delete ports all

delete mac-binding bvlan b2
delete mac-binding bvlan b1

delete fdbentry *MAC_PC* vlan s1
delete fdbentry *MAC_janet* vlan s1
```

The whole configuration for the TP Framework is shown in Appendix B on page 248. For more information, please refer to the “Technology Proxy Framework for AutoBAHN: User Guide” [TPFramework].

8.6 Test Description

The test consisted of the following stages. Only the first has been completed at the time of writing; the others are ongoing.

1. Testing of the Extreme BlackDiamond commands to support TP operations (completed), including obtaining access to the relevant switch commands and testing them independently of the TP.
2. Development of the TP (ongoing), including obtaining access to the Virtual Machine end host and setting up the TP there. Also making sure that the VM has proper access for automated configuration of the switches.
3. Testing of the TP operation and path creation (ongoing), including running the TP software, which sends configuration commands to the switches, and then running validation tests (using ping, iperf) among the path end points. The TP is fed a reservation request using a dedicated command line client.

8.7 Expected Results

The expected results of the testing activity were as follows:

- Verification of the compatibility of CGE technology for supporting the Bandwidth on Demand service, and in particular for cooperating with AutoBAHN tool deployment.
- Identification of the AutoBAHN / BoD features that can be supported by underlying CGE technology and the features that cannot be supported.
- Implementation of a prototype technology proxy for Extreme BlackDiamond equipment.
- Successful creation of circuits on demand between clients attached to the testbed network using a sample AutoBAHN deployment.

8.8 Results

The TP was successfully able to configure the testbed switches, set up a PBB-TE tunnel and enable Layer 2 connectivity between the desired end points.

The compatibility of CGE technology with the BoD service and with the AutoBAHN tool in particular was verified. CGE commands were used in Extreme BlackDiamond switches to set up fixed deterministic paths within the testbed carrying particular VLAN traffic from / to the connected edge points. However, some of the parameters of the BoD service were not available. In particular, limiting the bandwidth of created paths was not achieved. The authors believe this limitation to be equipment specific rather than technology specific.

A prototype TP was implemented and tested for successful configuration of the network devices.

Some tasks, including circuit creation between testbed edge clients, are still ongoing; further results are expected in July 2012.

8.9 Conclusions

This chapter has presented a high-level design of a CGE TP based on Extreme Networks BlackDiamond 12804 switches. The test confirmed that implementing such a CGE TP is viable, but that extensive configuration needs to be performed on each underlying device, as there is currently no Network Management System that could abstract these operations. The Domain Manager (DM) therefore needs to be aware of the detailed switch topology information and to pass a detailed request to the TP that it could then “translate” into the appropriate configuration commands to be sent to the BlackDiamond switches.

Concerning the bandwidth limiting issue, it was discovered that bandwidth limits could not be defined on the selected equipment (BlackDiamond 12800 series) since every available way of defining a desired bandwidth should be related to the “QoS profiles” feature of the switches, and the number of such available profiles is very small for supporting the requirements of the BoD service. Support for BoD in a CGE domain therefore has to be combined with over-provisioning / dimensioning of the service in terms of capacity.

9 Conclusions

9.1 Technologies Tested

This test report concludes the work carried out by JRA1 Task 1 during Y1–Y3 of the GN3 project. The work was divided into two parts: a theoretical part, where the Task identified transport technologies that qualify as Carrier Class Transport Network Technologies and that are considered to be relevant to the NREN community, and a second part where the selected technologies were exhaustively tested and assessed. The selected technologies were EoMPLS, Ethernet OAM, SE, PBB-TE, MPLS-TP, and OTN and GMPLS.

The main purpose of this work was to evaluate the different technologies, to test their capabilities and finally to assess how these capabilities can benefit NREN networks. The documented outcome of the investigation should help with the design and implementation of next-generation NREN infrastructure.

In order to have access to the newest hardware and features, the Task secured the involvement of a number of vendors. While in some cases, such as for the Ethernet OAM testing, it was possible to obtain hardware and software from the vendor, the tests were mostly performed at vendor premises, sometimes in the form of demonstrations. Nevertheless, the Task considers the results to be very useful and relevant, as it is hoped that this report has demonstrated. The main conclusions from the tests are presented below.

9.2 Test Results

The test results showed that both EoMPLS and PBB-TE are quite mature carrier-grade transport technologies, but while EoMPLS is designed for multi-domain environments, PBB-TE is more suitable for single-domain environments. EoMPLS technologies and services (i.e. VPWS and VPLS) fulfil the selected requirements for Carrier Class Transport Network Technologies identified in the Introduction (Section 1.1, page 8) and offer mechanisms for OAM, multi-domain topologies, protection and restoration, and efficient multicast distribution. In cases where the tested EoMPLS and MPLS techniques did not fully satisfy the requirements, alternative mechanisms were proposed to achieve the desired functionality. EoMPLS offers manageable and scalable transmission services with robust protection and restoration mechanisms. The technology can be used in research networks as well as carriers' networks to support data transmission services for the research community and other users.

As already mentioned, PBB-TE as a transport technology is better suited to single-domain applications. Multi-domain use of contiguous PBB-TE tunnels is possible but it does not comply with the MEF specification of the External Network-to-Network Interface and needs mutual knowledge of the MAC addresses of tunnel termination points, which goes beyond the normal method of management operations between domains.

Participants of the trial tend to agree with the current mainstream opinion that EoMPLS should be used in the core networks while PBB-TE could be used in access networks [HeavyReading]. PBB-TE could also be used in large campus networks, exploiting PBB-TE's traffic engineering features and fast protection switching whilst being applicable to the predominantly Ethernet-based expertise of many campus network managers.

A possible alternative to EoMPLS and PBB-TE is MPLS-TP. The demonstrations showed that MPLS-TP is still under development and the most complete and stable implementations are still to come. In the beginning, MPLS-TP was introduced as a revolutionary new technology but, in reality, it is a subset of MPLS with some added functionality for OAM and survivability operations. The results showed that this added functionality would be very beneficial for NREN networks. MPLS-TP is suited for NRENs and organisations with a strong transport culture in their operations, moving from legacy transport networks like SDH towards MPLS-based transport networks.

These three technologies are suitable for transporting Ethernet traffic, which makes Ethernet OAM features relevant and necessary. The main finding from the Ethernet OAM trial is that the functions embedded in the carrier-grade Ethernet equipment can be used for effective monitoring of the health and performance of wide-area Ethernet services from customer and provider perspectives. These functions are standardised by a number of IEEE, ITU-T and MEF specifications, and vendor implementations are close enough to those specifications to provide interoperability between OAM agents. The CyPortal software from Cyan, Inc., which stored and visualised monitoring data in diverse forms, was an important and useful element of the trial. It showed that this type of software is a crucial element in a provider's OSS environment when the provider offers managed and monitored Ethernet services to their customers. CyPortal will be used by the Task during Y4 for Ethernet OAM monitoring.

Another transport technology that was evaluated and that is slightly different from the three technologies discussed above is OTN. OTN is considered to be a promising technology to take over the established SDH networking layer and boost the network with more powerful switching, mapping and survivability functionality in the digital domain. In addition, it brings seamless integration to the optical domain and provides a common vehicle for mapping, switching and transporting all types of client signals. The testing of existing OTN platforms shows that the products are reaching market maturity. The most important functionalities, such as switching on different ODU levels and survivability based on different SNC parameters, including TCM, are already available. ODUflex is not fully implemented yet, but the tests confirmed that the basic functionalities to make ODUFlex possible are working well. The integration of a control plane into the OTN technology adds important functionality and intelligence and opens up possibilities for dynamic provisioning tools integration, which is a major requirement for transport technologies in NRENs' transport networks.

Finally, the test results for Synchronous Ethernet showed very positive results (as the objective of the test was achieved) but the possible uses for this technology in NREN networks are still uncertain. SE and PTP were configured and tested in both the CESNET lab and Cisco remote labs. It seems that R&E networks and the R&E community may be focused more on PTP deployment because PTP can provide both frequency and time distribution and no additional hardware is required (though the software must be PTP capable). SE, on the other hand, can distribute accurate frequency only and requires new hardware since standard Ethernet ports cannot work with SE ports and provide the required functionality. Future requests from the R&E community are needed to confirm this expectation.

9.3 Recommendations

The tests carried out proved that these technologies and their capabilities are able to meet the NRENs' requirements for Carrier Class Network Technologies capable of delivering reliable and flexible transport services. (The exception to this is SE, which requires further investigation.) It is up to the individual NRENs to decide – with this report and DJ1.1.1 to help them – how to integrate these technologies in their infrastructures and how to operate them. There seems to be a general tendency in the NREN environment towards EoMPLS which, with the new added features, is a perfect match for delivering carrier-class services. However, for some applications, OTN might be the right solution for achieving seamless integration between the optical and digital layer. Since next-generation core networks are going to be built with links of 100 Gbit/s and beyond, there is an obvious need for grooming and multiplexing of 1 GE and 10 GE links. The integration of features like Ethernet OAM and OTN TCM would require carefully thought out planning and design, and close collaboration between NRENs to achieve the smooth operation of their networks. OPEX reduction and more stable and reliable services are the main benefits claimed by the equipment providers. This is achieved by better end-to-end overview of the services in a multi-domain environment, possibility for precise allocation of the problem in the network and it can be used to trigger restoration mechanisms. However, these claims need to be validated by real experiences in real networks.

PBB-TE is a technology that, despite rumours to the contrary, is still in use. As already mentioned, PBB-TE is a good choice for single domains and campus networks, and several NRENs in Europe (such as SURFnet) have PBB-TE implemented in their networks. The future of MPLS-TP is still uncertain and it will take some time before the penetration of this technology in the market can be evaluated. MPLS-TP is a good choice for those organisations that have a strong transport tradition and culture in their operations. Finally, GMPLS will benefit the NRENs, adding intelligence to their networks and allowing survivability functions (automated restoration) as well as integration with BoD applications.

This report gives a good overview of the different technologies' status and identifies that there is still work to be done with regard to the integration of these technologies in multi-domain environments, including defining common operational procedures and best practices. In some cases the technology is still in development phase, which means that investigation needs to be continued in Y4 and future GÉANT projects.

During the GN3 project JRA1 Task 1 has collaborated with other Activities with similar interests, and as part of this collaboration JRA1 Task 1 participated in AutoBAHN testing with JRA2 Task 2, providing the test infrastructure. The testbed for AutoBAHN was provided by Essex University and JRA1 Task 1 participants from Essex University provided support during the process. The testing represents a technology implementation trial for a specific use case rather than for a more generic CCTNT implementation, as in the other tests, and is an example of the cross-Activity work that plays a key role in the Task 1's objectives and success criteria.

9.4 Future Work

Since there is still work to be done in the area of Carrier Class Transport Network Technologies, JRA1 Task 1 has been granted a one-year extension and in Y4 of GN3 the Task will focus on the following subjects:

- MPLS-TP – further testing.
- EoMPLS – further testing.
- Ethernet OAM – further testing and Service Assurance testing.
- Cross-Activity work with JRA2 Task 3 for perfSONAR extensions to support Ethernet OAM.
- Time-sensitive data applications study.
- OpenFlow.

Note: The last two items will be treated as sub-tasks within JRA1 Task 1, as their scope is slightly different from the rest.

Further information about all the above can be found in "Year 4 Activity Plans Overview" [GN3-Y4-Plans].

The Task will continue to disseminate the results of its work. To date, two white papers have been published [WhitePaper_CE, WhitePaper_OTN] and Task participants have presented at the 2010 TERENA Network Conference and 2011 NORDUnet Conference. JRA1 Task 1 work will conclude at the end of GN3 Y4.

Appendix A Trial Participants

The table below shows the name and NREN/institution of the participants in each trial, and the trial leader.

Name	NREN/Institution	EoMPLS	Ethernet OAM	SE	PBB-TE	MPLS-TP	OTN	X-Activity
Alberto Colmenero	NORDUnet		✓			✓[1]	✓[1]	
Anna Manolova Fagertun	NORDUnet/DTU						✓	
Bijan Rahimzadeh Rofoee	Essex University				✓			✓
Dave Tinkler	JANET		✓		✓			
Eduard Escalona	Essex University				✓			
Faris Ali	Lancaster University				✓			
Jac Kloots	SURFnet		✓					
Jan Radil	CESNET	✓	✓	✓[1]				
Jo Hoffmann	Lancaster University				✓			
Kurosh Bozorgebrahimi	NORDUnet/ UNINETT						✓	
Marcin Garstka	PSNC	✓[1]	✓					
Martin Dunmore	JANET		✓		✓	✓		
Mayur Channegowda	Essex University		✓					✓[1]
Paul Boyd	Lancaster University				✓			
Piotr Turowicz	PSNC						✓	
Ramanujam Jayakumar	Essex University		✓					✓
Rasmus Lund	NORDUnet						✓	
Reza Nejabati	Essex University				✓			
Victor Olifer	JANET		✓[1]		✓[1]	✓		

Table A.1: Trial participants

Notes:

1. Trial leader

Appendix B TP Framework Configuration

B.1 Introduction

This appendix presents the technology proxy framework configuration referred to in Chapter 8 *Cross-Activity Work*.

The following XML is used in order to correctly configure the TP Framework module, a generic AutoBAHN module for providing access to the heterogeneous underlying network infrastructure. The XML configuration is divided into four sections:

- **Credentials.** In this section the user is able to provide a set of credentials used for authentication when communicating with the devices.
- **Protocols.** Used for customisation of the protocols (e.g. telnet, ssh) used for the connections. Each protocol can be configured as many times as needed and then assigned to the appropriate device connection.
- **Devices.** This section defines a set of network devices handled by TP Framework. Only specified devices can be operated on.
- **Loaders.** This section relates to a collection of loaders responsible for applying configuration on particular devices. Whenever TP intends to apply any configuration on the device, it searches for an appropriate loader according to the information provided in the request. Loaders might be specialised according to vendor, model, OS name or version. Each TP loader defines methods for adding and removing configuration to set up or tear down the path on the given device.

The processing takes place as follows:

1. For each node from the path given in the request, TP searches for a suitable loader. A single loader can be reused for different nodes/devices but when it is executed it affects only a single node/device.
2. When the loader starts processing it selects the appropriate method to call.
3. The called method triggers a sequence of operations, e.g. to start connection to device, execute commands, check results, etc.

B.2 Configuration Code

```
<?xml version="1.0" encoding="utf-8" ?>
<tool xmlns:sh="http://tool.autobahn.geant.net/configuration/shell">

      <credentials>
         <grnet id="grnet">
               <username>XXXX</username>
               <password>XXXX</password>
         </grnet>
      </credentials>

      <protocols>
         <telnet type="telnet" defaultPort="23">
               <loginPrompt>UserName:</loginPrompt>
               <passwordPrompt>PassWord:</passwordPrompt>
               <prompt>></prompt>
               <superPrompt>#</superPrompt>
         </telnet>

      </protocols>

      <devices>
         <device address="192.168.8.251">
               <loopback>10.0.43.1</loopback>
               <connections>
                     <telnet credentials="grnet" protocol="telnet" />
               </connections>
         </device>
         <device address="192.168.8.252">
               <loopback>10.0.44.1</loopback>
               <connections>
                     <telnet credentials="grnet" protocol="telnet" />
               </connections>
         </device>
         <device address="192.168.8.253">
               <loopback>10.0.42.1</loopback>
               <connections>
                     <telnet credentials="grnet" protocol="telnet" />
               </connections>
         </device>
      </devices>

      <loaders>
```

```
   <sh:loader protocols="telnet">
         <sh:method id="add">
               <textParser/>
               <action>
                     <script>

   #if($constraintsIn.getConstraintForName('VLANS').value !=
$constraintsOut.getConstraintForName('VLANS').value)
                                 $this.failure('system')
                           #end
                     </script>
               </action>
               <if condition="$nodeType == 'internal'"
                     <shell>
                           <command id="create_vman">
                                 <execute>create vman
$parser.limit($resId, 32)</execute>
                           </command>
                           <command id="tag_vman">
                                 <execute>configure vman
$parser.limit($resId, 32) tag
$constraintsIn.getConstraintForName('VLANS').value</execute>
                           </command>
                           <command id="add_vman_ports">
                                 <execute>configure vman
$parser.limit($resId, 32) add port $ifceIn.name, $ifceOut.name
tagged</execute>
                           </command>
                     </shell>
               </if>
               <if condition="$nodeType == 'ingress' || $nodeType ==
'egress'">
                     <textParser/>
                     <shell>
                           <command id="create_svlan">
                                 <execute>create svlan
$parser.limit($resId, 32)</execute>
                           </command>
                           <command id="tag_svlan">
                                 <execute>configure svlan
$parser.limit($resId, 32) tag
$constraintsIn.getConstraintForName('VLANS').value</execute>
                           </command>
                           <command id="add_svlan_ports">
```

```
                                       <execute>configure svlan
      $parser.limit($resId, 32) add port $ifceIn.name, $ifceOut.name
      tagged</execute>
                                 </command>
                                 <!-- <command id="isid">
                                       <execute>configure svlan
      $parser.limit($resId, 32) isid
                                 </command> -->
                                 <command id="create_bvlan">
                                       <execute>create bvlan b</execute>
                                 </command>
                                 <command id="tag_bvlan">
                                       <execute>configure bvlan b tag
      1</execute>
                                 </command>
                                 <command id="add_bvlan_ports">
                                       <execute>configure bvlan b add ports
      **** tagged</execute>
                                 </command>
                                 <command id="add_svlan">
                                       <execute>configure bvlan b add svlan
      $parser.limit($resId, 32)</execute>
                                 </command>
                           </shell>
                     </if>
               <sh:method id="bandwith">
                     <textparser/>
                     <shell>
                           <command id="DSCP">
                                 <execute>enable diffserv examination ports
      all</execute>
                           </command>
                           <command id="QoS_profile_configuration">
                                 <execute>configure qosprofile all committed-
      rate $resParams.capacity ports all</execute>
                           </command>
                     </shell>
               </sh:method>
         </sh:loader>
      </loaders>

</tool>
```

References

[Accedian] http://www.accedian.com/

[AutoBAHN] https://forge.geant.net/forge/display/autobahn/Home

[CiscoMWR2941] http://www.cisco.com/en/US/products/ps9440/index.html

[Configs] http://www.ja.net/development/optical-networking/carrier-ethernet/configs/

[CyanInc] http://cyaninc.com/

[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

[DJ2.2.1] M. Balcerkiewicz, E. Escalona, V. Kapoulas, G. Kramer, R. Krzywania, R. Lund, J. Lukasik, A. Mackarel, B. Bach Mortensen, V. Papapanagiotou, V. Reijs, A. Sevasti, K. Stamos, “Deliverable DJ2.2.1: Specification of enhancements and developments for the AutoBAHN system” http://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-09-040-DJ2-2-1_Specification_of_Enhancements_and_Developments_for_the_AutoBAHN_System.pdf

[Extreme] Extreme Networks http://www.extremenetworks.com

[ExtremeXOS_CmdRef] Extreme Networks, Inc., ExtremeXOS 12.5.3 Command Reference http://www.extremenetworks.com/products/extreme-xos.aspx [access to document restricted to customers]

[ExtremeXOS_Concepts] Extreme Networks, Inc., ExtremeXOS 12.5.3 Concepts Guide http://www.extremenetworks.com/products/extreme-xos.aspx [access to document restricted to customers]

[GN3-Y4-Plans] “Year 4 Activity Plans Overview” https://intranet.geant.net/sites/Management/NREN_PC/Documents/GN3-11-356_20120314-NRENPC-Munich_Y4_AL_Activity_Overview_v1.7.pdf [restricted access]

[HeavyReading] http://www.heavyreading.com/insider/details.asp?sku_id=1880&skuitem_itemid=1047&&promo_code=&aff_code=&next_url=%2Finsider%2Fdefault.asp%3F

[IEEE 802.1ab] “802.1ab-2005 – IEEE Standard for Local and Metropolitan Area Networks: Station and Media Access Control Connectivity Discovery”, 6 May 2005 http://www.ieee802.org/1/pages/802.1ab.html

[IEEE 802.1ad] “802.1ad-2005 – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 4: Provider Bridges” http://standards.ieee.org/findstds/standard/802.1ad-2005.html

[IEEE 802.1ag] “802.1ag-2007 – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 5: Connectivity Fault Management”
http://standards.ieee.org/findstds/standard/802.1ag-2007.html

[IEEE802.1ah] “802.1ah-2008 – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 7: Provider Backbone Bridges”
http://standards.ieee.org/findstds/standard/802.1ah-2008.html

[IEEE 802.1aq] “802.1aq – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 8: Shortest Path Bridging”
http://www.ieee802.org/1/pages/802.1aq.html

[IEEE 802.1Q] “802.1Q-2005 – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks”
http://www.dcs.gla.ac.uk/~lewis/teaching/802.1Q-2005.pdf

[IEEE 802.1Qay] “802.1Qay-2009 – IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 10: Provider Backbone Bridge Traffic Engineering”
http://standards.ieee.org/findstds/standard/802.1Qay-2009.html

[Iperf] http://sourceforge.net/projects/iperf/

[ITU-T G.709/Y.1331] Recommendation ITU-T G.709/Y.1331 (12/2009) “Interfaces for the optical transport network (OTN)”
http://www.itu.int/rec/T-REC-G.709-200912-I

[ITU-T G.798.1] ITU-T Recommendation G.798.1 (04/2011) “Types and Characteristics of Optical Transport Network (OTN) equipment”
[restricted access]

[ITU-T G.873.1] Recommendation ITU-T G873-1 (07/2011) “Optical Transport Network (OTN) – Linear Protection”
[restricted access]

[ITU-T G.873.2] Recommendation ITU-T G873-2 “Optical Transport Network (OTN) – Ring Protection”
[not yet published]

[ITU-T G.808.1] Recommendation ITU-T G.808.1 (02/2010) “Generic protection switching – Linear trail and subnetwork protection”
http://www.itu.int/rec/T-REC-G.808.1-201002-I

[ITU-T G.8011] ITU-T Recommendation G.8011 / Y.1307 “Ethernet Services Framework”, 2004
http://www.itu.int/itudoc/itu-t/aap/sg15aap/history/g8011/g8011.html

[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)” – not yet approved
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” – not yet approved
http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=8148

[ITU-T G.Sup43] Supplement G.Sup43 “Transport of IEEE 10G Base-R in Optical Transport Networks (OTN)”
[restricted access]

[ITU-T Newslog1] Carrier network standards approved at Geneva meeting. 17 December 2011.
http://www.itu.int/ITU-T/newslog/Carrier+Network+Standards+Approved+At+Geneva+Meeting.aspx

[ITU-T Newslog2] Experts cast doubt on “jeopardise” Internet statement
http://www.itu.int/ITU-T/newslog/Experts+Cast+Doubt+On+Jeopardize+Internet+Statement.aspx

[ITU-T Study Group 15: MPLS] http://www.itu.int/oth/T0A0B00000C

[ITU-T Y.1563] Recommendation ITU-T Y.1563 (01/2009) “Ethernet frame transfer and availability performance”
http://www.itu.int/rec/T-REC-Y.1563-200901-I

[ITU-T Y.1564] Recommendation ITU-T Y.1564 (03/2011) “Ethernet service activation test methodology”
http://www.itu.int/rec/T-REC-Y.1564/en

[ITU-T Y.1731] Recommendation ITU-T G.8013/Y.1731 (07/2011) “OAM functions and mechanisms for Ethernet based networks”
http://www.itu.int/rec/T-REC-Y.1731/en

[JANET CE] http://www.ja.net/development/optical-networking/carrierethernetproject.html

[MbergLANTIMEM600] http://www.meinberg.de/english/products/lantime-m600-gps.htm

[MEF 6.1] MEF Technical Specification 6.1 “Ethernet Services Definitions – Phase 2”, April 2008
http://metroethernetforum.org/PDF_Documents/technical-specifications/MEF6-1.pdf

[MEF 10.2.1] MEF Technical Specification 10.2.1 “Performance Attributes Amendment to MEF 10.2”, 25 January 2011
http://metroethernetforum.org/PDF_Documents/technical-specifications/MEF_10.2.1.pdf

[MEF 10.2] MEF Technical Specification 10.2 “Ethernet Services Attributes Phase 2”, 27 October 2009
http://metroethernetforum.org/PDF_Documents/technical-specifications/MEF10.2.pdf

[MEF 26.1] MEF Technical Specification 26.1 “External Network to Network Interface (ENNI) – Phase 2”, January 2012
http://metroethernetforum.org/PDF_Documents/technical-specifications/MEF_26.1.pdf

[MEF-CGE] MEF, “Carrier Ethernet Services Overview”, August 2008
http://www.google.co.uk/url?sa=t&rct=j&q=a%20ubiquitous%2C%20standardised%2C%20carrier-class%20service%20and%20network%20defined%20by%20five%20attributes%20that%20distinguish%20it%20from%20familiar%20lan%20based%20ethernet&source=web&cd=2&ved=0CCsQFjAB&url=http%3A%2F%2Fmetroethernetforum.org%2FPPT_Documents%2FCarrier_Ethernet_Services_Overview.ppt&ei=lihvT5SdO8q_0QX4i7iOAg&usg=AFQjCNFEj5ySM5r8Z50jOrMPc3hMYG_8Cw

[MPLS-TP facts] http://www.itu.int/ITU-T/newslog/MPLSTP+The+Facts.aspx

[RFC4379] K. Kompella, G. Swallow, “Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures”, RFC 4379, February 2006
http://tools.ietf.org/html/rfc4379

[RFC 2205] R. Braden (ed.), “Resource ReSerVation Protocol (RSVP) – Version 1 Functional specification”, RFC 2205, September 1997
http://www.ietf.org/rfc/rfc2205.txt

[RFC 3473] L. Berger (ed.), “Generalized Multi-Protocol Label Switching (GMPLS) Signaling - Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions”, RFC 3473, January 2003
http://www.faqs.org/rfcs/rfc3473.html

[RFC 4665] W. Augustyn (ed.), Y. Serbest (ed.), “Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks”, RFC 4665, September 2006
http://tools.ietf.org/html/rfc4665

[RFC 5654] B. Niven-Jenkins (ed.), D. Brungard (ed.), M. Betts (ed.), N. Sprecher, S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009
http://www.rfc-editor.org/rfc/rfc5654.txt

[RUDE] http://rude.sourceforge.net/

[TechAnnex] GN3 Project "Annex I – Description of Work"
https://intranet.geant.net/sites/Management/PMO/Documents/GN3-09-004v2%204%20GN3%20Technical%20Annex%20-%20DoW%20-%20Annex%20I.pdf

[TPFramework] M. Giertych, J. Łukasik, "Technology Proxy Framework for AutoBAHN: User Guide"
https://forge.geant.net/forge/pages/viewpage.action?pageId=2556304

[WhitePaper_CE] J. Kloots, V. Olifer, "White Paper: Carrier Ethernet"
http://www.geant.net/Media_Centre/Media_Library/Media Library/Carrier Ethernet.pdf

[WhitePaper_OTN] A. Colmenero, R. Lund, A. Vasileva Manolova, "White Paper: OTN Capabilities in the NREN Environment"
http://www.geant.net/Media_Centre/Media_Library/Media Library/GN3-11-211_OTN-White-Paper_JRA1T1_v5.1.pdf

[Wireshark] http://www.wireshark.org/

Glossary

AIS	Alarm Indication Signal
ALU	Alcatel-Lucent
AMP	Asynchronous Mapping Procedure
APS	Automatic Protection Switching
AS	Autonomous System (an administrative domain in the BGP terminology)
ASBR	Autonomous System Border Router
BD	Extreme Networks BlackDiamond®
BDI	Backward Defect Indication
BEI	Backward Error Indication
BFD	Bi-directional Forwarding Detection
BGP	Border Gateway Protocol
BLAN	Bridged Local Area Network
BMP	Bit-Synchronous Mapping Procedure
BoD	Bandwidth on Demand
BTS	Base Transceiver Station
B-VID	Backbone Virtual LAN IDdentifier
B-VLAN	Backbone VLAN
B-VLAN	Bridged VLAN
CAPEX	Capital Expenditure
CBR	Constant Bit Rate
CCM	Continuity Check Message
CCP	Continuity Check Protocol
CCTNT	Carrier Class Transport Network Technology
CE	Customer Edge (device)
CET	Carrier Ethernet Transport
CFM	Connectivity Fault Management
CGE	Carrier Grade Ethernet
CIR	Committed Information Rate
CLI	Command Line Interface
C-PoP	Core Point of Presence
CV	Connectivity Verification
dDEG	Degraded Signal Defect
DM	Delay Measurement
DM	Domain Manager
DMM	Delay Measurement Message

DMp	Delay Measurement of ODUk path
DMR	Delay Measurement Response
DWDM	Dense Wavelength-Division Multiplexing [or Multiplexed]
eBGP	external Border Gateway Protocol (between Autonomous Systems)
EIR	Excess Information Rate
ENNI	External Network-to-Network Interface
EoMPLS	Ethernet over MPLS
EoT	Ethernet over Transport
EPL	Ethernet Private Line
EP-LAN	Ethernet Private LAN
EP-Tree	Ethernet Private Tree
ERS	Ethernet Relay Service
ESM	Ethernet Service Manager
ESMC	Ethernet Synchronisation Messaging Channel
ETH-AIS	Ethernet Alarm Indication Signal
EVC	Ethernet Virtual Connection
EVPL	Ethernet Virtual Private Line
EVP-LAN	Ethernet Virtual Private LAN
EVP-Tree	Ethernet Virtual Private Tree
EWS	Ethernet Wire Service
FC	Fibre Channel
FDB	Forwarding Database
FIB	Forwarding Information Base
FPGA	Field-Programmable Gate Array
FSP	Fibre Service Platform
FTFL	Fault Type and Fault Location
FTP	File Transfer Protocol
FWR	Fast Wavelength Restoration
G	Gigabit
Gbit/s	Gigabit per second
GCC	General Communication Channel
GE	Gigabit Ethernet
GFP	Generic Framing Procedure
GFP-F	Generic Framing Procedure – Transparent
GMP	Generic Mapping Procedure
GMPLS	Generalised Multi-Protocol Label Switching
GPS	Global Positioning System
GUI	Graphical User Interface
HD	High Definition
HO	High Order
HSLM	Hybrid Service Line Module
IB DDR	Infini Band Double Data Rate
IB SDR	Infini Band Single Data Rate
IEEE	Institute of Electrical and Electronics Engineers
IETF PWE3	Internet Engineering Task Force Pseudowire Emulation Edge to Edge working group

IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IL	Inner Label
IP	Internet Protocol
I-SID	Information Service Identifier
IS-IS	Intermediate System to Intermediate System
ITU-T	International Telecommunication Union – Telecommunication Standardisation Sector
JANET CE	JANET Carrier Ethernet Project
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 2	JRA2 Task 2, Hybrid Network Provisioning
JRA2 Task 3	JRA2 Task 3, Monitoring
L2VPN	Layer 2 Virtual Private Network
LAN	Local Area Network
LDP	Label Distribution Protocol
LEOS	LightningEdge Operating System
LER	Label Edge Router
LFIB	Label Forwarding Information Base
LLDP	Link Layer Discovery Protocol
LM	Loopback Measurement
LM	Loss Measurement
LO	Low Order
LSA	Link State Advertisement
LSC	Lambda Switch Capable
LSP	Label Switched Path
LSR	Label Switch Router
MA	Maintenance Association
MAC	Media Access Control
MD	Maintenance Domain
MEF	Metro Ethernet Forum
MEG	Maintenance Entity Group
MEP	MEG End Point
MERS	Metro Ethernet Routing Switch
MIP	MEG Intermediate Point
MPLS	Multi-Protocol Label Switching
MPLS-TP	MPLS – Transport Profile
MSTP	Multiple Spanning Tree Protocol
MTIE	Maximum Time Interval Error
MTU	Maximum Transmission Unit
MWR	Mobile Wireless Router
NDA	Non-Disclosure agreement
NGRTN	Next-Generation Routed Transport Network
NMS	Network Management System
NREN	National Research and Education Network

NSN	Nokia Siemens Networks
NTP	Network Time Protocol
OAD	One-way Average Delay
OADV	One-way Average Delay Variation (jitter)
OAM	Operation, Administration and Maintenance
OCh	Optical Channel
ODU	Optical Data Unit
OEM	Original Equipment Manufacturer
OL	Outer Label
OMS	Optical Multiplex Section
OPEX	Operational Expenditure
OPU	Optical Payload Unit
OSPF	Open Shortest Path First
OSPF-TE	Open Shortest Path First with Traffic Engineering
OSS	Operations Support System
OTN	Optical Transport Network
OTS	Optical Transport Section
OTU	Optical Transport Unit
P	Core router
PAA	Performance Assurance Agent
PB	Provider Bridges
PBB	Provider Backbone Bridge
PBBN	Provider Backbone Bridge Network
PBB-TE	Provider Backbone Bridge Traffic Engineering
PBT	Provider Backbone Transport
PCC	Protection Communication Channel
PDU	Protocol Data Unit
PE	Provider Edge (device)
PIR	Peak Information Rate
PL	Packet Loss
PM	Path Monitoring
PNC	Provider Network Control
P-NNI	Private Network-to-Network Interface
PSN	Packet-Switched Network
PT	Payload Type
PTP	Precision Time Protocol
PW	Pseudowire
QoS	Quality of Service
Ra	Measured rate
RAN	Radio Access Network
RDI	Remote Defect Indicator
Rg	Value of traffic injected into the Ethernet connection
RN	Regional Network
RNC	Radio Network Controller
ROADM	Reconfigurable Optical Add-Drop Multiplexer

RSS	Reconfigurable Switching System
RSVP	Resource Reservation Protocol
RSVP-TE	Resource Reservation Protocol with Traffic Engineering
RTP	Real-Time Protocol
RUDE	Real-Time UDP Data Emitter
SA	GN3 Service Activity
SA1	GN3 Service Activity 1, Network Build and Operations
SA2	GN3 Service Activity 2, Multi-Domain Network Services
SA2 Task 5	SA2 Task 5, Tools to Support Multi-Domain Workflows
SCN	Signalling Communication Network
SD	Signal Degradation
SD	Signal Degrade
SDH	Synchronous Digital Hierarchy
SE	Synchronous Ethernet
SF	Signal Failure
SLD	Service Level Description
SNC/I	Sub-Network Connection protection with Inherent monitoring
SNC/N	Sub-Network Connection protection with Non-intrusive monitoring
SNC/S	Sub-Network Connection protection with Sub-layer monitoring
SNCP	Sub-Network Connection Protection
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical Network
SP	Service Provider
SPB	Shortest Path Bridging
SSF	Server Signal Failure
SSH	Secure Shell
SSM	Synchronisation Status Message
STM	Synchronous Transport Module
STP	Spanning Tree Protocol
S-VID	Service Virtual LAN Identifier
S-VLAN	Service VLAN
TACACS	Terminal Access Controller Access-Control System
TAD	Two-way Average Delay
TADV	Two-way Average Delay Variation
TCA	Traffic Conformance Agreement
TCM	Tandem Connection Monitoring
TCP	Transmission Control Protocol
TCTE	Tandem Connection Terminating Element
TDM	Time Division Multiplexing
TE	Traffic Engineering
TIM	Trail Trace Identifier Mismatch
TLV	Type Length Value
TP	Technology Proxy
TRBD	Transponder Board
TS	Time Slot

TSS	Transport Service Switch
TTI	Trail Trace Identifier
TTT	Timing Transparent Transcoding
TVN	Thames Valley Network
UDP	User Datagram Protocol
UHD	Ultra High Definition
UNI	User Network Interface
UPSR	Uni-directional Path Switched Ring
VB	Virtual Bridge
VID	Virtual LAN Identifier
VLAN	Virtual LAN
VM	Virtual Machine
VoIP	Voice over IP
VPLS	Virtual Private LAN Service
VPMS	Virtual Private Multicast Service
VPN	Virtual Private Network
VPWS	Virtual Private Wire Service
WAN	Wide Area Network
WTSA	World Telecommunication Standardisation Assembly
WWP	World Wide Packets
XML	Extensible Markup Language