UNIVERSITAT POLITÈCNICA DE CATALUNYA

Centre Tecnològic de Telecomunicacions de Catalunya

Ph.D. Thesis

GMPLS-based provisioning of Ethernet connections over WSON with Quality of Service

Author:	Anica Bukva
Advisors:	Dr. Ramon Casellas
	Dr. Ricardo Martínez
Tutor:	Dr. Gabriel Junyent

Department of Signal Theory and Communications
Universitat Politècnica de Catalunya

October 2012, Barcelona

Abstract

The rapid growth of the data traffic, caused by the ever-growing number of Internet users and the services requiring huge bandwidth (e.g., videoconferences, IPTV, telemedicine, etc.) show that the existing transport network infrastructure is not optimal in terms of cost, flexibility and scalability. This problem is motivating the research community and the operators to explore new network architectures to the traditional transport technologies, providing higher capacity and flexibility, simpler operations and lower deployment costs. One of the most efficient ways to cope with the ever-growing data traffic is to deploy Wavelength Switched Optical Networks (WSON), which offer a huge transport network capacity thanks to the DWDM (Dense Wavelength Division Multiplexing) technology. Moreover, the utilization of all-optical switching (e.g., Optical Cross-Connects – OXCs) eliminates the electrical processing at the higher layers, which require expensive optical-electrical-optical (OEO) conversions, reducing, thus, the overall network cost.

On the other hand, next generation networks are expected to efficiently support packet-based services such as Ethernet transport, while keeping the same carrier-grade characteristics (e.g., reliability, scalability, operation and maintenance, etc.) of the traditional transport networks. In that sense, an appealing solution relies on deploying Connection-Oriented (CO) Ethernet, because of its advantages such as simplicity, interoperability, cost-effectiveness and bandwidth flexibility. The deployment of the CO Ethernet to become a high performance transport technology is being addressed by two different solutions, namely, the IETF/IEEE Provider Backbone Bridges – Traffic Engineering (PBB-TE) and the ITU-T MPLS- Transport Profile (MPLS-TP).

In view of all the exposed above, a candidate transport architecture for next generation networks is constituted by a dual layer infrastructure combining the benefits provided by flexible CO-Ethernet and the huge transmission capacity provided by WSON networks.

A key element for the deployment of such an integrated (dual-layer) network scenario is the introduction and use of the intelligence and automatic functions and mechanisms (e.g., real-time provisioning, routing, recovery, etc.) provided by a distributed control plane. In this regard, the Generalized Multi-Protocol Label Switching (GMPLS) protocol suite is a viable solution, since it is designed to support and handle a variety of network technologies (e.g., Packet, Time Division Multiplexing, Layer 2 Ethernet, DWDM, Fiber, etc.). Hence, a single GMPLS control plane instance may be used to control multiple switching capabilities.

In this dissertation, we focus on the GMPLS control for the provisioning and protection within a multi-layer network architecture constituted by both, CO-Ethernet and WSON switching layers.

An important concept in such a multi-layer network (MLN) is the Forwarding Adjacency (FA). The aim of FAs is to exploit the grooming decisions where existent lower-layer Label Switched Paths (LSPs) (i.e., FA LSPs) with sufficient unreserved

bandwidth are intelligently reused when accommodating and serving new upper-layer LSPs. By doing so, the established lower-layer LSPs form a Virtual Network Topology (VNT) for the provisioning of CO-Ethernet services. This yields to improve the scalability and efficiency of the network resource utilization of GMPLS-controlled networks, since multiple higher-layer LSPs may be nested/groomed over a single FA-LSP.

One of the main challenges in a MLN is how to accommodate dynamically requested higher-layer connections in the created VNT to optimally utilize the large capacity offered by a WSON. In this regard, we focus on the problem of path computation and LSP provisioning in a CO-Ethernet over WSON network, under the framework of GMPLS unified control plane. Firstly, we propose the usage of a *FA TE link timer*, which postpones the release of created FA TE link when there are no more upper-layer connections over such a link, rather than releasing it immediately. By doing so, we do reduce the signaling overhead due to consecutive triggering of FA TE links between the same pair of nodes. As a second contribution, we compare three approaches for dynamic VNT reconfiguration: *dynamic*, *semi-dynamic* and *virtual*, using our proposed online path computation algorithm. We show that with the dynamic approach the lowest connection blocking probability is obtained but at the expense of the increased setup delay. Finally, we have proposed a dynamic path computation algorithm that chooses the route depending on the current state of the network resources, favoring the usage/re-use of virtual and active FA TE links over establishing new optical LSPs occupying unused wavelength channels. The proposed algorithm provides better usage of the network resources, keeping the connection blocking probability low, comparing with selected algorithms proposed in the literature.

Next, we focus on schemes and mechanisms for dedicated path protection in a CO-Ethernet over WSON network. In that sense, after classifying recovery mechanisms for MLN found in the literature, we compare both *link-* and *Shared Risk Link Group (SRLG)- disjoint schemes*. These schemes aim at computing a link and SRLG disjoint backup path with respect to the computed working paths, respectively. It is shown that, although the *SRLG-disjoint scheme* provides higher survivability comparing with the *link-disjoint scheme*, the connection blocking probability is still significantly higher. As a second contribution regarding dedicated path protection, we present a protection scheme which tries to accommodate a requested connection over the least congested TE links (i.e., higher unused bandwidth), while establishing both, working and backup paths. The aim of this contribution is to provide an efficient usage of the network resources and to minimize the number of connections affected by a link failure. Finally, we propose a Suurballe algorithm-based scheme that aims at decreasing the connection blocking probability due to working path SRLG-joint links that are removed during the backup path computation.

The exhaustive performance evaluations of the above contributions are conducted through simulations under the assumption of dynamic traffic pattern and using well-known reference network topologies.

Table of Contents

Acronyms

BSS/OSS	Business Support Systems / Operations Support Systems
CO	Connection-Oriented
CR-LDP	Constraint Routing-Label Distribution Protocol
CWDM	Coarse Wavelength Division Multiplexing
DWDM	Dense Wavelength Division Multiplexing
ERO	Explicit Route Object
FA	Forwarding Adjacency
FSC	Fiber Switch Capable
GMPLS	Generalized Multi-Protocol Label Switching
IACD	Interface Adjustment Capability Descriptor
IETF	Internet Engineering Task Force
IS-IS	Intermediate System to Intermediate System
ISC	Interface Switching Capability
ISCD	Interface Switching Capability Descriptor
IP	Internet Protocol
LAN	Local Area Network
L2SC	Layer-2 Switching Capable
LMP	Link Management Protocol
LSC	Lambda Switching Capable
LSA	Link State Advertisement
LSP	Label Switched Path
LSR	Label Switching Router
LTII	LSP Tunnel Interface ID
MLN	Multi-Layer Network
MRN	Multi-Region Network
MPLS	Multi-Protocol Label Switching
MPLS-TP	Multi-Protocol Label Switching – Transport Profile
MTTR	Mean Time To Recovery
NHLFE	Next Hop Label Forwarding Entry

NMS	Network Management System
OEO	Optical-Electrical-Optical
OSPF	Open Shortest Path First
OSPF-TE	Open Shortest Path First-Traffic Engineering
OXC	Optical Cross-Connect
PB	Provider Bridges
PBB	Provider Backbone Bridges
PBB-TE	Provider Backbone Bridges – Traffic Engineering
PCC	Path Computation Client
PCE	Path Computation Element
PSB	Path State Block
PSC	Packet Switching Capable
QoS	Quality of Service
ROADM	Reconfigurable Optical Add-Drop Multiplexer
RSVP	Resource ReSerVation Protocol
RSVP-TE	Resource ReSerVation Protocol-Traffic Engineering
SRLG	Shared Risk Link Group
STP	Spanning Tree Protocol
TED	Traffic Engineering Database
TDM	Time-Division Multiplexing
TE	Traffic Engineering
VNT	Virtual Network Topology
WC	Wavelength Converter
WCC	Wavelength continuity constraint
WDM	Wavelength Division Multiplexing
WSON	Wavelength Switched Optical Networks

List of Figures

Chapter 1

Introduction

1.1 Motivation

The capacity of today's transport networks is being exhausted due to the emerging services and applications, and the ever-growing number of Internet users. Therefore, the transport networks have to enhance their capacity to accommodate the continuing expansion of bandwidth demand that keeps on fuelling the data traffic growth. One of the solutions is to deploy Wavelength Switch Optical Networks (WSON) since it relies on the Dense Wavelength Division Multiplexing (DWDM) technology, where multiple optical channels are multiplexed within a single optical fiber, increasing, thus, the total transport capacity. Moreover, with the appearance of all-optical switching (e.g., optical cross-connects - OXCs and Reconfigurable Optical Add Drop Multiplexers - ROADM), the necessity of optical-electrical-optical (OEO) conversions when switching at intermediate nodes is eliminated, leading to a reduction of the network cost. On the other hand, next-generation optical transport networks are evolving to integrate the so-called carrier-grade technologies, such as the Connection-Oriented (CO) Ethernet, including Provider Backbone Bridge - Traffic Engineering (PBB-TE) [1] or Multi-Protocol Label Switching with Transport Profile (MPLS-TP) [2]. By doing so, transport network infrastructures leverage the benefits of CO packet transport such as the finer bandwidth granularity (i.e., flexibility), scalability and QoS, as well as the cost-efficient advantage of Ethernet or IP protocols. In light of the above, an interesting solution for future transport networks relies on deploying a multi-layer network (MLN) which integrates both switching layers: CO packet switching provided by a carrier-grade Ethernet and lambda switching provided by a WSON.

In such an MLN, when the dynamic provisioning of services is required, it is necessary to provide intelligent cooperation among the involved switching layers. In that sense, one of the most suitable candidates to exploit such MLN TE strategies is to adopt a Generalized Multi-Protocol Label Switching (GMPLS) unified control plane [3]. Specifically, in a unified control plane, the status of the resources and the TE attributes regarding each switching layer are consolidated into a single TE database (TED) repository. Consequently, the usage of the network resources in all the layers can be optimized when setting up a connection (a label switched path – LSP, in the GMPLS context), encompassing both switching technologies.

The question of how to efficiently and effectively control the processes for the dynamic provisioning of connections in a MLN scenario, and particularly in CO Ethernet over WSON, has lately gained increased interest. Searching for an optimal solution in such a network leads to facing up many open issues that need to be resolved. In that sense, the most important ones are: the correlation of the layer costs, the problem of the scalability and the performance issues that may arise when the number of TE links significantly grows, the development of new path computation algorithms considering aggregation and grooming decisions, the problem of choosing the switching layer at which a request will be accommodated, etc.

Moreover, another important problem in such a MLN is how to recover this transport network in case of a failure of one or more network elements (i.e., links and nodes). In general, the recovery approaches can be classified into protection and restoration. Deploying the former one, it is needed to compute both, a working and a backup path, once a connection request arrives. Deploying the latter one, a backup path is computed once a failure that affects the working path actually occurs. In this work, we concentrate on how to protect a MLN from failures, satisfying specific requirements and objectives in an efficient way. Therefore, the macroscopic objectives addressed in this thesis are two: on the one hand, devising path computation algorithms and LSP provisioning schemes and, on the other hand, end-to-end (e2e) dedicated protection (i.e., path protection) in CO Ethernet over WSON network, under the framework of GMPLS unified control plane. In the following section, we present the structure of the thesis, giving a brief overview of the contents of each Chapter.

1.2 Structure of thesis

The thesis is divided into five Chapters. At the end of the thesis, the design and implementation details of the simulator used for the experimental analysis of our work are detailed in the Appendix.

Chapter 1: Introduction

This chapter presents the motivations and rationales for deploying and investigating this topic, i.e. LSP provisioning and protection in GMPLS-controlled MLN infrastructures. Then, we detail the outline of this thesis dissertation. Finally, the research contributions and publications in terms of peer-reviewed journal, international and national conferences are enumerated.

Chapter 2: Introduction to GMPLS unified control plane for multi-layer networks

The second chapter is devoted to the most important features and mechanisms of a GMPLS unified control plane for LSP provisioning in a MLN, providing an overview of the technology and industrial standards in which the research work that we have carried out is scoped. In that sense, after presenting the two considered transport network technologies (i.e., CO-Ethernet and WSON) used in this work, the GMPLS concepts of Forwarding Adjacency (FA) and Virtual Network Topology (VNT) are introduced. Finally, the control signaling procedure for the automatic LSP establishment in a MLN is explained in detail.

Chapter 3: LSP Provisioning in MLN

This chapter deals with the problems of VNT configuration and path computation algorithms in a CO Ethernet over WSON, controlled by a GMPLS unified control plane. First, we overview the state of the art related to the path computation algorithms for VNT reconfiguration and LSP provisioning. Afterwards, we present the attained contributions related to this topic. Specifically, we can highlight three contributions:

The first contribution is presented in Section 3.5. It deals with the problem of the signaling overhead due to consecutive establishment of FA TE links between the same pair of nodes. To resolve this problem, we introduce the *FA TE link timer* that postpones the releasing of a FA TE link for some pre-defined time period, after the last L2SC LSP connection over such a link is torn down. Through numerical results, it is shown that the timer value is directly related to the generated signaling overhead due to the FA TE link creation.

The second contribution, presented in Section 3.6, deals with a simple path computation algorithm for a MLN that takes into account both, virtual and FA TE links. The algorithm is used for comparing three different approaches for dynamic VNT reconfiguration: *dynamic*, *semi-dynamic* and *virtual*. It is shown that the dynamic approach outperforms the other two in terms of connection blocking probability, but at the cost of higher setup delay due to the dynamicity of FA TE link creations.

Finally, as the third contribution, we have proposed, implemented and evaluated in the simulator, an on-line path computation algorithm for MLN controlled by GMPLS unified control plane. Such an algorithm relies on a novel path cost function which accounts different attributes and parameters such as the number of traversed hops, and the state of the network resources (i.e., links and nodes). We have shown that the proposed algorithm outperforms selected algorithms found in the literature, in terms of the connection blocking probability, targeting a more efficient balancing of the usage of the network resources at both involved layers. The proposed path computation algorithm and the numerical results obtained in the simulator are presented in Section 3.7.

Chapter 4: End-to-end dedicated protection in MLN

This chapter is dedicated to e2e recovery methods in a MLN network controlled by GMPLS unified control plane. Specifically, we address the problems regarding dedicated path protection, giving an overview of protection schemes in the current literature. Then, we present the contributions of this dissertation regarding this topic.

In Section 4.4, we compare the performance of two different approaches for e2e dedicated protection, regarding connection blocking probability and survivability. The first approach satisfies only link-disjointness, while the second approach takes into account SRLG disjoint links when computing a backup path (i.e., *link-disjoint scheme* and *SRLG-disjoint scheme*, respectively). It is shown that the *link-disjoint scheme* outperforms the *SRLG-disjoint scheme* in terms of connection blocking probability, but at the expense of worsening the network survivability.

In Section 4.5, we present a protection scheme which relies on computing a two-step Dijkstra-based algorithm. The algorithm aims at establishing both working and backup paths favoring less congested TE links. The goal of the scheme is to balance the network resources usage at both involved layers, providing high level of survivability, but keeping the blocking probability low. It is shown that the scheme obtains good performance, even with very high dynamicity of link failures.

Finally, the aim of the scheme presented in Section 4.6 is to decrease the connection blocking probability in a MLN due to temporal removing of working path SRLG-joint links from the network graph when computing the associated backup path. In order to relax this problem, we proposed a *SRLG-scaled scheme*, in which, when calculating the backup path, the SRLG-joint links are not totally removed from the network graph, but the SRLG-disjoint links are still prioritized. We have used a Suurballe-based path computation algorithm adjusted for a MLN. Through the set of simulations it is shown that, deploying the proposed *SRLG-scaled scheme*, we obtain much lower connection blocking probability, but still at the expense of worsening the survivability, when comparing with a MLN protection scheme found in literature which systematically eliminates all the links (SRLG-joint) forming the working path at the time of computing the associated backup path.

Chapter 5: Conclusions and Future work

In Chapter 5, we summarize the main contributions of this thesis dissertation, highlighting the main achievements and results obtained in this document. This task results essential to derive potential recommendations for the network operators when deploying the expected MLN network infrastructures in the mid-term. Finally, several topics and research lines for the future work are introduced.

Appendix

The simulator used in this thesis is implemented in the OPNET Modeler [4]. The implementation details are presented in the Appendix at the end of this document. After introducing the basic functionalities of the OPNET Modeling Domains, we explain the implementation of the model objects that we have designed and developed for the targeted research objectives. Finally, we explain in detail the implemented

GMPLS RSVP-TE signaling protocol functionalities [5] for the provisioning of CO Ethernet services over WSON networks.

1.3 Research contribution

Journals

- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "A Dynamic Path-Computation Algorithm for a GMPLS-Enabled Multi-layer Network", OSA Journal of Optical Communications and Networking (JOCN), Vol. 4, No. 6, pp. 436-448, June 2012.

International conferences

- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "An On-Line Path Computation Algorithm for a Protected GMPLS Enabled Multi-Layer Network", in Proceedings of 17th European Conference on Networks and Optical Communications (NOC 2012), 20-22 June 2012, Vilanova i la Geltrú (Spain).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "A dynamic path computation algorithm for e2e dedicated protection in a GMPLS controlled multilayer (Ethernet/WSON) network", in Proceedings of International Workshop on Reliable Networks Design and Modeling (RNDM 2011), 5-7 October 2011, Budapest (Hungary).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "A dynamic on-line path computation algorithm for VNT configuration in GMPLS controlled multi-layer (Ethernet/WSON) network", in Proceedings of IEEE Conference on High Performance Switching and Routing (HPSR 2011), July 4-6 2011, Cartagena (Spain).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Evaluation of Multi-Layer (Ethernet/WSON) GMPLS-Based Path Computation Policies for Dynamic VNT Configuration", in Proceedings of 16th European Conference on Networks and Optical Communications (NOC 2011), 20-22 July 2011, Newcastle upon Tyne (UK).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Enhanced Dynamic VNT configuration in GMPLS Controlled Ethernet over WSON with timer-based lightpath holding time", in Proceedings of 13th International Conference on Transparent Optical Networks (ICTON 2011), 26-30 June, 2011, Stockholm (Sweden).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Static versus Dynamic Virtual Network Topology Configuration within GMPLS-enabled

Connection-Oriented Ethernet over WSON", in Proceedings of the Bone-Tiger2 Summer School 2010, 6-10 September 2010, Budapest (Hungary).

- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "RSVP-TE signaling for dynamic provisioning in multilayer networks (Ethernet/WSON): Enhancements for Forwarding Adjacency operation", in Proceedings of the IX Workshop in G/MPLS Networks, 5-6 July 2010, Girona (Spain).
- R. Martínez, R. Casellas, R. Muñoz, A. Bukva, "Evolution of the GMPLS control plane of the ADRENALINE testbed : overview towards multi-layer (Ethernet/WSON) capabilities", in Proceedings of RT-Multilayer Workshop, June 14-16 2010, Vilanova i la Geltrú (Spain).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "OPNET implementation of RSVP-TE based signaling for GMPLS unified control of connection-oriented Ethernet transport over WSON", in Proceedings of RT-Multi-layer Workshop, December 2009, Vilanova i la Geltrú (Spain).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Unified GMPLS-based control of Carrier-grade Ethernet over DWDM", in Proceedings of 14th European Conference on Networks and Optical Communications (NOC 2009), 10-12 June 2009, Valladolid (Spain).

National conferences

- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Evaluation of a dynamic path computation algorithm in a GMPLS controlled multi-layer (Ethernet/WSON) network", in Proceedings of the Red FIERRO Workshop, 7-8 July 2011, Cartagena (Spain).
- A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Challenges and requirements of multi-layer GMPLS control of Ethernet over WDM technologies", in Proceedings of RT-Multi-layer Workshop, February 2009, Vilanova i la Geltrú (Spain)

Chapter 2

Introduction to GMPLS unified control plane for multi-layer networks

This chapter presents the basic concepts of GMPLS unified control plane for multi-layer networks. Firstly, we describe the main features of two transport network technologies used in this thesis: CO-Ethernet and WSON. Then, we explain the notion of a multi-layer network (MLN) and multi-region network (MRN) [6]. Finally, we present the main characteristics and functions of the GMPLS unified control plane.

2.1 Transport networks

With the appearance of new emerging IP packet-based services (e.g., high definition IPTV, video-conferences, etc.) and applications with various bandwidth and QoS requirements, it is necessary to improve the current transport networks so they can provide higher capacity, reliability and QoS in a cost-efficient manner. Indeed, new technology options are needed, improving the operational model that we have now. Therefore, one of the interesting solutions for future transport networks relies on deploying a MLN which integrates the following switching layers:

- **Packet switching:** We present two main solutions for connection-oriented Layer2 Packet Transport Technologies for Ethernet service delivery: the IEEE Provider Backbone Transport (PBT) [1] and the joint ITU-T and IETF solution referred to as MPLS Transport Profile (MPLS-TP) [2]. It should be born in mind that the latter one, however, can be applied to any Layer2 Packet Transport Network, not exclusively to Ethernet.
- **Circuit switching:** In this thesis, we focus on WSON transport networks, where the switching is performed based on wavelengths within a fiber. In such a network, a circuit-switched connection between the end nodes forms a lightpath that is used as a carrier for IP traffic.

In the following, we will explain these two technologies.

2.1.1 CO Ethernet

Ethernet is a dominant technology in the Local Area Network (LAN). However, as it was deployed in the beginning, Ethernet did not meet carrier requirements such as fault tolerance, OAM capabilities, scalability, QoS and protection. In consequence, major standard organizations were involved in the improvement of CO Layer2 Packet Transport Technologies for Ethernet service delivery so it could satisfy all the requirements. To this end, Ethernet migrated from a connectionless to a provisioned connection-oriented transport technology.

The deployment of the CO Ethernet to become a high performance transport technology is being addressed by different solutions. In this thesis, we present a PBT, a solution that directly extends the Ethernet, as well as MPLS-TP, a Layer2 Packet Transport Network technology, that can be used for provisioning of Ethernet connections.

Provider Backbone Transport

PBT is an extension to the IEEE Provider Backbone Bridges (PBB) [7], which is, in turn, an extension of the Provider Bridges (PB) [8]. PB, as well as PBB, is a connectionless carrier grade Ethernet technology. Figure 1 shows how, through the encapsulation, the original Ethernet frame (i.e., 802.1) has evolved to the PBB frame.

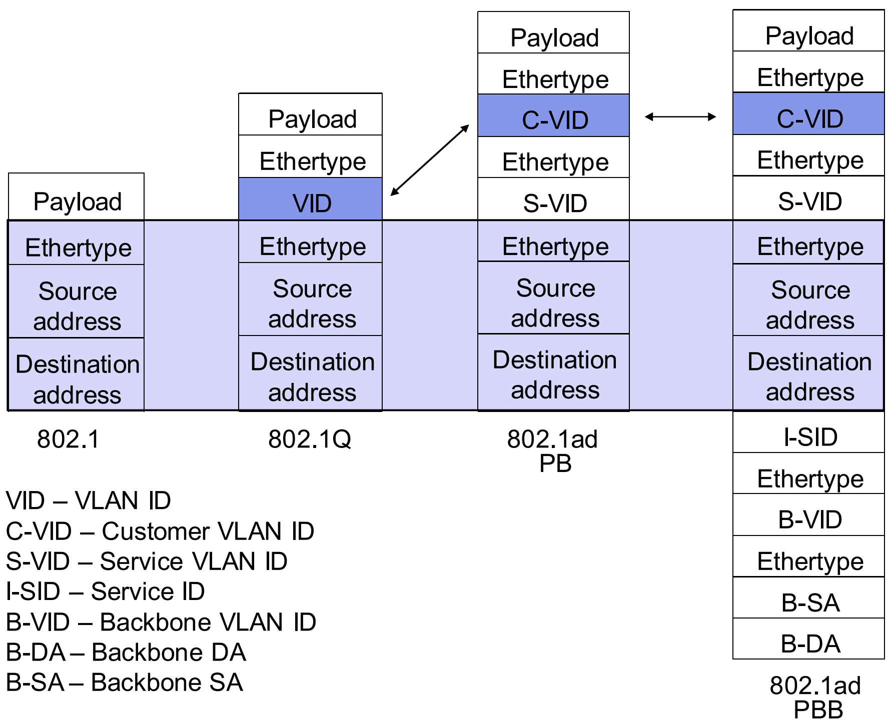


Figure 1. Ethernet frame

In order to overcome provider scalability problem in 802.1Q networks, PB, also known as Q-in-Q, enables VLAN stacking that supports the appending of multiple VLAN tags to the same Ethernet frame [8]. By doing so, a second level of VLANs is introduced, so a service provider network, interconnecting customer networks, can be decoupled from the latter, which use their own VLANs. However, although PB separates the customer and the provider networks, the entire provider backbone Ethernet bridges need to learn customer MAC addresses and to maintain their forwarding tables, which limits its applicability in large networks.

PBB addresses these issues and defines a new encapsulation (i.e., MAC-in-MAC), by adding a MAC header dedicated to the service provider. By doing so, PBB adds both a Backbone source and destination MAC addresses (i.e., a Backbone VLAN ID (B-VID) and a Backbone Service ID (I-SID) [7]) to the PB frame. Using the Backbone and the customer MAC addresses, there is a clear separation between service provider and customer networks. Therefore, the switches in the provider networks do not need to maintain the forwarding tables of the customer networks. It is worth noting that the potential use of 24 bits for the I-SID tag, allows up to 16 million of possible instances in the provider networks. This, in turn, solves the scalability problems of the PB [7].

One of the problems present in both PB and PBB networks is that both rely on the family of Spanning Tree Protocol (STP) that can be very slow in large networks. In PBT [1], however, Ethernet flows are traffic engineered by adding connection-oriented functionalities. In PBT networks, an external management or control plane is responsible for establishing, maintaining and releasing the connections. Therefore, there is neither MAC learning or STP mechanisms in PBT networks. The 12 bits B-VID and 48 bits B-DA fields together form a 60 bits globally unique identifier which is used to forward the Ethernet frames (i.e., a virtual circuit identifier), resulting in PBT tunnels. Deployment of PBT is enabled by making a small alteration to the normal Ethernet behavior. Therefore, PBT can be easily implemented on existing Ethernet hardware.

MPLS-TP

MPLS Transport Profile (MPLS-TP) came out as a result of the joint agreement between the IETF and the ITU-T. The standardization work was initiated by the ITU-T to define a new transport oriented packet network standard referred to as Transport MPLS (T-MPLS). The idea behind MPLS-TP is to join the functionalities of MPLS connection-oriented packet switched networks and the operations and capabilities of transport networks. It is defined as a profile of MPLS-TE [9] and the pseudo-wire [10] architecture. As a Layer2 Packet Transport Technology, MPLS-TP can be used for Ethernet service delivery. The main characteristics of MPLS-TP are:

- It allows the data plane and control plane separation, which was not the case with MPLS, operating under either a centralized Network Management System (NMS) or a distributed control plane.

- It supports a wide range of provisioning, OAM and recovery capabilities, equivalent to those of SONET/SDH networks, and no necessarily dependant of the control plane.
- The MPLS-TP messages are sent in line (differentiated from the data), and can work in the absence of IP protocol support in the transport equipment.

2.1.2 WSON

The DWDM technology relies on encoding the signal that is being transmitted using different optical frequency (wavelength). Each wavelength forms a different optical channel that can carry the information that is independent of the information that other optical channels within the same fiber are carrying. The wavelength channels are switched at the intermediate nodes, either in electrical or in optical domain. The former one requires the optical-electrical-optical (OEO) conversions for the wavelength switching. Therefore, a node disposes transceivers that are necessary to decode and regenerate the signal. On the other hand, all-optical switching eliminates the well-known electronic bottleneck, using network elements such as Optical Cross-Connects (OXCs) and Optical Add Drop Multiplexers (OADMs). By doing so, OEO conversions when switching at intermediate nodes is eliminated, leading to the reduction of the network cost.

WSON are transport networks which rely on DWDM technology, and which are capable of switching wavelengths in optical domain. Since multiple optical channels are multiplexed within a single optical fiber, the total transport capacity of the fiber is significantly increased. Therefore, WSON are one of the most efficient ways to cope with the demand for a higher transport capacity.

2.1.3 Multi-layer networks

A MLN can be defined as a TE domain comprising multiple data plane switching layers either of the same Interface Switching Capability (ISC) (e.g., TDM) or different ISC (e.g., TDM and PSC) [6]. A particular case of MLN is a multi-region network (MRN), defined as a TE domain supporting at least two different switching types (e.g., PSC and TDM) [6]. As it is mentioned before, in this thesis we focus on a MLN, which integrates the following switching layers: connection-oriented packet switching provided by Carrier-grade Ethernet and lambda switching provided by WSON. Having such a MLN, we exploit the benefits of connection-oriented packet transport such as finer bandwidth granularity (i.e., flexibility), scalability, QoS and the cost-efficient advantage of Ethernet as well as a high transport capacity provided by WSON.

A Label Switched Path (LSP) always starts and ends at the interfaces of the same switching type. In a MLN, network nodes can be either *single switching type capable* or *multi switching type capable* [6]. A single switching capable node can terminate data links supporting only one ISC. On the other hand, a multi switching capable node can terminate data links with different switching capabilities. Furthermore, multi switching capable nodes can be classified as *simplex* or *hybrid nodes*. In a *simplex node*, each data link is connected to the node by a separate link interface. Therefore, this type of nodes can terminate data links with different switching capabilities, but it cannot interconnect the different switching capabilities since there is a separate link interface for each switching capabilities. On the other hand, a hybrid node disposes an internal link connecting two different switching elements. Therefore, since the two switching elements are internally interconnected within the hybrid node, it is possible to terminate the resources of one switching capability and to provide adjustment for other switching capability. One example of a hybrid node supporting PSC and LSC is shown in Figure 2.

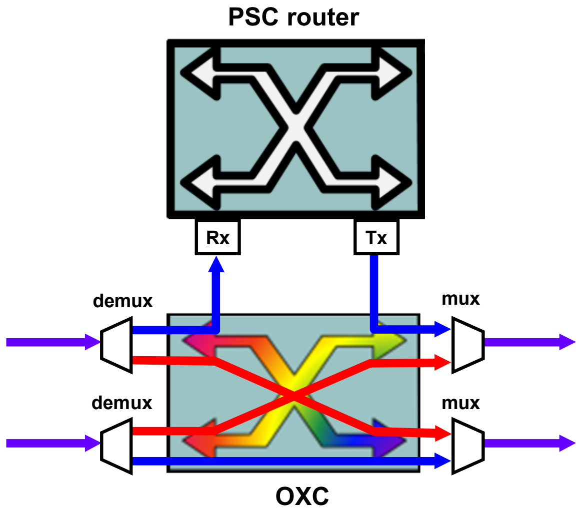


Figure 2. Hybrid node

2.2 GMPLS-enabled control plane

The main idea behind the GMPLS was to develop a set of protocols that would allow a dynamic service provisioning in transport networks. Moreover, the target was to provide a control plane that would be able to control networks supporting different switching technologies. In order to understand the concept and the functionalities of the GMPLS control plane, we will give a brief overview of MPLS.

2.2.1 From MPLS to GMPLS

MPLS is a mechanism for data forwarding standardized by IETF [11]. The forwarding of each packet is based on an associated label. At each router across the network, the packet is forwarded according to the incoming label, which is then swapped by an outgoing label. Routers that perform routing based on labels inside a MPLS networks are called Label Switch Routers (LSR). The nodes in a MPLS network (i.e., LSR) maintain a look-up table (i.e., Next Hop Label Forwarding Entry - NHLFE) containing a mapping of {incoming interface, incoming label} to {outgoing interface, outgoing label} [11]. Therefore, looking up the values in the NHLFE, the packet is forwarded and a new label replaces the incoming one.

In that sense, it is noticed that the operation used in MPLS, from {input label, incoming interface} to {output label, outgoing interface}, can be applied to other switching technologies, not only to Packet Switched Capable (PSC) and Layer2 Switching Capability (L2SC) networks. For example, timeslots in Time-Division Multiple (TDM) networks, or lambda in WDM networks, can be represented as a label used in MPLS. Therefore, the notion of label can be generalized, and the work was named Generalized MPLS. However, MPLS still needed some extensions so it could be applicable to different switching technologies. Therefore, there were added other functionalities to GMPLS in order to overcome the limitations in MPLS (e.g., bidirectional connections, separation of data and control plane, etc). Finally, GMPLS was extended to support different switching type of technologies [3] such as PSC, L2SC, TDM, Lambda Switch Capable (LSC) and Fiber-Switch Capable (FSC).

Unlike MPLS that manages both, data and control plane, GMPLS operates only within the control plane. A control plane can be divided into several functional components responsible for network elements discovery, connectivity and network topology distribution, link management, path computation, distribution of the information about available resources, resource allocation, connection set up, maintenance and releasing, protection and restoration. To this end, GMPLS control plane specifies the following protocols:

- Resource Reservation Protocol - Traffic Engineering (RSVP-TE) [5] and Constraint-Routing Label Distribution Protocol (CR-LDP) [12] as signaling protocols. However, we should stress that the work on the latter one has been stopped by IETF in February 2003 [13], announcing, thus, the RSVP-TE as the only GMPLS signaling protocol. RSVP-TE is responsible for set up, maintenance, modification and termination of a LSP in a GMPLS network. GMPLS controllers exchange the signaling messages between themselves, reserving the resources at the data plane. The whole signaling procedure for the LSP establishment in a MLN will be explained in details in the Section 2.2.6.
- Open Shortest Path First (OSPF-TE) [14] and Intermediate System - Intermediate System (ISIS-TE) [15] as routing protocols. GMPLS routing protocols are responsible for distribution of information that will be used for the path computation necessary for the LSP establishment. Network resources needed for the path computation are modeled as a TE links. Attributes and capabilities of all the TE links within one control domain are stored in a TE Database (TED). Therefore, each node has a full visibility of the network state and resources. Finally, using the information stored in a TED, a path computation algorithm calculates the route from the source to the destination node.

- Link Management Protocol (LMP) [16], responsible for the neighbor nodes discovery and link management. This protocol operates between two adjacent nodes discovering the capabilities and identifiers of the links that connect them. The protocol covers the following tasks: control channel management, link connectivity verification, link property correlation and fault isolation [16].

In PSC networks, control messages can be sent over the same links as the data traffic. In that case, when a switch receives a packet and reads the packet header, it will differentiate if the packet belongs to a control message or it is the data information that is being transmitting. This is called *in-band* model. On the other hand, in some cases, control messages are transmitted over separate channels, links or even network. For example, in WDM networks, one optical channel can be reserved for control messages, while the rest of the optical channels are reserved for the data transmission (*in-fiber-out-of-band* model). The model where are used different links of network for control plane messages is called *out-of-fiber-out-of-band* model [3]. It should be born in mind that GMPLS does not specify the exact implementation of the control channels.

2.2.2 Peer, Overlay and Hybrid model

In a MLN controlled by GMPLS control plane, not all the information is shared between the involved layers. Depending on how much information is distributed, GMPLS control plane can be deployed in three different architectural models:

- **Peer or unified model:** All the nodes in the network have full visibility of the network topology and the resources availability. Therefore, deploying this model in a MLN, a single control plane is used for the establishment of a LSP. The main advantage of this model is that, the usage of the network resources in all the layers can be optimized when setting up a LSP, encompassing both switching technologies. However, this model faces the scalability problems in networks with high number of nodes [3].
- **Overlay model:** The TE information is shared only between the nodes belonging to the same network layer. Therefore, the nodes do not have the full visibility of the network resources. For each switching layer a different control plane is deployed, preserving a strict separation between routing and signaling protocols. Although this model does not have the scaling problem, it is still needed to manage the interactions between two control planes.
- **Hybrid or augmented model:** This model represents the mixture of the peer and the overlay models, where some information can be interchanged between the layers, such as topology of the involved layers, or some of the TE information.

2.2.3 Unified control plane for MLN networks

The main goal of a MLN is to optimize the establishment of end-to-end connections, favoring the collaboration among the involved switching layers. Moreover, the objective is to avoid congested network elements during the provisioning of LSP, increasing the network resource efficiency of all the involved layers. This is referred to as multi-layer traffic engineering (TE) [17]. A unified control plane architectural model enables full TE strategies and decisions since provides a single TED repository at each node to collect the TE information relative to all the layers. Consequently, efficient path computation strategies can be devised aiming at optimizing the network resources utilization across the whole MLN.

It is worth mentioning that, even though the GMPLS control plane was developed to support networks with different types of switching capabilities, the research community has recently begun to work on the problem of internetworking these networks. In this context, in the following section, one of the most important GMPLS control plane concepts for the multi-layer networks are addressed: the Forwarding Adjacency (FA) and virtual links.

2.2.4 Virtual Network Topology

In order to favor the cooperation between the switching layers, the set of lower-layer connections (i.e., LSPs) forms a VNT for the upper-layer [6]. Figure 3 shows an example of a VNT. Signaling messages are exchanged within the control plane, in order to set up connections at the data plane formed by Ethernet and WSON layer. The established lightpaths (i.e., red and blue) in the WSON layer provides a necessary connectivity to deliver Ethernet traffic. By doing so, WSON forms a VNT for the Ethernet layer, connecting the nodes 10.1.0.1 and 10.1.0.5, and 10.1.0.1 and 10.1.0.6. The VNT can be accomplished through exploiting FA and / or virtual TE links. In the following sections, these two concepts will be explained in details.

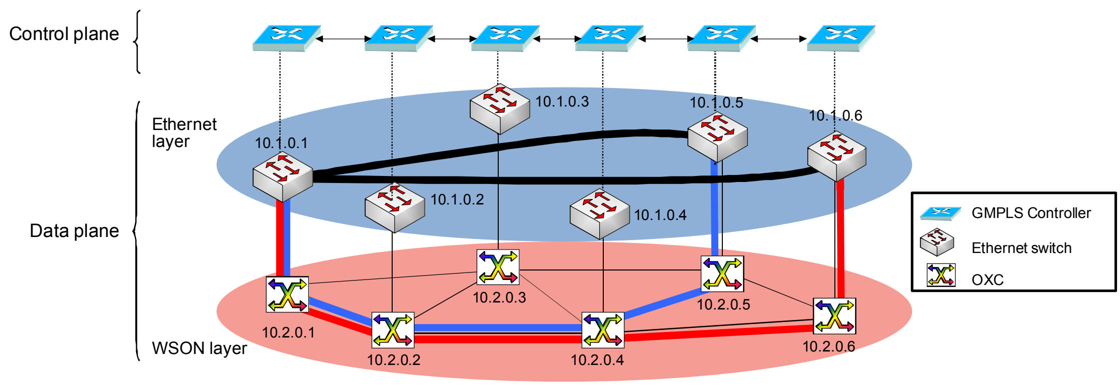


Figure 3. An example of VNT in a MLN

2.2.5 FA TE and virtual links

A FA is a control plane concept where the lower-layer LSPs may be used for forwarding the upper-layer LSPs [6]. For instance, first an optical LSP (i.e., lightpath) is set up. Then, higher-layer LSPs are established / nested over such an optical connection as long as sufficiently unused bandwidth is available in the optical LSP. Indeed, once the optical LSP is established, it becomes a FA-LSP. This, in turn, is advertised by the routing protocol as a TE link at the upper-layer (i.e., Ethernet layer). In consequence, subsequent path computations can use this TE link when routing incoming Ethernet LSP requests. It is worth mentioning that a FA does not require (even discourages) the maintenance of a routing adjacency between the higher-LSP head-end (i.e., source node) and the tail-end (i.e., destination node). In this way, the FA-LSP enables us to implement multi-layer LSP network control mechanism in a distributed manner. This yields several advantages such as more efficient use of the networks resources, simplification of the complexity of the control plane interactions between the layers, unification of the addressing space, etc. In other word, FAs result a useful and powerful tool for improving the scalability of the GMPLS MLN networks.

A virtual link is defined as a lower-layer FA LSP which is pre-computed, but not established (i.e., it is not signaled at the control plane nor switched at the data plane) [6]. That means that a virtual link represents the potentiality to set up a FA LSP in the lower layer to support the TE link that has been advertised [18]. A virtual link can be either fully or partially determined before its establishment. In case it is fully determined, the path of underlying LSC LSP is entirely pre-computed and stored in the TED. In case it is partially determined, only some of the information is known (e.g., the head-end and the tail-end nodes of the link, some of the nodes of the entire path, etc.) allowing flexibility for the establishment of the associated lower-layer LSP. The establishment of the corresponding lower-layer LSP will be triggered by an upper-layer LSP which needs the advertised link for its establishment. Once the lower-layer resources are actually occupied following the standard GMPLS procedures, the established virtual link can be used as a regular link to accommodate the upper-layer LSPs. As soon as there is no upper-layer LSPs over the virtual link, the GMPLS tear down signaling message is sent and the corresponding lower-layer resources are released. Therefore, the virtual link is no longer active, but it is still advertised as an upper-layer TE link by the routing protocol and stored in the TED. Hence, when needed, it can be again signaled and used.

The main advantage of using virtual links instead of pre-established FA TE links in VNT configurations is that, in the latter, the associated optical FA LSPs occupy resources (e.g., wavelength channels) that may not be optimally computed. This may preclude the use of such optical resources to set up other, eventually more appropriate, LSC LSPs. Conversely, virtual links, as mentioned, are just pre-computed but not established. Therefore, no resources are occupied if no upper-layer LSPs exist over such a link. Finally, although this is not our case, it is worth noting that one of the important aspects of virtual links and FA TE links is that they convey data plane potential connectivity information to the upper layers, especially when the TED is not unified.

2.2.6 RSVP-TE signaling for MLN

GMPLS signaling mechanism supports the dynamic creation of FAs in MLN scenarios with dynamic negotiation of link local and remote identifiers by introducing the LSP_TUNNEL_INTERFACE_ID (LTII) object in both RSVP-TE Path and Resv messages. The LTII object typically includes an unnumbered TE link composed of a node identifier (IPv4 address) and an interface identifier (32-bit non-zero integer) [6]. Such an unnumbered TE link is used to identify both ends of the created point-to-point upper-layer FA TE link. Recall that the attributes of such a FA TE link are derived from the established lower-layer FA LSP that may encompass several nodes and links at the optical domain. In a dynamic VNT configuration, unnumbered interface identifiers are dynamically assigned during the establishment of a FA TE link.

Figure 4 shows an example of the L2SC LSP establishment in a MLN composed of both L2SC switches and OXCs. It is assumed a pre-computed virtual link between the nodes 10.1.0.3 and 10.1.0.4. Recall that the advertisement of the virtual TE link does not preclude setting up the associated lower-layer FA LSP.

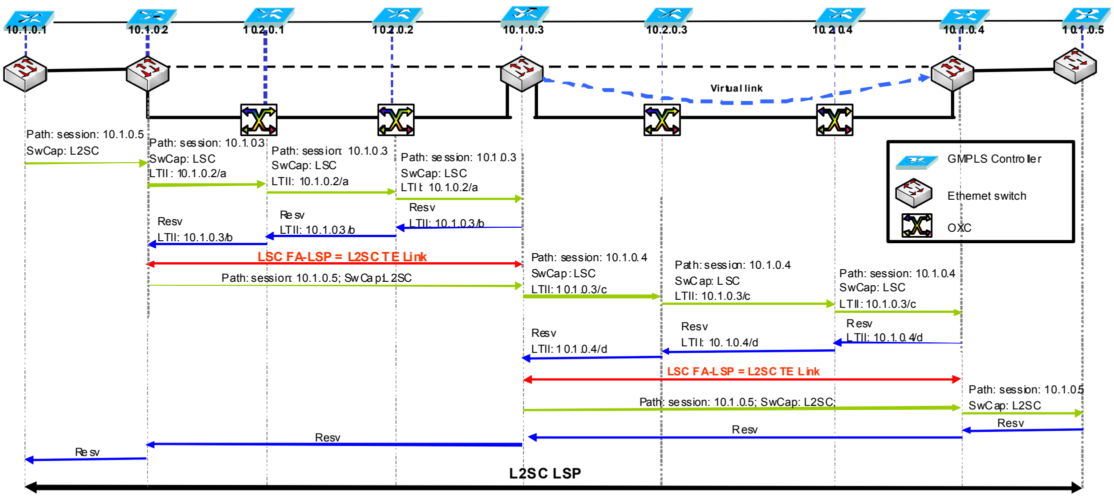


Figure 4. L2SC LSP establishment in a multi-layer network

In the example, a L2SC LSP between nodes 10.1.0.1 and 10.1.0.5 is requested. The computed path using the TED information as input is formed by the following nodes: 10.1.0.1, 10.1.0.2, 10.2.0.1, 10.2.0.2, 10.1.0.3, 10.1.0.4 and 10.1.0.5. Observe that, to reach the 10.1.0.5 node, the computed LSP needs to cross the boundary from the upper layer (L2SC) to the lower layer (LSC) between the nodes 10.1.0.2 and 10.1.0.3. On the other hand, as mentioned before, there is an advertised virtual link between the nodes 10.1.0.3 and 10.1.0.4 whose TE attributes (e.g. unreserved bandwidth, ingress node, egress node, interface identifiers, etc.) are already stored in the network TED. Taking into account that a single GMPLS control plane instance (i.e., routing dissemination) is used, all the nodes have the complete picture of the whole network. Thereby, the path across this MLN can be obtained. The computed route is then passed to the signaling protocol as an EXPLICIT_ROUTE_OBJECT (ERO), which is inserted in the RSVP-TE Path message.

When the Path message reaches the node 10.1.0.2, it will realize that there is a change of layer in the computed path. Then, using the TED information and the computed route, it resolves the other end of the change of layer, which results the node 10.1.0.3. Consequently, a LSC FA LSP needs to be established between 10.1.0.2 and the 10.1.0.3. Following the standard RSVP-TE signaling procedures for MLN, the lower-layer LSP will be established inserting the LTII object to create the upper-layer FA TE link between 10.1.0.2 and 10.1.0.3 nodes. Once the LSC FA LSP is set up, the TE attributes of the newly created FA TE link are flooded by the routing protocol and stored in the TED of all the network nodes. After that, the node 10.1.0.2 will directly send a RSVP-TE Path message to the node 10.1.0.3, at the L2SC switching capability, using the created FA TE link. When the Path message reaches the node 10.1.0.3, this node uses the remaining computed ERO to determine that the virtual link connecting 10.1.0.3 and 10.1.0.4 needs to be used. This requires that the associated lower-layer FA LSP (i.e., through the nodes 10.1.0.3, 10.2.0.3, 10.2.0.4 and 10.1.0.4) is actually activated / set up occupying the wavelength channels resources at each underlying optical link. After the virtual link is established / activated, the L2SC LSP can be set up using the established FA TE link (i.e., the virtual link). Using the ERO information, a RSVP-TE Path message is sent from the 10.1.0.4 towards 10.1.0.5 to set up the L2SC LSP. Finally, a RSVP-TE Resv message will be sent backwards from the destination to the source node, reserving the resources across the route.

2.2.7 Static and dynamic VNT configuration

VNT configuration can be defined statically [19][20][21][22][23][24][25], before any request, or dynamically, triggered by a signaling request from the upper – layer [17][27][28][29][30][31][32][33][34][35]. The former approach relies on designing a VNT (i.e., set of pre-established FA LSPs) for a given initial traffic demand. The pre-established FA LSPs are not released even if there are no upper-layer connections using the corresponding FA TE link. In dynamic VNT configuration, the resources are reserved online, that is, optical resources are occupied as the signaling mechanism is accommodating the incoming upper–layer LSP requests. A LSC LSP is torn down and the resources are released when there are no upper-layer connections using the corresponding FA TE link.

The static VNT configuration approach provides a network that is more stable, if the given traffic demand does not vary significantly during the time. Therefore, this approach might be the preferable choice for the network operators. Moreover, static VNT configuration does not require expensive reconfigurable equipment and the total signaling complexity is much lower comparing with the dynamic VNT configuration approach. Nevertheless, a VNT that is designed and optimized for a given traffic demand may not be able to satisfy dynamic and unpredictable traffic changes. In other words, the resources at the lower-layer being pre-defined may not be optimally reserved and, thus, they might be used for the establishment of more appropriate lower-layer (FA) LSPs. On the other hand, using the dynamic VNT configuration approach, LSC LSPs are established and torn down dynamically, depending on the requested traffic. Therefore, this approach responds more efficiently to dynamic traffic changes and it provides better usage of the network resources at both involved layers. Moreover, in case of a network element failure, the dynamic VNT configuration approach provides more efficient real time recovery, since the LSC LSP affected by a failure can be dynamically re-routed if there are enough resources at the lower-layer. Finally, for these reasons, in this thesis dissertation we focus on online VNT reconfiguration, under dynamic traffic changes. However, in some cases, we still consider pre-established FA TE links and/or pre-computed virtual links.

The main problems and challenges related to the dynamic VNT reconfiguration approach and LSP provisioning in a MLN are explained in details in Chapter 3.

2.3 Chapter Summary

This chapter provides an overview of the GMPLS control of a MLN formed by CO Ethernet and WSON. Initially, a short description of the transport networks used in this thesis is given. Then, we have presented the evolution of MPLS to GMPLS control plane, as well as the main features and functionalities of GMPLS control plane for MLN. In this regard, we have explained the concepts of VNT through FA and virtual links, and the signaling procedure for their establishment.

In the following chapters, we present some of the challenges in a MLN controlled by GMPLS unified control plane, regarding dynamic path computation algorithms and LSP provisioning (Chapter 3), as well as protection schemes and mechanisms (Chapter 4). After describing and classifying some of the proposed solutions found in the current literature, we present our contribution regarding these topics.

Chapter 3

LSP Provisioning in MLN

In this chapter, the main problems and challenges for LSP provisioning in a MLN (CO Ethernet over WSON network) when a GMPLS unified control plane is deployed are identified. Consequently, we address the problems regarding dynamic VNT reconfiguration and path computation in such a MLN. Then, we review and classify path computation algorithms found in the literature. Finally, we present the contributions of this dissertation regarding path computation and LSP provisioning in a MLN.

3.1 Introduction

By introducing the concept of MLN/MRN in [6], an end-to-end LSP can be established crossing the switching layers several times as long as the switching capability constraints (i.e., switching adaptation among the layers) are satisfied. In order to fully utilize the high capacity provided by WSON networks, it is important to make efficient TE decisions and to develop effective strategies, so the upper-layer traffic is dynamically accommodated over the lower-layer. Moreover, an efficient real-time path computation algorithm is needed, which will maximize the network throughput, increasing the total amount of traffic that a MLN can transport and minimizing the total network cost.

3.2 Path computation

Path computation is the process of computing a route from the source node to the requested destination node. A path computation algorithm uses the topology and network resources information stored in the network TED. The data stored in the TED is modified and updated every time any network state change occurs. This is accomplished by a GMPLS routing protocol (i.e., OSPF-TE) which immediately floods the current network state information (e.g., availability of the network resources, topology connectivity, etc.) to all the network nodes. Therefore, having a full visibility of the current state of the network topology and resources, it is possible to compute an efficient path to the destination, so a LSP can be established.

There are two different strategies for a route computation in GMPLS controlled networks: fully distributed source-based routing and by means of a centralized Path

Computation Element (PCE) [26]. In the former one, each node that gets a connection request computes a path to the destination node. The main drawback of the source-based routing is that it may impose too much burden in the network nodes, maintaining the TED in every node, especially in large MLN/MRN infrastructures. Moreover, it cannot benefit from information that is not distributed via the routing protocol and it is harder to integrate with operator policies and BSS/OSS. In the latter one, however, a central PCE entity is responsible for the path computation, which must construct its own TED. Therefore, using PCE for path computation purposes, network nodes are relieved from the routing burden. Moreover, having the dedicated hardware (i.e., PCE), it is possible to compute more complex and more efficient algorithm. However, the main drawback of using PCE is increased time needed for LSP establishment, due to PCE-PCC (Path Computation Client) communication, comparing with distributed source-based routing. Moreover, in highly dynamic scenarios, the PCE is not able to update its TED so quickly, leading, thus, to inefficient path computation.

Traffic Engineering information (i.e., attributes and state of the network nodes and links) distributed by a GMPLS routing protocol allows a path computation algorithm to select routes across a MLN, satisfying particular requirements and constraints imposed by a connection request. The simplest path computation algorithms (e.g., Dijkstra and Bellman-Ford) calculate a route taking into account only the number of traversed hops. Therefore, the algorithm always chooses the route with the minimum number of traversed nodes. In GMPLS controlled network, the simplest algorithm looks at the TE metric link attribute. Therefore, the algorithm chooses a route with minimal number of hops if TE metric of each network TE link is equal to one. A more efficient path computation algorithm, however, should deal with other constraints and network element attributes to provide better paths. For example, one of the basic requirements of a constrained algorithm is that a path computation algorithm should avoid using the TE links that do not have enough resources for carrying the requested traffic. Furthermore, in the context of WSON, the algorithm should check whether there is an available wavelength converter at an intermediate node when the WCC cannot be satisfied. In MLN, such a constraints-based path computation algorithm needs to exploit the dynamic cooperation between the layers, in order to enhance the overall network resource efficiency. It should provide efficient TE strategies, increasing the network resource efficiency of both involved layers.

Once the route is determined, it is passed as an ERO object to the Path message of GMPLS signaling protocol (i.e., RSVP-TE), so the LSP can be established as it is explained in the Section 2.2.6.

3.3 Traffic Engineering challenges in MLN

The LSP establishment in a MLN where different switching technologies are integrated and controlled by a single instance of GMPLS control plane brings many challenges. Herein, we will introduce the most important ones, regarding VNT reconfiguration and MLN path computation.

Scalability problem

In general, with dynamic LSP provisioning, the more information is available for the path computation, the more efficient / optimal route can be obtained. Therefore, deploying a GMPLS unified control plane for controlling a MLN is a straightforward way to ensure that each node has a full visibility of all the network layers. This, in turn, leads to obtain optimal decisions when computing end-to-end LSP paths. However, maintaining such a huge amount of information, especially with ever-increasing Internet traffic and dynamic VNT reconfiguration, scalability problems may become a bottleneck. Moreover, each node is required to flood the information about any change in the network that occurs to all the neighborhood nodes, increasing, thus, the overall control plane overhead. On the other hand, using a control plane per layer / region (i.e., overlay architectural interconnection model), each control plane is responsible for the network resources optimization for that layer / region. Therefore, each node maintains the TED only within a layer, reducing, thus, the scalability problem. Nevertheless, when computing an end-to-end path having a control plane per layer, although a route segment is optimal from the point of view of one layer, it still may result as an inefficient path from the perspective of the entire network. Furthermore, without knowing about the state of the network resources of other layer / region, sometimes it is even impossible to find a route, causing, thus, a higher connection blocking probability.

Traffic grooming

In order to improve the network performance, it is important to be able to group multiple upper-layer connections requiring low bandwidth into a lower-layer high-capacity connection. This strategy of efficiently grouping multiple and flexible lower-bandwidth (e.g., packet, Ethernet) LSPs from different clients over higher-bandwidth circuit (e.g., optical) LSPs is known as *traffic grooming*. Therefore, the total bandwidth of a wavelength channel is shared between client connections. By doing so, the network capacity is significantly increased leading to better usage of the overall network resources. However, one of the issues regarding traffic grooming is, thus, how to address the computation for optimizing the provisioning of clients LSPs through the lower-layer (lambda) LSPs. In other words, the challenge is how to find the most efficient way to manage the resources at both layers in a collaborative way, satisfying as many connection requests as possible, and minimizing the total network cost under the dynamic traffic patterns.

Pre-provisioning vs. triggered LSP

In a MLN constituted by two switching layers (i.e., L2SC and LSC), when an LSP crosses the boundary from an upper to a lower layer, it may be nested into a lower-layer LSP. Moreover, the LSP can change layer several times, as long as the switching capability constraints (i.e., switching adaptation among the layers) are satisfied.

From the signaling point of view, there are two alternatives for LSC LSP establishment that induces a forwarding adjacency: static (pre-provisioned) and dynamic (triggered) [6]. A pre-provisioned FA-LSP is initiated and configured statically, before any request from the L2SC layer arrives. On the other hand, the lower-layer LSP may be established dynamically, triggered by a signaling request from the upper-layer, traffic demand changes, topology configuration changes or network failures [6]. As it is explained in the Section 2.2.5, the established lower-layer LSPs is then announced as an upper-layer FA TE link, and its TE attributes and characteristics are inherited from the corresponding lower-layer FA LSP. By doing so, the WSON layer provides the connectivity information that will be used for path computation at the CO Ethernet layer as long as the established FA TE link has enough available resources. One of the challenges is, thus, to find an optimal solution in choosing between pre-provisioned and triggered signaling, regarding network performance and quality of service (i.e., delay, loss/disruption, link utilization, residual capacity, reliability, etc.). The main drawback of using the pre-provisioning signaling is that the resources occupied by statically established LSC LSPs could be used by an eventually more efficient LSC LSP. However, static VNT configuration provides a network that is more stable when there are enough resources for accommodating Ethernet connections.

Dynamic path computation algorithm

The main issue related to MLN path computation is how to choose which layer is more suitable for an LSP request accommodation, for given resource limitations and traffic demands. This is even more difficult when the dynamic VNT configuration is required. During the path computation, different constraints have to be taken into account, such as the link interface switching capability for the adaptation between layers, the WCC at the optical layer, the availability of sufficient unreserved bandwidth in the TE links, etc. Moreover, an open issue includes choosing an optimal route that can balance bandwidth utilization and traffic granularity adaptation between the layers. The issue includes development of dynamic routing algorithms to compute path across different technologies that form a MLN.

In the following Section, we review the produced literature and classify their proposed solutions for some of the mentioned challenges specifically related to real-time path computation algorithms and dynamic VNT reconfiguration.

3.4 Related work

In the past few years, dynamic VNT reconfigurations have been studied intensively. Some papers propose the use of a centralized PCE [26] for finding a suitable route (e.g., [27][28][29][30]), while other papers support source-based algorithms (e.g., [17][31][32][33][34][35]). Regardless of the PCE-based or source-based routing it is commonly accepted that a GMPLS unified control plane aims at attaining a better network resource usage across the entire network than when an independent control plane per layer (i.e., different routing protocol instances, separated TED information) is used.

In early research work published in [17], the authors proposed two routing policies for the VNT reconfiguration. Both policies first try to accommodate the upper-layer LSP over an existing lower-layer (i.e., LSC) LSP which directly connects the source and the destination nodes. If such an LSP does not exist or its resources are not available, the first policy will try to establish an upper-layer LSP over multiple lower-layer LSPs. If this is not possible, it will establish a direct lower-layer LSP between the end nodes as long as there are enough available resources. On the other hand, the second policy will first try to establish a new lower-layer LSP which directly connects the source and the destination nodes and, if this is not possible, it will accommodate the connection over multiple upper-layer LSPs. The results show that the number of packet-switched capable ports within the PSC switch is a key factor in choosing one of the proposed policies.

The authors of [30] developed a simple so-called *min-phys-hop* algorithm, in which each physical link has a weight (i.e., GMPLS TE metric) equal to one, so each established lower-layer (optical) LSP has a weight equal to the number of corresponding underlying physical links. The authors discussed several layer-preference policies that are used for comparing the proposed algorithm with the min-hop algorithm where all the physical and FA TE links in the MLN have a weight equal to one. They choose the layer on which the request will be accommodated depending on the configurable parameter α [30]. It is shown that the *min-phys-hop* algorithm reduces the connection blocking probability with almost all possible selections of the input parameter.

The previously mentioned papers take into account only the number of hops as the most important criterion for computing a suitable route. In other research work, additional constraints were taken into account, such as the congestion of optical links and / or LSC LSPs, switching capability penalties, etc. For example, the authors in [31] proposed an algorithm that chooses the layer on which the traffic will be accommodated comparing the capacity utilization of the virtual / logical and physical topology. They proved that this algorithm outperforms the algorithm proposed in [30] in terms of both, connection blocking probability and resource utilization.

In [36], the authors have proposed a distributed VNT reconfiguration method based on a simple heuristic algorithm. In this work, the algorithm adds a new lower-layer LSP if the traffic demand is greater than a pre-defined threshold parameter, which represents the maximum after which is assumed that the lower-layer LSP is congested. On the other hand, underutilized optical LSPs are torn down to favor future connection demands, after confirming that the traffic over the associated FA TE link is below a defined parameter. Before tearing down the link, all the traffic over such a link will be rerouted over an existing VNT. Finally, the impact of the variability of these two parameters on the distributed VNT reconfiguration was investigated. The work in [36] is extended in [37] assuming dynamic traffic changes and proposing a statistical traffic processing scheme. Such a scheme involves traffic measurement, where the statistical traffic rate information is gathered in the network TED. By doing so, the authors have compared the impact of the traffic demand changes on the VNT reconfiguration time.

There are series of path-computation algorithms for MLNs that rely on the auxiliary graphs which are constructed from a network graph. In that sense, the authors in [33][34][35] have transformed a network graph into a channel graph by adding virtual links and nodes, taking into account the switching capability attribute of the TE link interfaces. Specifically, the authors in [34] have developed a graph model which consists of auxiliary vertexes and edges. By manipulating the weight of the edges of the graph, the model can achieve various objectives exploiting different grooming policies and taking into account the following constraints: wavelength-conversion capability, grooming capability, number of wavelengths, and number of transceivers. The work in [34] was slightly extended and formalized in the context of GMPLS and TE attributes in [35]. Specifically, the Interface Adjustment Capability Descriptor (IACD) and the interface switching capability descriptor (ISCD) GMPLS attributes of the routing protocol are used enabling the algorithm proposed in [34] to be applied within GMPLS-controlled MLN. Using this TE information, the authors defined four different policies which assign a different TE metric to the FA TE links. They showed that under different network scenarios, the utilization of these attributes attains an improvement of both the connection-blocking probability and the overall network resource utilization. Finally, we can conclude that, in general, algorithms that rely on auxiliary graphs significantly improve the network performance but at the expense of increasing the complexity, which may become an important restriction when dealing with large-scale networks.

Although it deals with a PCE per layer computation [38], it is worth mentioning the algorithm proposed in [28], since it takes into account both FA TE links and virtual links. The authors proposed a multiplicative and an additive TE metric for virtual and FA TE links in an MLN. In the paper, a VNT manager (VNTM) simplifies the interaction between PCEs that manage two different layers. Results show that the additive policy is more effective in reducing the number of required cross-connections while not affecting network resource utilization. Finally, it should be born in mind that, since there is an upper-layer PCE in charge for computing paths in the upper-layer network, and a lower-layer PCE in charge for lower-layer path computation, when deploying this algorithm it is not possible to change a layer.

In the following, we present the main contributions of this dissertation regarding the conceived path computation algorithms and the performance evaluation of LSP provisioning in a CO Ethernet over WSON controlled by a GMPLS unified control plane. It can be classified into three parts.

- In Section 3.5, we present our work published in [39]. We propose the usage of a *FA TE link timer*. The main idea behind the timer is to reduce the control plane processing overhead caused by successive establishment of a FA TE link between the same pair of nodes. The timer postpones the release of the established FA TE link, when there are no upper-layer connections over such a link for a pre-defined time period.

- In Section 3.6, we present a simple path computation algorithm where both, virtual and FA TE links are deployed, providing the connectivity to CO Ethernet layer. The main objective behind this contribution is to find an adequate solution in choosing between pre-provisioned and triggered signaling, regarding quality of service (e.g., setup delay) and the network performance (e.g., connection blocking probability). Using the proposed algorithm, we have compared three approaches for VNT configuration: *dynamic*, *semi-dynamic* and *virtual*. The work has been published in [40]. The proposed algorithm has been extended during the thesis, and we present its final version in Section 3.7.
- The algorithm presented in Section 3.7 uses a path cost function that depends on the current state of the network resources. Adjusting the parameters in the cost function, a good balance in the usage of network resources at both involved layers can be achieved. Moreover, compared with the algorithms proposed in the literature, it provides efficient grooming decisions attaining the lowest connection blocking probability. The work has been published in [41].

3.5 FA TE Link Timer

As explained in Chapter 2, in a dynamic VNT reconfiguration in a MLN, the creation of a new FA may be triggered by an upper-layer (i.e., CO Ethernet) signaling request. If there are available resources at the optical layer, the FA LSP will be established, and used as an upper-layer TE link. The advertised FA TE link can be then re-used by other upper-layer LSPs as long as enough unused bandwidth is available. When there are no upper-layer connections occupying resources over such FA TE links, the associated lower-layer LSP needs to be torn down releasing all the occupied resources which are then updated in the TED. If, after a while, there is a new higher-layer connection request, which would eventually induce the creation of a new TE link between the same source and the destination nodes, a new signaling request for the re-establishment of the associated lower-layer LSP will be again introduced. If there are available resources, the FA TE link will be established.

The above procedure may cause two main problems when a dynamic VNT reconfiguration is required. First, it can provoke excessive control plane overhead due to the successive setup and teardown of the associated FA LSP between the same pair of nodes. Second, it may increase the setup delay for the Ethernet connections that trigger the underlying FA LSP establishments. Both problems are especially noticeable under highly dynamic traffic patterns. To the best of our knowledge, there are no contributions in the literature that deal with this problem. Therefore, the aim is to avoid successive FA LSP creations between the same edge nodes in order to minimize the two outlined problems. As a possible solution, we propose the usage of a *FA TE link timer*. This timer extends the lifetime of the FA after the last client LSP is released. In other words, when there are no upper-layer LSPs that occupy resources (e.g., bandwidth) over the FA TE link, instead of immediately releasing such a link, the *FA TE link timer* will be started postponing the release of the link until the time expires. By doing so, the FA LSP can be potentially re-used for an eventual establishment of other upper-layer LSPs. It is clear that the actual (optimal) value of the timer will depend on the actual traffic patterns. Consequently, for the different values of the timer, we compare this solution with the standard one, when a FA TE link is released immediately when there are no more actual reservations.

3.5.1 Simulation parameters

The simulations are carried out by a developed simulator which is made using OPNET Modeler [4], allowing us to obtain exhaustive numerical results. More information about the implementation of the simulator can be found in Appendix at the end of this dissertation.

Figure 5 illustrates the used network topology for the performance evaluation. We have implemented selected RSVP-TE functionalities needed for the signaling in a MLN. In this model, simplified RSVP-TE objects are used and the RSVP-TE refreshes are not implemented. As it is explained in previous sections, we consider a peer interconnection model with a unified control plane, and therefore, all the nodes in the network have full visibility about the topology and network resources.

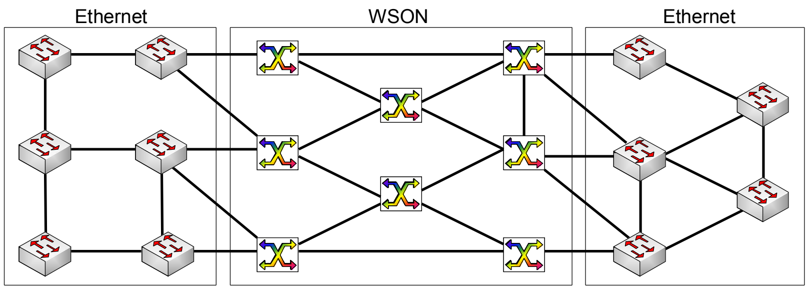


Figure 5. Network topology

L2SC LSP requests arrive to the network following a Poisson arrival process with the average inter arrival time of 1s. Each request starts at a randomly chosen source node belonging to one of the two Ethernet areas, with the destination uniformly chosen among the nodes from the other Ethernet area. FA LSPs are triggered on demand. Every OXC in the WSON area has one wavelength converter. Each CO Ethernet switch has full grooming capability, or, in other words, the switches have enough ports for originating and terminating LSC LSPs. Therefore, there is not restriction on the number of lightpaths to be handled by each Ethernet node. The random policy is used to assign the wavelength and a Dijkstra-based algorithm to choose the shortest path. Each physical link has cost (i.e., GMPLS TE metric) equal to one, while an established FA TE link has a cost derived from the number of underlying physical links it is constituted. The algorithm always tries to satisfy the WCC, and when this is not possible, it checks if there is a wavelength converter at the given node. Furthermore, the WCC is assured during the signaling using the RSVP-

TE LabelSet object. Finally, we assume that during the signaling message forwarding, the only source of delay is the link propagation delay, which is the same for each control link, 3.5ms.

The LSP holding time is exponentially distributed and it varies in different simulation sets to evaluate the network behavior under various traffic loads. Every optical link supports ten wavelength channels. The maximum reservable bandwidth of the higher-layer (i.e., between two Ethernet nodes) TE links is 1Gb/s, while in the optical links is 2.5Gb/s. Since we assume that the number of ports at each CO Ethernet node is equal to the number of wavelengths of optical link, the Ethernet-OXC links and OXC-OXC links have the same maximum reservable bandwidth (i.e., the bandwidth of an optical channel multiplied by total number of wavelengths). For every connection, the requested bandwidth is 100Mb/s. Each data point is obtained simulating $7.5x10^5$ L2SC LSP requests.

3.5.2 Results

One of the key issues is the selection of the FA timer, which is related to the arrival traffic process and the network topology. Herein, we evaluate the network performance having several scenarios in which we use different FA link timer values.

We have used four metrics to evaluate the network performance:

- Setup delay as a function of a traffic load with the fixed value of the timer;
- Setup delay as a function of the timer values with the fixed traffic load;
- Total number of the created FAs during one simulation;
- Connection blocking probability.

Setup delay

Figure 6 shows the setup delay as a function of the traffic load in Erlangs. We have used different values for the timer: 0s, 11s, 22s and 100s. From the figure, it can be observed that the usage of the timer marginally improves the setup delay, and the improvement is slightly higher at the low traffic load. This is because, at low traffic loads, setting up and tearing down FA LSPs happens much frequently than at higher traffic load, since the connection holding time is smaller. Therefore, having the timer, a FA stays longer “alive”, so the other L2SC LSPs can re-use it, without a need for establishing a new FA. As it is expected, in the low traffic load, avoiding the creation of new FAs by postponing their release, the amount of signaling control is lower, as well as the setup delay. On the other hand, we can observe that at high traffic loads, the timer usage has less influence on the setup delay than at low traffic load. The reason behind this is that at high traffic load the connection holding time is higher, and, thus, the L2SC LSPs last longer, as well as their corresponding FA LSPs. Also, at the high traffic load there is a higher probability that a FA link will be re-used by some other L2SC LSPs, since the established FA TE links are less frequently released. Therefore, the FA TE link timer does not have such an impact on the setup delay like at the low traffic load. Finally, we can observe that, having the higher values of the timer, the setup delay is even lower.

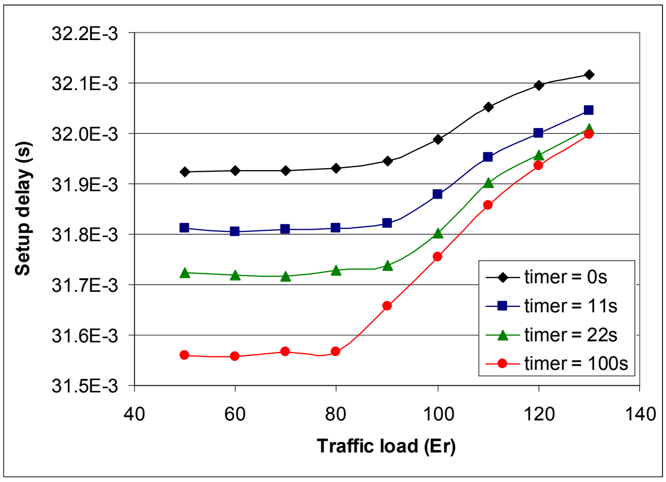


Figure 6. Setup delay

It is worth noting that, in each case, the setup delay is increasing having higher traffic load, which is a consequence of the creation of new FA LSPs which will satisfy the increased load.

Setup delay depending on the timer value

The influence of the timer value on the setup delay is depicted in Figure 7.

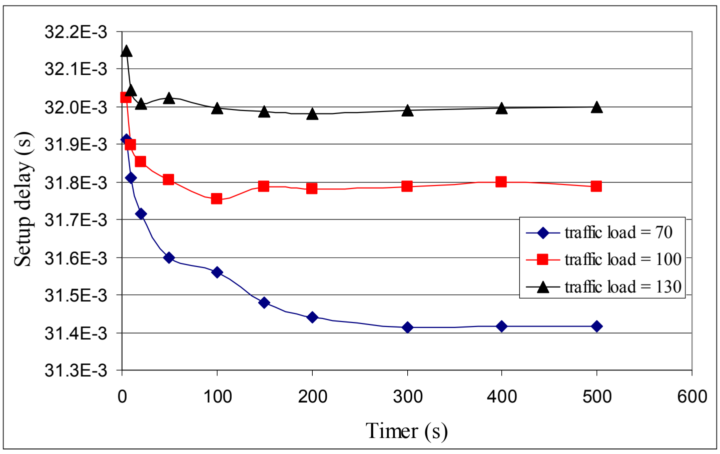


Figure 7. Setup delay as a function of the timer value

The simulation has been run for three different values of traffic load: 70Er, 100Er and 130Er, taking into account the following values of the FA link timer: 5s, 10s, 20s, 50s, 100s, 150s, 200s, 300s, 400s and 500s. From the figure, we can observe that, by increasing the value of the timer, at some point, the setup delay will converge. It can be seen that, with the lower traffic load (i.e., 70Er), the setup delay converges slower while increasing the timer value, than with the higher traffic load. That is because of the fact that the timer has bigger influence on the setup delay with the low traffic load, since the creation of the FA LSPs happens more often. Having higher traffic load (i.e., 100Er and 130Er), by increasing the value of the timer, the setup delay converged much faster. Finally, we can successfully verify that, after some point a higher timer value has no influence on the setup delay and that the network will behave as it had static VNT configuration.

Total amount of triggered FA TEs

In order to comparatively evaluate the control plane overhead, which is directly related to the number of established and released FAs, Figure 8 shows the total number of the created FAs during the simulation.

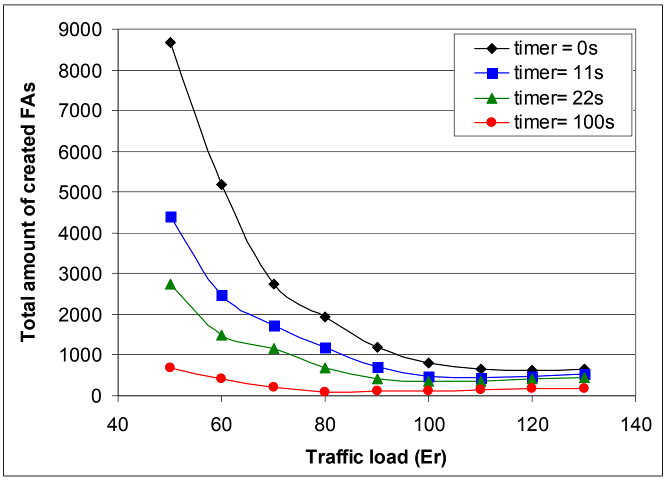


Figure 8. Total amount of the created FAs

The results are obtained running four sets simulations, with the different values of the timer: 0s, 11s, 22s and 100s. It is shown that, without using the timer, the total amount of the created FAs is much higher than in the case when the timer is used (i.e., from ~3 to 9 times at the low traffic load). Therefore, we can conclude that, when using the timer, there are less created links, and they are more used than in the case when the timer is not used. Furthermore, we have confirmed that, by increasing the timer value, the number of the total amount of created FAs is even lower, which means that the amount of control signaling processing and overhead is less.

Blocking probability

As a direct consequence of the approach, too high timer value may be reserving the resources at the optical layer (i.e., wavelength), which could be used for establishing a new, eventually better-utilized, FA LSP. Therefore, one needs to carefully evaluate whether the usage of the timer has a notable impact on the connection blocking probability due to the increased lifetime of the underlying lightpaths. Figure 9 shows the connection blocking probability as a function of a traffic load with and without the timer usage. From the figure, we can observe that, by increasing the timer value, and, thus, keeping the corresponding optical resources unavailable for longer time, the usage of the timer in this scenario does not affect the connection blocking probability, since the optical links dispose of enough resources.

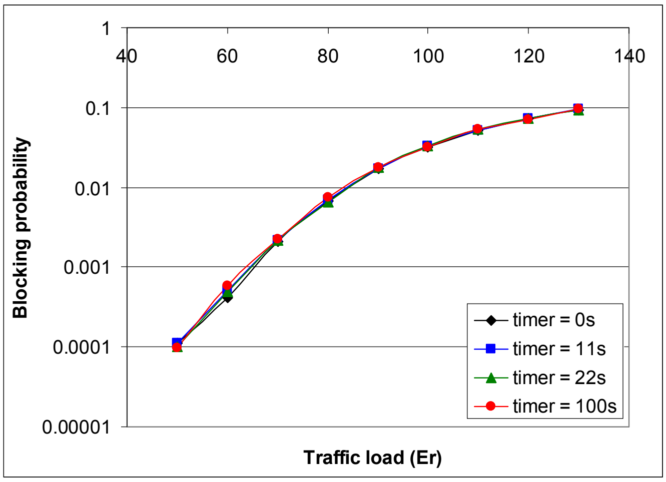


Figure 9. Blocking probability having ten wavelengths per optical link

However, if we consider the same network topology, and for the sake of simplicity, if each optical link has only two wavelength channels and each node at the WSON layer has one wavelength converter, the impact of the timer on the connection blocking probability is high and appreciable. In the Figure 10, it is depicted the connection blocking probability as a function of the traffic load. The results are obtained after running the set of three simulations with the following values of the FA link timer: 0s, 20s and 100s.

We can observe that the usage of the timer has a significant impact on the connection blocking probability, especially with the low traffic loads. Without the timer, we reserve the resources at the optical layer for a FA TE link just for the time that there are L2SC LSPs over such a link. When there are no higher-layer connections using the FA LSP, the resources at the optical layer will be immediately released. Therefore, the connection blocking probability is lower in the case when we do not use the timer, since we are not reserving the resources at the optical layer that are not needed at that moment. Furthermore, the impact of the timer usage is higher with low traffic load since the creation of FAs happens more frequently. Likewise, by increasing the value of the timer, the connection blocking probability is even higher since we are keeping a FA "alive" for a longer period, reserving the corresponding wavelength that might be use for the establishment of eventually more appropriate FA TE links.

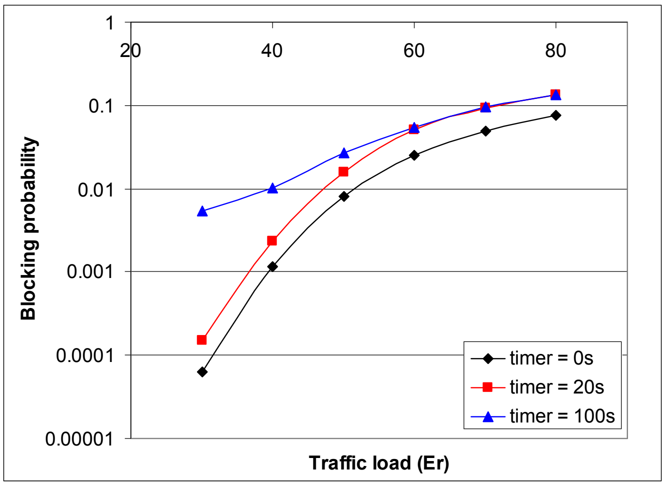


Figure 10. Blocking probability having two wavelengths per optical link

3.5.3 Conclusion: FA TE link timer

The main objective of using FA TE link timer is to decrease the amount of the signal processing due to the creation of a new FA link between the same pair of nodes. Moreover, by postponing the release of an established FA TE link, the setup delay is slightly improved. However, although the difference between the number of created FA LSPs with and without use of timer is high (up to 9000 FA LSPs), this still has a small influence on the total signaling, since there are around 750000 L2SC LSP requests and the connection blocking probability does not exceed 10% (i.e., more than 90% of requested connections are established).

Excessively long timers tend to behave asymptotically as if the VNT was statically configured. As a result, it negates the benefits of dynamic establishment of connections increasing, thus, the connection blocking probability, since the optical layer resources might be reserved for a FA that is no longer being used by any L2SC LSP. Nevertheless, if there are enough resources at the WSON layer, the FA TE link timer does not worsen the total network throughput (i.e., the connection blocking probability remains the same as when the timer is not used). Finally, we can conclude that the challenge states a trade-off between the setup delay and the blocking probability, considering the different values of the timer, and taking into account that the usage of the timer reduces the overall control plane processing.

3.6 Path computation algorithm for MLN: Dynamic, Semi-Dynamic and Virtual Approach

In the Section 3.4, we have summarized the state of art regarding path computation algorithms conceived for a MLN controlled by GMPLS control plane. We could see that, to the best of our knowledge, only the work proposed in [28] considers both virtual links and FA TE links. Moreover, some of the papers (e.g., [17][32]) establish an upper-layer LSP using exclusively either the virtual or the physical topology. The novelty in our work is that we propose a simple path computation algorithm for an MLN in which a VNT can be reconfigured dynamically, integrating both layers involved. Furthermore, the algorithm takes into account both virtual and FA TE links, favoring their usage over a new lower-layer (LSC) LSP establishment. The main objective of this contribution is to find an adequate solution in choosing between pre-provisioned and triggered signaling, regarding quality of service (e.g., setup delay) and the network performance (e.g., connection blocking probability). In case of pre-provisioned signaling, it is very important to maximize the benefits provided by the usage of pre-established FA LSPs and pre-computed virtual links.

For the VNT configuration, we consider three strategies or policies: *dynamic*, *semi-dynamic* and *virtual*:

- **Dynamic**: all the FA TE links are computed dynamically, online and on demand, triggered by upper-layer LSP requests. If there is an FA TE link that is not used by any upper-layer connection, the link will be removed from the TED and the occupied resources will be released.
- **Semi-dynamic**: before serving any upper-layer LSP request, a set of LSC LSPs is pre-established (i.e., the optical resources are occupied at the data plane), inducing an associated set of FA TE links. The routing protocol disseminates these FA TE links as regular links which are stored in the TED. Consequently, these FA TE links can be used for the MLN path computation. Moreover, if the existing FA TE links cannot accommodate an upper-layer request (e.g., lack of available bandwidth), it is possible to establish a new lower-layer LSP which, consequently, creates a new FA TE link. Therefore, this scenario combines both pre-established FA TE links and dynamically created FA TE links. Observe that pre-established FA TE links are never released even if they are not being used.
- **Virtual**: the set of virtual links between border nodes is chosen and pre-computed, representing potential lower-layer LSPs that can be established. Their establishment will be triggered by upper-layer requests when needed, following a pre-defined route. Therefore, the corresponding lower-layer resources are neither reserved nor cross-connected until the signaling is triggered by the upper-layer LSP request.

Finally, the underlying optical LSPs will be released when there are no connections over the corresponding virtual link, but still, the virtual link will be stored in a network TED as a potential LSC LSP. Similar to the semi-dynamic scenario, apart from virtual links in the network, it is possible to dynamically establish new LSC LSPs, and thus create the associated FA TE links.

To this end, in the following section we propose a MLN TE path computation algorithm for the evaluation of these three approaches.

3.6.1 Path computation algorithm for dynamic VNT reconfiguration

In our model, the higher-layer source nodes are responsible for deciding whether to establish a new LSP or not. When the source node receives an incoming L2SC LSP request, according to the TED information, like in [17], it firstly checks if a direct FA TE link between the source and the destination nodes with enough unreserved bandwidth exists. If not, it checks if there is a pre-computed but not established virtual link which directly connects the edge nodes. If such a virtual link exists, the source node starts the signaling for the establishment of the corresponding LSC LSP (i.e., lightpath) following the associated pre-computed route. Once the LSC LSP is established, the L2SC LSP will be routed over the newly established FA TE link. Observe that virtual links are used only in the virtual approach.

If there is neither existing FA TE link nor pre-computed virtual link that directly connects the source and the destination nodes, the on-line path computation algorithm is executed aiming at maximizing the benefits provided by the usage of the already pre-established FA LSPs (i.e., upper-layer TE links) and pre-computed virtual links. The considered path computation algorithm relies on computing the shortest path cost. Illustratively, an example of the TE link costs is shown in Figure 11. Observe that between the node A and the node D there is a pre-computed but not established virtual link, as well as a FA TE link. The path costs are calculated as follows:

- The TE link metric used for an established FA TE link is set to the number of the corresponding underlying optical links (i.e., *metric (FA TE link) = num_of_underlying_opt_links*). In the example shown in the Figure 11, the cost of the FA TE link is set to three, since there are three optical links that constitute the link.
- The TE link metric used for a non-active virtual link is set to the number of underlying optical links plus one (i.e., *metric (non-active virtual links) = num_of_underlying_opt_links + 1*). This favors the use of FA TE links with respect to the virtual links, since in the former the resources are already signaled and occupied. Observe that the cost of the virtual link in the Figure 11 is set to four, one more than the cost of the established FA TE link which has the same number of the underlying optical links as the virtual link.

- Finally, the TE metric used for a physical (optical) link is set to 1. Additionally, every change of the layer computed by the routing algorithm is "punished" with plus one in the total path cost in order to favor the usage of FA LSP over a new LSP establishment (i.e., *metric (new_LSC_LSP) = number_of_optical_links + 2*). From the Figure 11, it can be seen that, for a new LSC LSP establishment between the nodes A and D, three optical links will take a part in the route and the layer is changed twice, so the metric is 3 + 2 = 5.

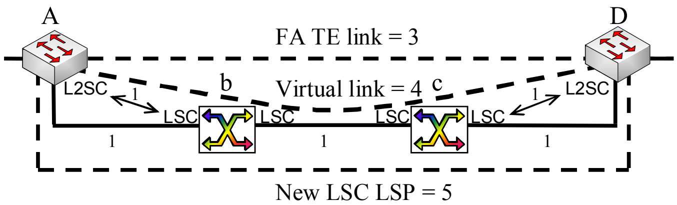


Figure 11. TE link cost

Apart from the TE metrics, other constraints are taken into account. The algorithm always tries to satisfy the WCC in the optical segments, but if this is not possible, it will check if there are enough wavelength converters at the intermediate OXC. The WCC is assured through the usage of RSVP-TE LabelSet object during the signaling. If there are two paths with the same cost between a source and a destination node, the algorithm will always choose the path which minimizes crossing the layer (i.e., new LSC LSP creation). Therefore, the algorithm prioritizes the L2SC establishment over multiple FA TE links and / or virtual links, checking firstly if there are available resources (e.g., unused bandwidth) to accommodate the request. Consequently, if there are two paths with the same cost between the edge nodes, the algorithm will favor FA TE links over virtual links.

Finally, taking into account the explained link cost / metric, the minimum cost path in the network is computed using Dijkstra-based algorithm.

3.6.2 Evaluation of the MLN approaches

For the evaluation of the VNT approaches using the proposed algorithm, we have used the topology shown in the Figure 12. The network is constituted by 20 *single switching capable* nodes (i.e., eight CO Ethernet switches and twelve OXCs). Like in the Section 3.5, we assume that the Ethernet nodes have full grooming capability. Each OXC has two wavelengths converters. The nodes are connected by bidirectional optical links multiplexing 10 wavelength channels.

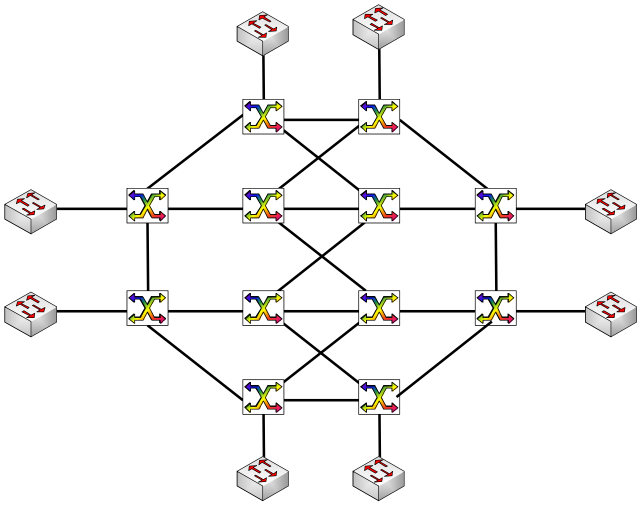


Figure 12. Network topology

The LSP requests arrive at the network following a Poisson distribution with the average inter-arrival time of 1s, with the requested bandwidth of 100 Mb/s. The source and the destination node are randomly chosen among the Ethernet nodes. The LSPs have an exponentially distributed holding time which varies in the different simulation set, in order to evaluate the behavior under various traffic loads. We assume that the processing time needed for path computation is 20 ms. The wavelength channel bit rate is 2.5 Gb/s and the link propagation delay is the same for each network link and set to 3.5 ms. The data points are obtained simulating around 7×10^5 L2SC LSP requests.

Each Ethernet node has either a pre-computed (for virtual approach) or pre-established (for semi-dynamic approach) unidirectional TE link with other four randomly chosen CO Ethernet switches. The TE links are pre-computed / pre-established before any signaling request from the upper-layer using Dijkstra-based algorithm and, in case of the semi-dynamic scenario, they are signaled following standard GMPLS signaling procedures.

3.6.3 Numerical evaluation

The network performance is evaluated considering the connection blocking probability and the end-to-end average setup delay as the key performance indicators for comparing the three VNT configuration approaches.

Blocking probability

The connection blocking probability as a function of traffic load is shown in the Figure 13. From the figure, we can observe that with the semi-dynamic policy we get significantly higher blocking probability compared to the other two policies. That is because in the semi-dynamic approach the set of pre-established FA LSPs may not be optimal, and the occupied resources at the optical layer may be used for the establishment of other LSC LSPs which would eventually result more efficient. In a fully dynamic scenario, the blocking probability is the lowest, since new LSC LSPs can be established whenever the existing FA LSPs are exhausted, and there are enough unused resources at the optical layer (i.e. wavelength channels). The blocking probability in the dynamic approach outperforms the blocking probability in the semi-dynamic approach up to 95% at the low traffic load. Increasing the traffic load, the difference is decreasing.

The virtual approach outperforms the semi-dynamic one, but its blocking probability is still higher than in the dynamic approach. The reason behind this is that in the semi-dynamic approach the pre-established FA LSPs are occupying the resources at the optical layer which leads to sub-optimal VNT configuration. On the other hand, in the virtual approach, the FA LSPs are pre-computed but not established. Therefore, the resources are not occupied nor blocked due to a specific set of LSPs, which does not constraint the computation and establishment of future LSPs. However, compared to the dynamic approach, the virtual approach worsens the blocking probability. This is due to the fact that in the virtual approach, when a virtual TE link needs to be established, the considered route is the one that is statically pre-computed. It is very likely that this route is not the optimal one with regard to the network state, which may lead to waste of network resources. Moreover, if, for a LSP request, there is a direct virtual link that connects the source and the destination nodes, but the LSP is not established yet, the signaling messages used for the establishment of this virtual link will be triggered following the pre-computed route, even though there are no available resources at the optical layer through that path. Consequently, the signaling request will be failed, even if there was a feasible and alternative route to reach the destination. However, it is worth noting that the edge node will firstly check if there is an already established FA LSP between the source and the destination node with enough resources.

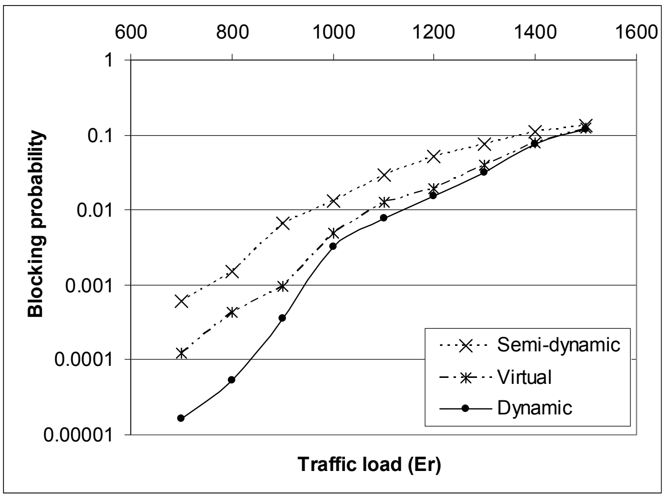


Figure 13. Connection blocking probability

As the traffic load increases, the blocking probability also increases in all three approaches, since the optical layer resources are scarce and it is more difficult to accommodate the L2SC LSP requests. Indeed, at some point, the blocking probability of all three approaches will become equal. This is due to the fact that at the high traffic loads the resources are less available, so the path computation algorithm cannot find the route and the destination cannot be reached, no matter if there were pre-established LSC LSPs or not.

Setup delay

In the Figure 14, it is shown the average setup delay as a function of the traffic load. The lowest setup delay is obtained by the semi-dynamic approach. Indeed, in such an approach, the set of LSC LSPs are pre-established and not released if there are no upper-layer LSP that uses the pre-defined FAs. Therefore, the time needed to set up such LSC LSPs is saved when establishing the L2SC LSP. On the other hand, in the dynamic approach, the FA LSPs are set up and torn down dynamically to accommodate the L2SC connections. Consequently, the VNT configuration changes more frequently. Therefore, comparing the setup delay of these three approaches, the worst performance (i.e., higher setup delay) is obtained by the dynamic approach. Finally, the setup delay obtained by the virtual approach is lower than the dynamic approach, but still higher compared to the semi-dynamic one. That is because, in the virtual approach, for a pre-defined virtual link there is no need for the executing path computation if the link directly connects the source and the destination nodes. Indeed this route is pre-computed, and thus, this allows outperforming (i.e., lowering) setup delay obtained by the dynamic approach. On the other hand, those virtual links still need to be signaled. This time is not consumed by the semi-dynamic case since the pre-defined links are already established. Note that, although the differences among these three setup delays are very small, it can still reflect the advantage of using pre-established and / or pre-computed LSC LSPs.

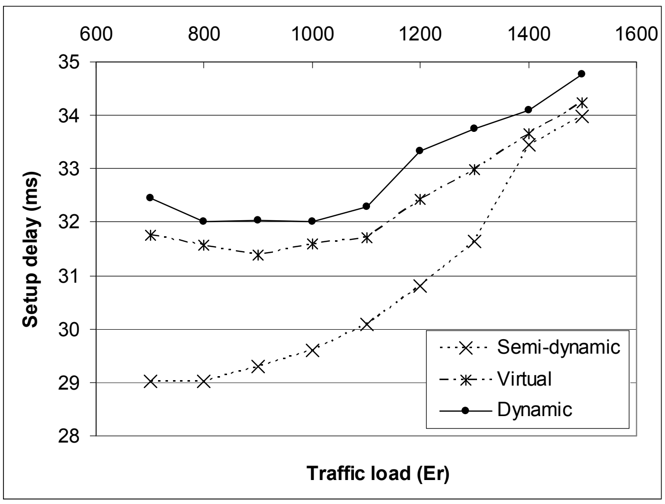


Figure 14. Setup delay

By increasing the traffic load, the setup delay is also increased. This happens because when traffic load is high, many connection requests are blocked due to the lack of available resources, and those ones that actually succeed, tend to be traversed through path entailing a large number of hops, which in turn does increase the average setup delay. Although it cannot be observed from the Figure 14, it is expected that at some point the setup delays of all the three approaches converge. That is because, at high traffic loads, the FA TE links last longer, so the network for all three policies will converge to the fully static network.

3.6.4 Conclusion: Path computation algorithm for MLN

The main goal of this contribution is to evaluate three different approaches for the dynamic VNT configuration, where both, FA TE links and virtual links, were taken into account. Moreover, we have devised a path computation algorithm that deals with both FA TE / virtual links, considering different TE link metrics. The algorithm favors the usage of FA LSPs and virtual links over a new LSC LSP establishment. We have shown that the dynamic approach lowers the connection blocking probability due to its better usage of the network resources when compared to both the virtual and the semi-dynamic approaches. However, this is attained at the expenses of increasing the average setup delay. On the other hand, the semi-dynamic approach performs the worst in terms of the blocking probability, due to not optimal pre-established FA TE links which are reserving / occupying resources that may not be used. Nevertheless, the obtained setup delay by the semi-dynamic approach outperforms the obtained by the virtual and the dynamic approach. Finally, given that the differences in the setup delay are relatively minor, it seems reasonable to focus on the pure dynamic approach. However, in real scenarios it is expected that network operators combine the dynamic policy with FA TE link and virtual links, to provide a network that is more stable. In light of these results, we conclude that the challenge is on choosing one of the three approaches for VNT configuration, taking into account the trade-off between the connection blocking probability and the average setup delay.

3.7 Enhanced path computation algorithm for dynamic VNT reconfiguration

The goal of a path computation algorithm in a MLN is to attain an optimal use of the network resources providing efficient cooperation and MLN TE decisions. The idea behind this is to devise a dynamic MLN path-computation algorithm which would aim at balancing the network resource utilization at all the involved layers through exploiting both the virtual and FA TE link concepts. The proposed algorithm is an extension of the strategy evaluating the three different VNT reconfiguration approaches explained in the previous Section 3.6. It has been extended several times during the thesis, and, herein, we present its final version. Specifically, the proposed routing algorithm uses a cost function which depends on the state and the TE attributes of the network elements (i.e., nodes and links). Moreover, it favors and promotes when possible the re-use of both virtual and FA TE links over triggering a new LSC LSP. The algorithm is evaluated in all three approaches for a VNT configuration: dynamic, semi-dynamic, and virtual. We should remind that, apart from pre-defined virtual links, it is possible to dynamically establish new FA LSPs in all three approaches. Our work is compared with the algorithms proposed in [17] and [30], and with the algorithm used for the evaluation of the VNT configuration approaches explained in the Section 3.6.

3.7.1 Path computation algorithm

Similarly to the algorithm explained in the Section 3.6, we assume a source-based path computation, although the same approach can be done in a dedicated PCE that has visibility of both layers. When the source node receives an incoming L2SC LSP request, according to the TED information, like in [17], it first checks if a direct FA TE link between the source and the destination nodes with enough unreserved bandwidth exists. If not, it checks if there is a pre-computed but not established virtual link which directly connects the edge nodes. If such a non-active virtual link exists, the source node starts the signaling for the establishment of the corresponding LSC LSP following the associated pre-computed route. Once the LSC LSP is established, the L2SC LSP will be routed over the newly established FA TE link. Conversely, if there is neither a direct FA TE link nor a virtual link that directly connects the source and the destination nodes, a path computation algorithm is executed. The algorithm relies on computing the shortest path cost between the pair of source–destination nodes. Its objective function aims at minimizing the cost, efficiently using available network resources. The input information of the algorithm is stored in the TED, and is as follows:

- Switching capability of the TE link interfaces.
- Number of underlying (lower-layer) links, for both virtual links and FA TE links.
- Availability of the network resources (i.e., wavelength converters, wavelength channels, unreserved bandwidth).
- Network resources utilization (i.e., load of the network TE links and the ratio of the used and the total wavelength converters at the given node).

3.7.2 Pseudo-code for the proposed algorithm

In order to explain the proposed algorithm, we use the following notation:

E: Set of the physical links in the network

FA: Set of all currently established FA TE links

VL: Set of all pre-computed virtual links

$E' = (E \cup FA \cup VL)$

$G = (V, E)$: Network graph G, with the set of nodes V and the set of links E'

s: Source node

d: Destination node

P: Path from the source s to the destination d

i: Any node from the set V

$e = (u,v)$: Any link $e \in E'$ between any two nodes u, $v \in V$

req_band: Requested bandwidth

dist[i]: Array which contains the distance from s to i

pred[i]: Array which contains the predecessor nodes of i

Q: Set of all the adjacent nodes that are not yet visited

unreserved_band(e): Unreserved bandwidth of any link $e \in E'$. If $e \in E$ and the physical link is between two OXCs, we consider that the unreserved bandwidth is equal to the number of available wavelengths multiplied by the rate of a wavelength. We assume that the number of ports at Ethernet nodes p is equal to the number of wavelengths W within an optical link. In other words, CO Ethernet nodes can handle as many wavelengths as an OXC-OXC link supports. Therefore, the total capacity of an Ethernet-OXC node is equal to the total capacity of an OXC-OXC link.

WC(u): Number of wavelength converters at node u

WCC: Wavelength continuity constraint

C(u): cost of any node $u \in V$

$C(e_{uv})$: cost of any link $e \in E'$ between the nodes u and v, where u, $v \in V$

C(P): cost of the path P

The proposed path computation algorithm calculates the route with the minimum cost, denoted *C(P)*, according to the following cost function:

$$\mathrm{C}(P) = \mathrm{C(u)} + \mathrm{C}(\mathrm{e}_{uv}) \quad (1)$$

The node and the link cost are defined as:

$$\mathrm{C(u)} = \sum_{u} \alpha a_{u} \quad (2)$$

$$\mathrm{C}(\mathrm{e}_{uv}) = \sum_{uv} (\beta b_{uv} + \chi c_{uv} + \delta d_{uv}) \quad (3)$$

The coefficients α, β, γ and δ are used to adjust the values of the parameters a, b, c, and d, and the parameters are defined as follows:

- Parameter a specifies the ratio between the number of wavelength converters (WCs) used and the total number of WCs equipped at a

WSON (OXC) node *u*. By doing so, highly congested optical nodes with scarce available WCs tend to be avoided.

- Parameter *b* represents the current traffic load in the network TE link. For the physical optical links, the traffic load is defined as the ratio between the number of occupied wavelength channels and the total multiplexed wavelength channels at the link. For the established FA TE links, this value presents the ratio between the bandwidth used and the maximum reservable bandwidth of the link. For pre-computed but not established virtual links, parameter b is equal to zero.
- Parameter *c* depends on the switching type of the TE link. That is, if both ends of a TE link support the same switching capability (e.g., L2SC), the parameter is equal to zero. Otherwise (i.e., a TE link has at one end L2SC and at the other end LSC), *c* is set to one. By doing so, every change of layer is penalized. In a dual-layer network, the change of layers traversed by a path always occurs in pairs (i.e., from L2SC to LSC and from LSC to L2SC). Thus, if γ is set to 1, the path cost due to layer changes is increased by two.
- Parameter *d* reflects a number of physical hops of a TE link. Therefore, for each optical link it is equal to one. For established FA TE links, it represents the number of underlying optical links. And, finally, in the case of virtual links, it is equal to the number of underlying optical links plus one. By doing so, we favor the usage of already established FA TE links over pre-computed but not established virtual links. Moreover, if the parameters γ and δ are set to one, the algorithm will rather choose either an FA TE link or a virtual link than setting up a new LSC LSP which would have the same number of hops as the existing FA / virtual link. Nevertheless, for values of α and δ being different from zero, the current state of the node resources and the links will be taken into account while choosing whether to establish a new LSC LSP or not.

Finally, taking into account the explained metric, the minimum-cost path in the network is computed using a Dijkstra-based algorithm, and its pseudo-code is shown in Figure 15. Bear in mind that, during the path computation, apart from the cost function, other network constraints are taken into account. Links that are fully occupied (i.e., no sufficient available bandwidth) are excluded from the path computation. The path computation always tends to satisfy the WCC in the optical segments, which is then ensured by the RSVP-TE protocol using the LABEL_SET Object. However, if this is not possible, it will check if there is an available wavelength converter at a given node. If, during the path computation, there are two paths with the same cost between the source and the destination nodes, the algorithm will always choose the path which minimizes crossing the layer (i.e., a new LSC LSP creation), accommodating the request over multiple FA TE and / or virtual links. Consequently, if there are two paths with the same cost between the edge nodes, the algorithm will favor FA LSPs over virtual links.

```
INPUT
     G, V, E, FA, E', VL, s, d, req_band
OUTPUT
     P
Shortest path based algorithm
     for ∀i | i ∊ V
        dist[i] ← INFINITY
        pred[i] ← NULL
     dist[s] ← 0
     Q ← s
     While Q is not empty
        u ← extract min dist (Q)
        for ∀e,v | e=(u, v) ∊ E'
          if unreserved_band(e)<req_band
             continue
          if e ∊ E
             if WCC is not satisfied and WC(u)==0
                continue
          new_dist ← dist[u] + C(u) + C(euv)
          if new_dist < dist(v)
             dist[v] ← new_dist
             pred[v] ← u
             Q ← v
     end while

     u ← d
     while pred[u] is not NULL
        insert u at the beginning of P
        u ←  pred[u]
     end while
     return P
```

Figure 15. Pseudo code of path computation algorithm

With the proposed algorithm, the virtual topology is adjusted dynamically to suit for traffic-demand changes. The main advantages of the proposed algorithm are as follows:

- It is a simple path computation algorithm, based on a calculation of the shortest path.
- By introducing the parameters *c* and *d*, paths with minimal physical hops have advantages, avoiding the possibility that two FA TE links in a path share the same optical link.
- Congested links and nodes are avoided by introducing the parameters *a* and *b*.
- By adjusting the coefficients α, β, γ and δ, the influence of the parameters *a*, *b*, *c*, and *d* on the cost function can be manipulated, in order to satisfy different constraints.
- It is possible to dynamically change the layer several times along the path as long as the switching capability constraints are satisfied.

3.7.3 Performance evaluation

We compare the proposed algorithm with the three following algorithms found in the literature:

- The first algorithm presents the *Policy 1* proposed in [17] explained in the Section 3.4. Since the algorithm first tries to accommodate the request over multiple FA TE links, for the sake of simplicity, we call this algorithm the upper-layer-first algorithm (ULFA). Recall that in this algorithm, there is no changing of the layers, but the requests are accommodated at either the upper layer (i.e., L2SC) or the lower layer (i.e., LSC).
- The second algorithm is proposed in [30]. The algorithm chooses the path that has minimal hops in the optical layer, and the authors called the algorithm *min_phys_hop*, or MPH.
- The third algorithm is similar to the algorithm we used in Section 3.6. Specifically, the cost of an FA TE link is set to the number of minimal physical hops, while pre-computed but not established virtual links have a metric equal to the number of underlying optical links plus one. Finally, when a new LSC LSP is triggered, every optical link has cost equal to one, and each change of the layer is punished with plus one. The algorithm extends MPH so it supports not only FA TE links but also virtual links and, thus, we call it the extended minimal physical hops algorithm, or EMPH.

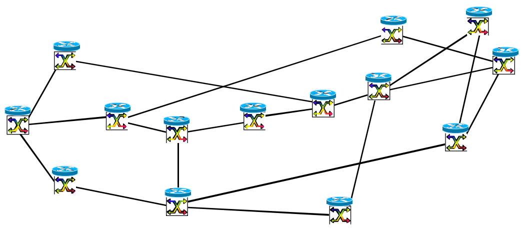
In all three algorithms, ULFA, MPH, and EMPH, the WCC must be assured, and when this is not possible, the algorithms will check if there is a wavelength converter at the given optical node. Finally, for the sake of simplicity, we call the proposed algorithm integrated multi-layer path computation, or IMLPC.

In order to evaluate the performance of the proposed algorithm, we take into account two network topologies: NSFNET [42] and the basic reference topology for the Pan-European network [43]. The algorithms are compared using the three different VNT configuration approaches explained in the Section 3.6

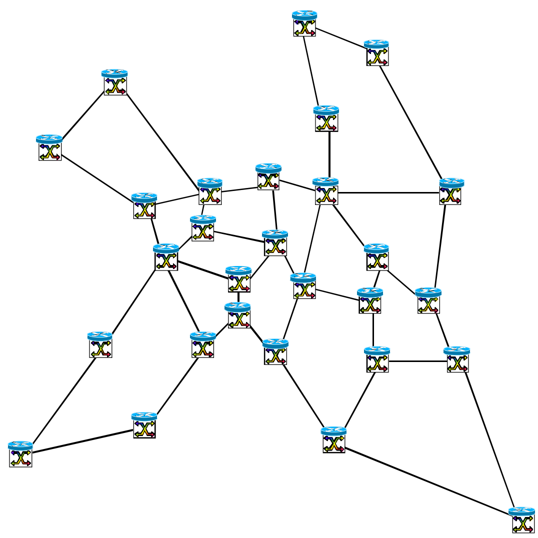
3.7.4 Simulation parameters

All the nodes in the network are single switching capable nodes, which can be either CO-Ethernet (L2SC) switches with full grooming capabilities or optical nodes (i.e., LSC OXC) with a limited number of wavelength converters. The NSFNET topology is constituted by 14 OXC nodes, with one L2SC Ethernet node attached to each OXC (there are 28 nodes in total), where the distance between them is negligible. The Pan-European topology is constituted by 28 OXC nodes, and, similar to NSFNET topology, with one L2SC switch node connected to each OXC node. For the sake of simplicity, in Figure 16, only OXC nodes are depicted with a corresponding GMPLS controller. Every optical node has a fixed number of available wavelength converters: 4 per node in the NSFNET network and 15 per node in the Pan-European network. Every optical link is bidirectional, and each of them supports 10 wavelengths in the NSFNET network and 20 wavelengths in the Pan-European network. The rate of every wavelength channel is 2.5 Gb/s, while the requested bandwidth for every upper-layer LSC is 100 Mb/s. The link propagation delay is assumed equal for every control link, 3.5 ms. As mentioned before, for each network topology (i.e., NSFNET and Pan-European) there are considered three scenarios / approaches: dynamic, semi-dynamic and virtual.

The L2SC switch nodes generate connection requests dynamically, following a Poisson arrival process with an average inter-arrival time of 1s. Destinations are chosen uniformly, among the rest of the L2SC nodes. The LSP holding time is exponentially distributed and varies in different simulation sets from 2400 s to 3000 s in the NSFNET network, and from 6000 s to 9000 s in the Pan-European network. In case of the semi-dynamic and the virtual scenarios, each L2SC node has two pre-established / pre-computed FA / virtual TE links in the NSFNET network, and four in the Pan-European network, with a randomly chosen destination node. All the LSPs, and thus, the virtual and FA TE links in the network are unidirectional. Each data point is obtained simulating around 7.5×10^5 L2SC LSP requests.



a)



b)

Figure 16. Used network topologies: a) NSFNET b) Pan-European

For the sake of simplicity, in the proposed algorithm, we set the values for all the coefficients equal to one; that is $\alpha = \beta = \gamma = \delta = 1$. Nevertheless, the values of the parameters can vary in order to stress different constraints, e.g., to avoid using highly congested OXCs, optical links, etc. In that sense, depending on the network topology and the resources, by adjusting the values of the coefficients we can achieve a good balance in the usage of the network resources. For example, for the lower values of the ratio of the number of optical link wavelength channels over the data rate of a wavelength channel, it is more appropriate to avoid the creation of a new FA TE link. In other words, since the rate of a channel is high, it is better to accommodate packet requests over the existing FA TE links than to trigger the establishment of a new LSC LSP, using optical resources less efficiently. Therefore, a high value of γ may be required to achieve this goal, favoring, thus, the best utilization of existing FA TE links. Similarly, in a translucent network with few wavelength regenerators per node, it may be better to set a higher value of α, in order to reduce their usage.

The proposed algorithm can provide a good balance among the network TE links utilization by adjusting the coefficient β. The higher the value of the β coefficient, the smaller the probability that a very congested TE link will be taken into account for the path computation. Finally, recall that the proposed algorithm integrates both layers, which may lead to dynamically changing the layer several times along the path as long as the switching adaptation constraints are satisfied.

3.7.5 Results

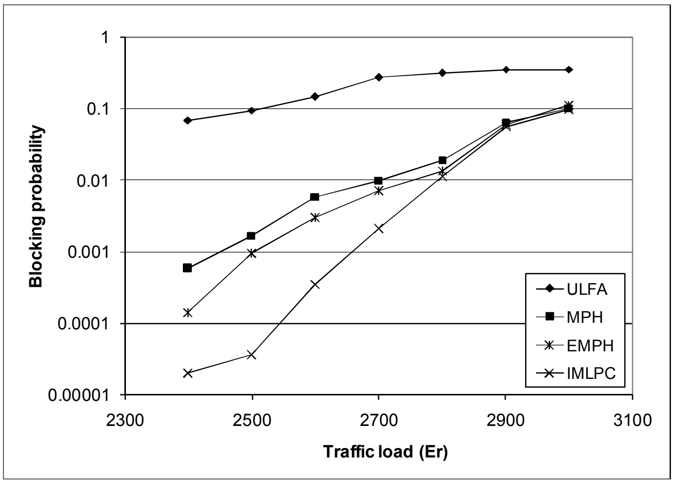
In order to evaluate the performance of the proposed MLN path-computation algorithm, the following figures of merit are used to evaluate the network performance:

- Connection blocking probability
- Average number of active wavelength channels per optical link
- FA TE link usage: The average load of an FA TE link during its lifetime
- Average number of hops at the upper layer

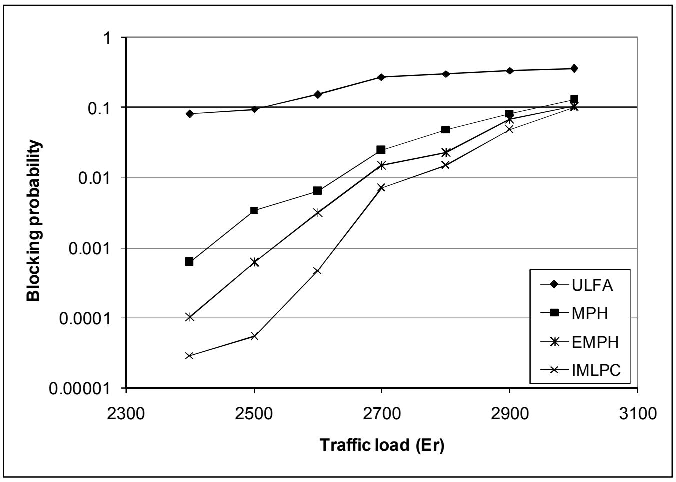
Blocking probability

Figure 17 shows the connection blocking probability as a function of the traffic load (expressed in Erlangs, Er) in an NSFNET network, using the (a) dynamic, (b) semi-dynamic, and (c) virtual approach.

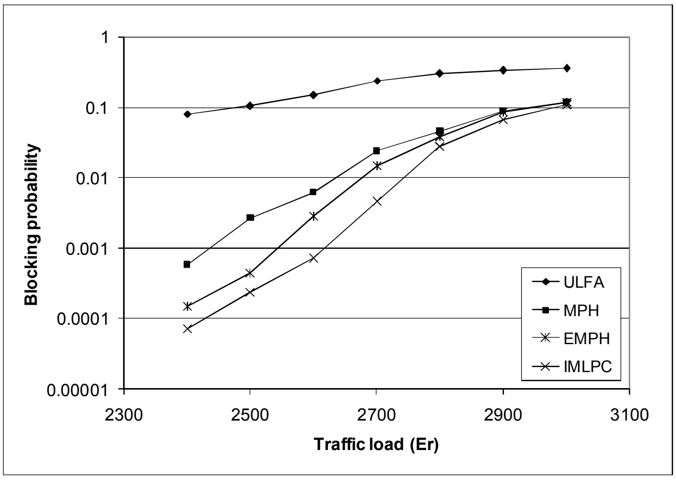
From the figure, we can observe that the proposed IMLPC algorithm attains the lowest blocking probability in all three scenarios. That is because the algorithm includes all the considered node and link states and attributes, and thus, chooses a path with less loaded optical / FA TE links along with the nodes having more available WCs than in the other three algorithms. The ULFA performs the worst, since it tries to accommodate all requests first at the upper layer, restricting the route computation. In other words, this algorithm does not allow a path to change the layer at intermediate nodes, and thus, no collaboration between the layers exists, causing the worst blocking probability.



a) dynamic



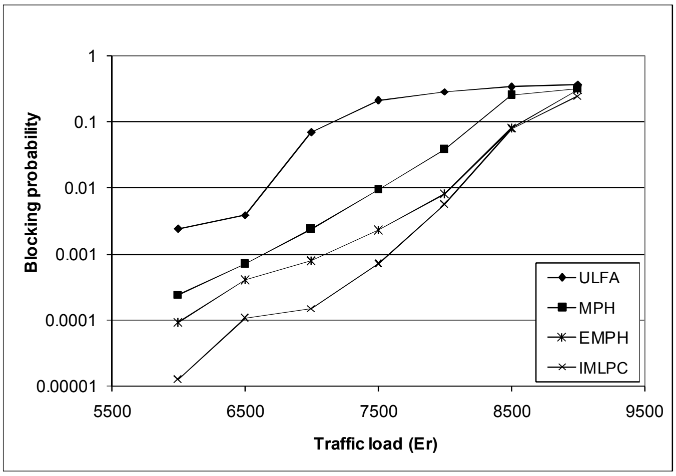
b) semi-dynamic



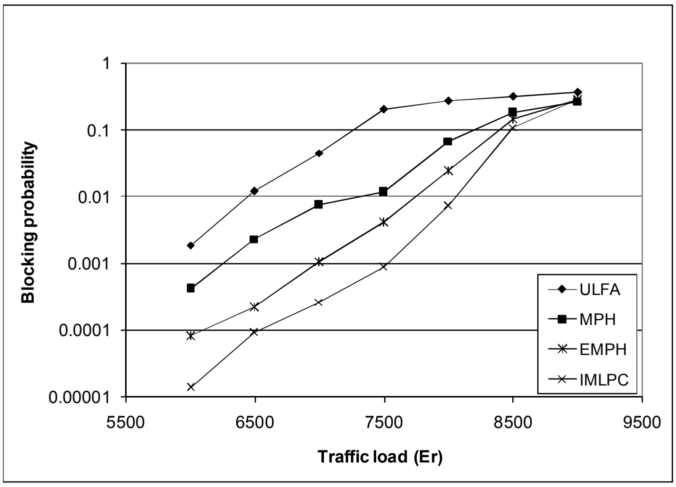
c) virtual

Figure 17. Connection blocking probability in NSFNET topology

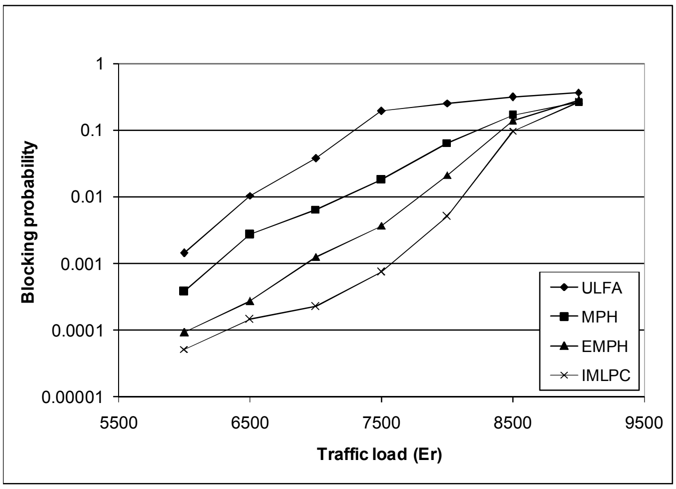
Finally, the EMPH algorithm outperforms the MPH algorithm. In the MPH algorithm, the path cost for the establishment of a new LSC LSP is the same as the cost of an already established FA TE link whose underlying LSP traverses the same number of hops. On the other hand, the EMPH algorithm "punishes" every change of the layer, and thus, it tends to reuse an already established FA TE link with up to two hops more than a new LSC LSP. Therefore, the EMPH algorithm aims at saving the resources at the optical layer at the expense of exploiting the reuse of already established FA TE links. Since in our scenarios the nodes are equipped with a limited number of wavelength converters, with the EMPH algorithm we have obtained better connection-blocking probability than when the MPH algorithm is used.



a) dynamic



b) semi-dynamic



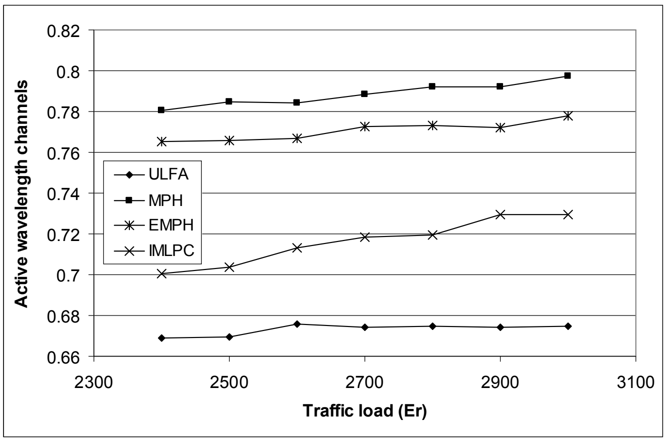
c) virtual

Figure 18. Connection blocking probability in Pan-European topology

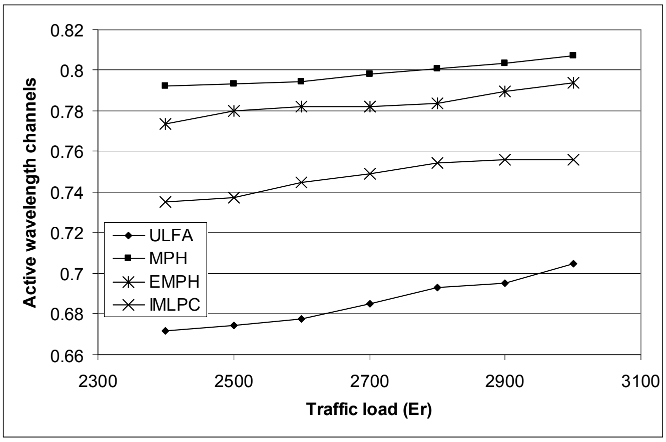
We obtain similar results in the Pan-European topology (Figure 18). We can observe that the IMLPC algorithm also outperforms the ULFA and the MPH and EMPH algorithms in all three scenarios. Nevertheless, the connection blocking probability in all three scenarios is increasing faster in the NSFNET topology than in the Pan-European topology (note that in the NSFNET network the traffic load is increasing by 100 Er, while in the Pan-European network it is increasing by 500 Er). That is because, since the Pan-European network is bigger, it has more resources, and thus, it needs higher traffic load, so the connection-blocking probability of the algorithms converges slower than in NSFNET topology, in all three scenarios.

Average number of active wavelength channels per optical link

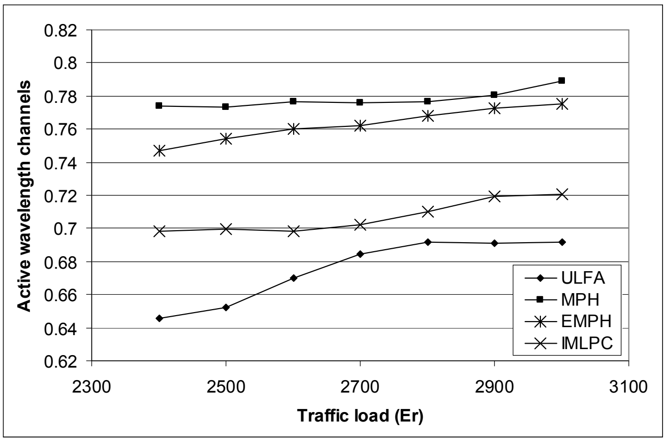
Figure 19 depicts the average usage of wavelength channels per optical link, scaled from 0 to 1, using the (a) dynamic, (b) semi-dynamic, and (c) virtual approach, in the NSFNET topology.



a) dynamic



b) semi-dynamic



c) virtual

Figure 19. Active wavelength channels in optical links, NSFNET topology

From the figures, we can observe that the ULFA uses the least wavelength channels compared to the other algorithms, in all three scenarios. This is because this algorithm always tries to accommodate a new request over multiple existing FA TE links, if no direct FA TE link between the source node and the destination node actually exists. Therefore, this algorithm triggers the least number of LSC LSPs. The IMLPC algorithm uses fewer wavelength channels than the MPH and EMPH algorithms. This is because these two algorithms do not take into account the congestion of the network resources, but the main criterion for L2SC LSP establishment is the number of hops (i.e., MPH and EMPH) and if the layer has been changed (i.e., EMPH). On the other hand, the IMLPC algorithm tries to avoid highly congested links (both optical and FA TE links) and the usage of wavelength converters. By limiting the usage of wavelength converters, we are reducing crossing the layers more than when the MPH and EMPH algorithms are deployed. Therefore, the IMPLC algorithm triggers fewer LSC LSPs than the MPH and EMPH algorithms, and thus, the optical physical links are less occupied. That is why the IMPLC algorithm has fewer wavelength channels used than the other two algorithms. Finally, the EMPH algorithm uses fewer wavelengths than the MPH algorithm, because the

EMPH algorithm punishes every incurred change of layer, while the MPH algorithm chooses a path with the minimal number of physical hops.

Comparing these three scenarios, we can see that in the semi-dynamic scenario the optical resources are still slightly more used than in the other two scenarios, with all four algorithms. The reason behind this is, as explained, that in the semi-dynamic scenario there are pre-established FA TE links occupying optical resources that might not be optimal. Therefore, although there are pre-established FA TE links, sometimes it is more efficient to establish a new one than to use the existing ones, resulting in more FA TE links being established, and thus, having more wavelength channels occupied.

The average number of active wavelength channels in optical links in the Pan-European network is depicted in Figure 20. Since we obtain similar results in the cases of the semi-dynamic and virtual approaches, only the results obtained when the dynamic scenario is deployed are depicted.

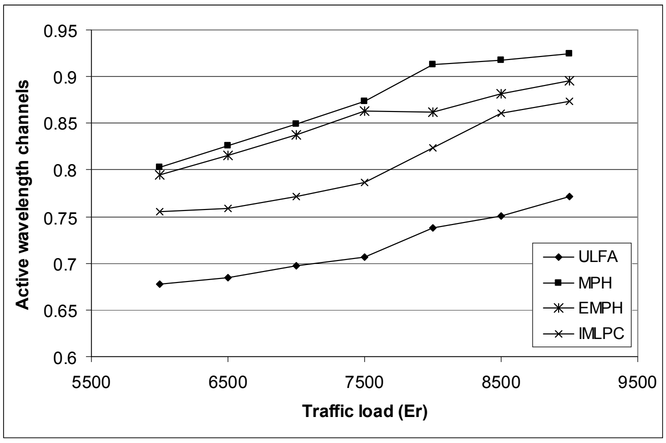


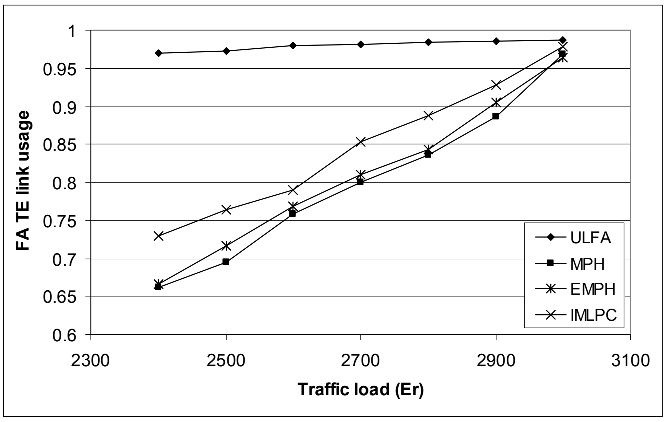
Figure 20. Active wavelength channels in optical links, Pan-European topology-dynamic scenario

FA TE link usage: The average load of an FA TE link during its lifetime

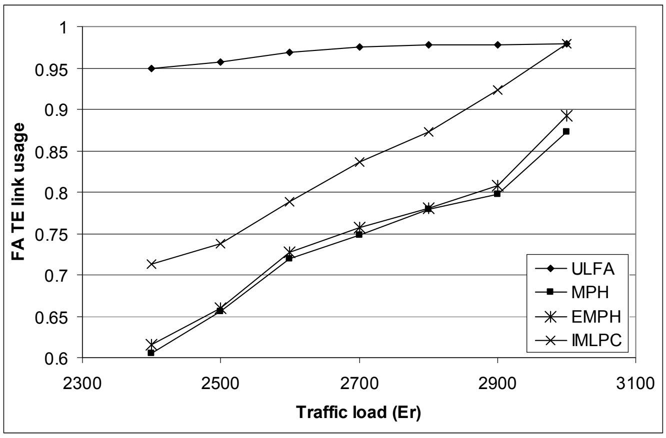
Figure 21 shows the average usage of FA TE links as a function of traffic load in the NSFNET topology having (a) dynamic, (b) semi-dynamic, and (c) virtual VNT configuration.

Since the ULFA tends to accommodate requests first over the upper layer, it is the algorithm that leads to use FA TE links the most. Indeed, the algorithm uses almost 100% of the FA TE link capacity (from ~95% to 98.5%, in all three scenarios). On the other hand, both the MPH and the EMPH algorithms barely use the FA TE link capacity. Observe that the difference between them is relatively low. In particular, using the MPH algorithm, FA TE links are less used than in the EMPH strategy, since the former takes into account only the number of physical hops. Therefore, the chances to accommodate an LSP request over an already established FA TE link with the same number of hops as a new LSC LSP establishment is the same. However, the

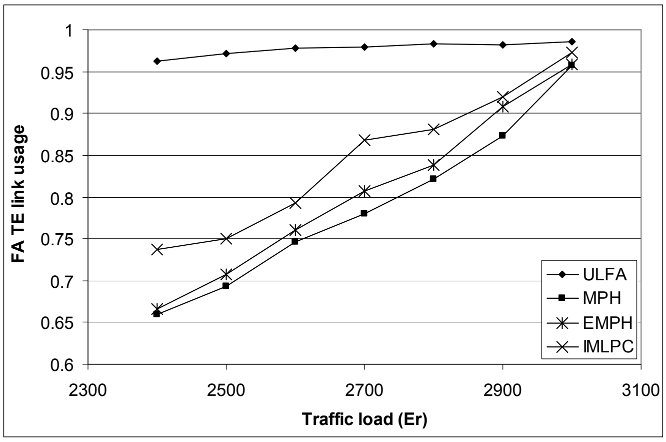
EMPH algorithm punishes every change of layer, which leads to slightly higher usage of the established FA TE links. Finally, the IMLPC algorithm uses more FA TE links with respect to the EMPH and MPH algorithms, since the IMPLC algorithm aims at avoiding congested optical links and nodes along with punishing the change of layer. Consequently, the utilization of the established FA TE link is increased.



a) dynamic



b) semi-dynamic



c) virtual

Figure 21. FA TE link usage, NSFNET topology

Comparing these three scenarios, one can observe that, in the semi-dynamic scenario, the established FA TE links are, on average, less used. This is because, as is shown in Figure 21, in the semi-dynamic scenario, there are more wavelength channels used than in the other two scenarios, constituting more FA TE links. Consequently, the established FA TE links are slightly less used.

Figure 22 depicts the FA TE link usage as a function of traffic load in the Pan-European network, having the dynamic scenario.

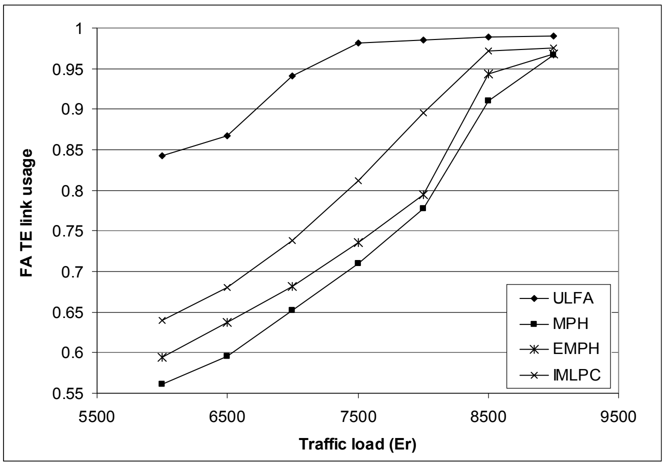


Figure 22. FA TE link usage, Pan-European topology

The figure confirms the result obtained in Figure 21. However, we can observe that the FA TE links are less used in low traffic load. Specifically, the Pan-European network is a bigger network with a huge amount of resources. Consequently, the utilization of FA TE links becomes significant when the network resources start to be exhausted, and this occurs at high traffic load.

Average number of hops at the upper layer

Figure 23 shows the average number of hops for L2SC LSP establishment as a function of traffic load in the NSFNET network. We will show the results obtained only deploying the dynamic VNT configuration scenario, since we get similar results for the other two scenarios. For this statistic, we take into account the average number of hops while establishing only L2SC LSPs. Therefore, the number of hops during LSC LSP establishment is not taken into account.

From Figure 23 we can observe that the ULFA chooses the longest paths. This is because, if there is not a direct FA TE link with enough resources between the source and the destination nodes, this algorithm tries to route a path over multiple established FA TE links. If this is not possible, it computes a direct LSC LSP. As is expected, the MPH and EMPH algorithms compute paths with the least number of hops, since they rely on this metric. Finally, the proposed IMLPC algorithm computes slightly longer paths, since other constraints have been taken into account. However, the differences in the average number of hops of these three algorithms are insignificant.

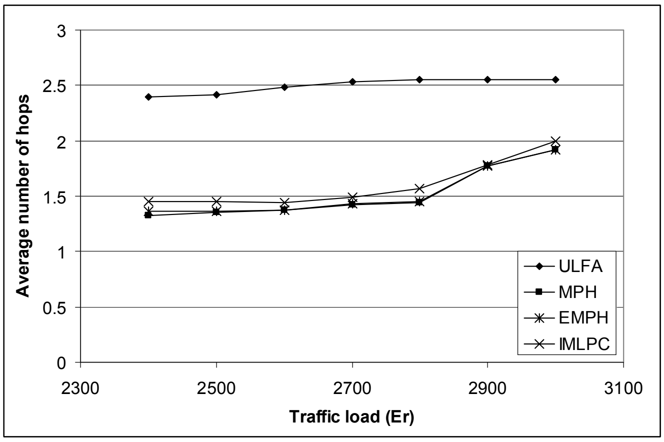


Figure 23. Average number of hops at the upper-layer, NSFNET topology-dynamic scenario

To validate the results obtained in the previous figure, Figure 24 shows the average number of hops in the Pan-European network. Although we have obtained similar results to those for the NSFNET topology, one still may observe that in the Pan-European network the paths are slightly longer. This is because the network is bigger, and it is more difficult to establish L2SC LSPs with a small number of hops. However, since the average number of hops directly reflects the delay for end-to-end establishment, we can conclude that deploying the ULFA algorithm requires the longest time for setting up L2SC LSPs.

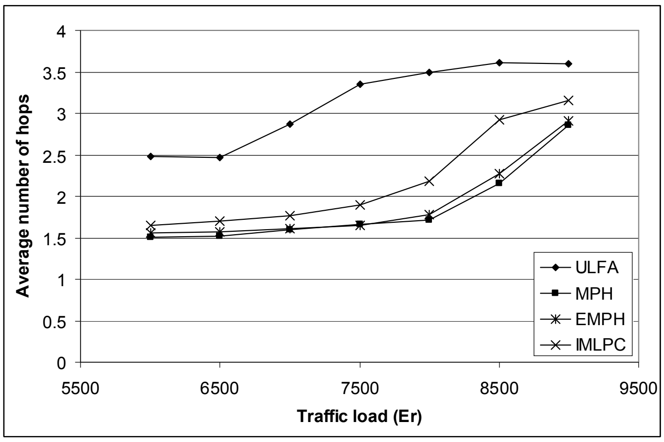


Figure 24. Average number of hops at the upper-layer, Pan-European topology-dynamic scenario

3.7.6 Conclusion: Enhanced path computation algorithm for dynamic VNT reconfiguration

The main goal behind this contribution was to devise an efficient, integrated MLN path-computation algorithm for dynamic VNT configuration which improves the network performance when compared with selected algorithms proposed in the literature. The proposed algorithm relies on calculating the shortest path using an adjustable cost function, in which both virtual and FA TE links are taken into account. Moreover, the algorithm considers various constraints such as the WCC, the availability of network resources (i.e., wavelength converters, wavelength channels, FA TE links), and their utilization. Adjusting the values of the cost function parameters, it trades off the number of hops and the congestion of the network resources. Deploying the proposed algorithm, it is possible to dynamically establish L2SC LSPs with multiple layer changes, including virtual and FA TE links, and triggering an LSC LSP whenever it is needed.

The algorithm is compared with three algorithms found in the literature, using two network topologies: NSFNET and Pan-European. It is shown that by deploying the proposed IMLPC algorithm the connection blocking probability is significantly reduced in all three scenarios for VNT configuration: dynamic, semi-dynamic and virtual. Moreover, it is demonstrated that it achieves almost the same network resource usage and it better exploits grooming decisions compared to the MPH and EMPH algorithms. Nevertheless, compared with the ULFA, the MPH and EMPH algorithms, the proposed algorithm performs better in terms of the connection blocking probability and the network resource usage, even with other network topologies and under different network configuration. This is because the ULFA leads to accommodate the requested connection over one or multiple FA TE links. This does waste the resources at the upper layer. The other two algorithms choose the route with the minimal number of physical hops, but this does not mean that the computed path is optimal. The proposed algorithm integrates both involved layers aiming at balancing their resources. Moreover, it is possible to attain different goals even when there is a different ratio of number of wavelength channels and wavelength channel rate, or when there are few / many wavelength converters per OXC. Finally, we can conclude that the proposed algorithm yields much more flexibility than the other three algorithms.

3.8 Chapter Summary and Conclusions

In this Chapter, we have discussed about the main issues of MLN combining both CO Ethernet over WSON and controlled by a GMPLS unified control plane. In particular, we concentrate on aspects such as the online path computation algorithm, the dynamic VNT reconfiguration and the multi-layer LSP provisioning. Then, we have reviewed the algorithms for VNT configuration found in the literature. Finally, we have presented three contributions of this PhD thesis dissertation regarding this topic.

As a first contribution, we have introduced the concept of *FA TE link timer*. The main idea behind this is to reduce or limit the systematic GMPLS signaling overhead caused by successive establishment / deletion of FA LSPs between a given pair of nodes. We have shown that, using the FA TE link timer, the total amount of the created FA LSP is reduced which, in turn, does decrease the amount of control signaling processing. However, we have shown that the timer has a minor influence on the average setup delay. Finally, we can conclude that the timer value should be chosen according to the trade-off between the setup delay and the connection blocking probability.

In the second contribution of this dissertation, it is presented a novel online path computation algorithm for a MLN where a VNT constituted by both virtual and FA TE links, can be dynamically reconfigured. The algorithm favors the usage of FA TE and virtual links over setting up new LSC LSP with the possibility to change the layer several times as long as switching adaptation constraints are satisfied. We have used the algorithm to evaluate three different approaches for online VNT reconfiguration: *dynamic*, *semi-dynamic* and *virtual*. It has been shown that the *dynamic* approach obtains the lowest connection blocking probability, but at the cost of the increased setup delay, due to dynamicity of FA TE link creation.

Finally, we have presented a new devised MLN path computation algorithm that, using an adjustable cost function, computes MLN routes aiming at achieving an efficient use of the resources in the all the involved layers. To this end, path computation algorithm takes into account the current availability of the network resources and various constraints (i.e., WCC, switching adaptation capability constraints, etc). Through the set of simulations, we have proved that the algorithm provides a good balance between the network resources, integrating both involved layers, and exploiting efficiently the grooming decisions.

Chapter 4

End-to-end Dedicated Protection in MLN

In the first part of this chapter, recovery mechanisms and schemes for MLN are reviewed. In general, they can be classified into protection and restoration scheme. In this dissertation we focus on end-to-end (e2e) protection schemes for MLN under the framework of GMPLS unified control plane. The main features and problems related to single-layer and multi-layer protection managed by GMPLS unified control plane are presented. Then, we present proposed protection schemes found in the literature. Finally, we present the contributions of this thesis regarding path protection in MLN.

4.1 Introduction

In a high-capacity transport network, a failure of one single network element (e.g., a fiber link) can affect several optical channels, leading to a great amount of data loss. Therefore, network survivability and recovery methods are the most critical issues in any core network. There are typically two approaches / schemes that deal with the network reliability: protection and restoration [44][45][46][47][48][49][50][51][52][53]. Protection is an approach based on computing, provisioning and reserving the resources for the backup (protecting) path, at the moment of setting up the working path [54][55][56]. Therefore, if a failure occurs, the traffic will be re-routed over the already pre-defined backup path. The protection approaches can be classified into dedicated and shared protection. Deploying the dedicated protection approach, the resources of the established backup path cannot be used by the working paths of other connections [57][58]. Furthermore, dedicated protection can be classified into 1+1 and 1:1 protection. In dedicated 1+1 protection, the data is being simultaneously transmitted through both, working and backup paths at the same time, and the receiver at the destination node decides from which path will take the data. In case of 1:1 protection, the resources are reserved for both, working and backup path, but the data will be transmitted over a backup path once the failure is detected. Deploying the shared protection approach, a protecting resource (e.g., wavelength channel) can be shared among the working paths of other connections in some specific conditions / cases (e.g., low priority connections), as long as respective working paths are link-disjoint [59][60][61].

The restoration approach is based on computing a backup path once a failure that affects the working path occurs [62][63][64]. The main advantage of restoration over protection is that the network resources are reserved only when it is needed. On the other hand, with the protection strategy, the resources are occupied even if a failure never occurs, precluding, thus, their efficient usage. Nevertheless, the advantage of using the protection approach is a very short Mean Time To Recovery (MTTR). In case of 1+1 protection, the MTTR is the smallest (i.e., <50ms), since the destination node will take the data from the backup path, as soon as a failure is detected. Even in case of 1:1 or shared protection, the MTTR is very low, since the backup path in both cases is pre-defined, before any failure in the network. On the other hand, the restoration approach needs significantly more time since a backup path is triggered and established upon failure is actually detected. Finally, note that both approaches, protection and restoration, can be applied also to a link / segment recovery, in which the traffic is re-routed only over a link / segment affected by a failure [65][66][67][68]. In this dissertation, we focus on the survivability problem in an integrated MLN (CO-Ethernet over WSON) using dedicated path protection.

The main problem in a MLN with dedicated path protection is how to choose, efficiently, a pair of working and backup paths for the requested LSP, targeting specific survivability requirements and objectives in terms of network resource utilization. A common requirement is that a single failure must not affect both paths. To this end, when computing a backup path, only links that are SRLG (Shared Risk Link Group) disjoint with the links used in the working path (working links) are taken into account [69]. The SRLG concept specifies a group of network elements (i.e., links) that share a common risk of failure. For the links belonging to the same SRLG group as the working links, we say that they are SRLG-joint, differentiating from the links that are SRLG-disjoint.

4.2 Challenges in GMPLS controlled MLN with e2e dedicated protection

In this section, we present some of the challenges in MLN dedicated path protection. The shared protection and restoration approaches are out of scope of this dissertation.

Inefficient usage of the network resources

One of the main issues in MLN with e2e dedicated protection is that the resources are occupied by a backup path even though a connection is never hit by a failure, leading, thus to inefficient usage of the network resources [57][58]. Therefore, the main challenge in a MLN with dedicated protection is to provide a recovery scheme which will optimize/enhance the usage of the network resources at the involved layers, when establishing both, working and backup paths. Moreover, it should be born in mind that both paths need to satisfy the same constraints required for that connection request (e.g., WCC, switching capability adaptation constraints, etc), which may create additional problems in a MLN/MRN.

Two-step Dijkstra vs. Suurballe algorithm

In order to compute two SRLG-disjoint paths, there are two main/reference routing algorithms: the two-step algorithm, based on the Dijkstra algorithm, and the Suurballe algorithm [70]. The two-step Dijkstra algorithm firstly computes a minimal-cost working path, and then it computes a minimal-cost backup path, once the working path SRLG-joint links are removed from the graph. The complexity of this algorithm is very low (i.e., $O(2n^2)$, where n represents a number of network nodes) [73], but it may fail in the so-called "trap" topology problem [71]. This problem specifies the situation where a solution for the computation of SRLG-disjoint working / backup paths exists, but the algorithm is unable to find out such a solution.

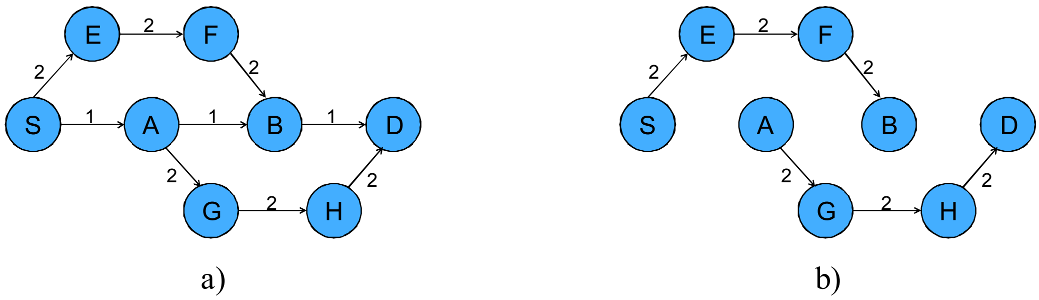


Figure 25. Trap topology

An example of trap topology is shown in the Figure 25. If there is a connection request between the nodes *S* and *D*, the two-step Dijkstra will select the route *S-A-B-D* for the working path, since it provides the shortest distance to the destination (Figure 25. a). Then, in order to compute the route for the backup path, all the working path SRLG-joint links will be removed from the network graph (i.e., the links *S-A*, *A-B* and *B-D*). Finally, after removing the links (Figure 25. b), the algorithm is unable to find the backup path to the destination node.

The trap topology problem is avoided using the Suurballe algorithm. The algorithm firstly finds the shortest path between the nodes *S* and *D* (Figure 26. a), *P1*. Then, it reverses the direction of the links of the path *P1* (Figure 26. b) and it finds the shortest path between the nodes *S* and *D*, *P2*, in the new topology, resulting the path: *S-E-F-B-A-G-H-D* (Figure 26. c). The algorithm will then remove all the common links between the *P1* and *P2* (i.e., link *A-B*, Figure 26. d). Finally, on the new topology (Figure 26. d), it will run the two-step Dijkstra algorithm to find two disjoint paths between the nodes *S* and *D*, resulting the paths: *S-E-F-B-D* and *S-A-G-H-D*.

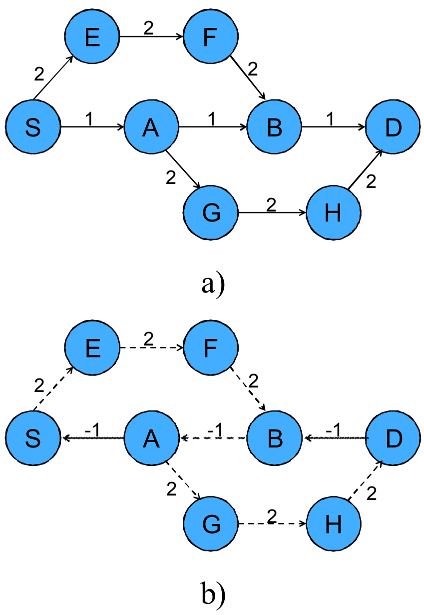


Figure 26. Suurballe algorithm

As it has been shown, the main advantage of using the Suurballe algorithm is that it avoids the trap topology, finding two disjoint paths even when the two-step Dijkstra algorithm is not able to find. However, the complexity of the Suurballe algorithm is higher than using the two-step Dijkstra algorithm, and it is $O(n^2 \log n)$ [72]. Finally, we should stress that, the Suurballe algorithm assures the minimal total cost of both paths (i.e., the sum of two costs). On the other hand, the two-step Dijkstra algorithm provides computing the working path with the minimal cost, which in some specific cases can be desirable [57].

Connection blocking probability due to the SRLG diversity constraint

In order to ensure 100% survivability, the main criterion is that a single network failure does not affect both paths. This can be achieved through fulfilling the SRLG diversity constraint between the working and the backup path. To this end, GMPLS enables the routing protocol to disseminate the SRLG attribute tied to each TE link in the network. A TE link can belong to one or more SRLG groups. In particular, in a MLN network, the SRLG set of every created FA TE link is usually constituted by the union of all the SRLGs of the corresponding underlying optical links. This may lead that a FA TE link is composed of a large set of SRLGs depending on the number of optical links inducing the FA TE link. This, in turn, may complicate the computation of the working/backup paths ensuring the SRLG-diversity constraint.. Consequently, if a route for either working or backup path cannot be found, the requested connection will blocked. However, we can conclude that, in some network states, it is not always feasible to exclude all the SRLG-joint links for every single connection request. In other words, it is not always practical or even necessary to fully protect every connection in a MLN from a failure that might never happen.

Link vs. SRLG disjoint path computation

As it is explained before, the main goal of an efficient protection scheme is to provide 100% survivability in case of a single failure. This is even more difficult in a MLN due to the fact that although, at the upper layer, we compute a link-disjoint working/backup paths, it is very likely that the used FA TE links for each path may share the same underlying lower-layer link. This can happen, for example, when only the upper-layer TE Database (TED) is visible. Thus, only link-disjoint working / backup paths can be computed without guarantees that both path satisfies the SRLG-diversity constraint. Moreover, in some cases, it is not always possible or feasible to collect the SRLGs along the underlying lower-layer path (i.e., in case of loose routing). Therefore, ensuring only link-disjoint working / backup paths at the upper layer, a single failure of an optical link may affect specific and separated FA TE links forming the working and backup paths. In other words, the connection survivability fails. To address this problem, one needs to guarantee full SRLG disjointness applying an adequate path computation algorithm as long as the information related to both layers in the network is known.

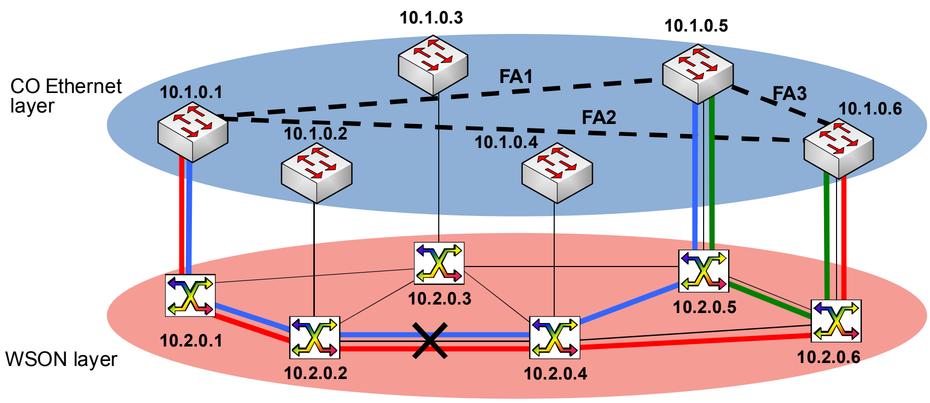


Figure 27. A single failure in a MLN affecting two FA TE links

One example is shown in Figure 27. From the figure, it can be observed that there are three established FA TE links (i.e., FA1: 10.1.0.1 - 10.2.0.1 - 10.2.0.2 - 10.2.0.4 - 10.2.0.5 - 10.1.0.5, FA2: 10.1.0.1 - 10.2.0.1 - 10.2.0.2 - 10.2.0.4 - 10.2.0.6 - 10.1.0.6 and FA3: 10.1.0.5 – 10.2.0.5 – 10.2.0.6 – 10.1.0.6) which are disjoint at the Ethernet layer, but at the lower layer FA TE links FA1 and FA2 shares the same optical link (i.e., 10.2.0.2 – 10.2.0.4). When a request between the nodes 10.1.0.1 and 10.1.0.6 arrives, the source node (i.e., node 10.1.0.1) computes the routes for both the working and protecting paths. Since FA TE links are disjoint at the upper-layer, the working path will be established through the direct FA TE link FA2, while the protecting path will be signaled through the links FA1 and FA3. If a failure of the optical link 10.2.0.2 – 10.2.0.4 occurs, this will affect the links FA1 and FA2, and hence both, working and protecting paths.

4.3 Review of the end-to-end protection works in the literature

In last few years, several strategies for e2e protection in MLN have been proposed. They can be classified in strategies that protect only at the lower-layer, only at the upper-layer, and a multi-layer protection strategy. In the lower-layer protection strategy, for every path at the lower layer, a protecting path at the same layer ensuring optical link-disjointness is established. Therefore, when a physical link failure occurs, the affected connections at the upper-layer (i.e., the connections over the induced FA TE links) will be routed over the already established lower-layer protecting paths. This strategy requires the minimal number of recovery actions, since protection is provided at the layer of coarsest granularity. Consequently, it provides a good performance regarding fast MTTR (<50ms), since lower-layer does not need to propagate the information about the failure to the upper-layer. However, the main drawback of the lower-layer protection strategy is that backup paths cannot be used for the establishment of other working paths, which leads to inefficient usage of the network resources. Some of the papers related to the protection at the lower layer are [46][74][75].

The upper-layer protection strategy is based on the establishment of a backup path for each working path exclusively at the upper-layer (i.e., L2SC). This strategy attains better network resource (i.e., wavelength channels) usage since the lower-layer LSPs constituting the VNT for the upper-layer are not dedicated to either working or backup paths. However, the main drawback is that it requires many recovery actions, due to a finer granularity of upper-layer connections [51]. For example, let us assume that one optical link at the WSON constitute more than one FA TE links belonging, thus, to the same SRLG group. Therefore, if a failure of such an optical link occurs, it will affect all the corresponding FA TE links. Therefore, the protection scheme needs to recover from as many failures as there are connections over the affected FA TE links [51][76][77], leading, thus, to a higher complexity (e.g., signaling notifications) and extended MTTR. An example of two FA TE links (i.e., FA1 and FA2) that share an optical link (i.e., the link between the nodes 10.2.0.2 and 10.2.0.4) has been shown in Figure 27.

Finally, the multi-layer strategy protects the connections through exploiting the cooperation / collaboration of both involved layers [51][52][53][78][79][80][81]. This strategy should provide a dynamic VNT reconfiguration for recovery purposes, exploiting the network resources at the involved layers during the establishment of both, working and backup paths. Such a recovery strategy requires overall view of the multi-layer network (i.e., topology and resources) as well as advanced path computation algorithms. The main advantage is that it achieves a more efficient usage of resources but at the cost of higher implementation complexity. An particular application of the multi-layer strategy relies on involving a partial protection strategy, where some of the upper-layer links are statically protected at the lower-layer [79][80]. In general, the decision at which layer a connection will be protected is made based on different pre-defined criterion (e.g., required restoration time, affected traffic, etc), which also represents one of the challenges.

In the following, we will present contributions of this dissertation regarding e2e dedicated protection in a MLN. It can be divided in three parts:

- In Section 4.4, we compare two approaches for e2e dedicated protection. In the first approach, all the TE links (either FA TE links or physical links) that constitute a working path are temporarily removed from the network graph, in order to compute the backup path. On the other hand, in the second approach, all the SRLG-joint links forming the working path are removed for the backup path computation. The main goal is to evaluate the performance of these two approaches regarding connection blocking probability and survivability. The work has been published in [82].
- In Section 4.5, we present a dedicated protection scheme that aims at balancing the usage of the network resources at both involved layers, providing a high level of total network throughput and a high connection survivability even with very high dynamicity of link failures. The work has been published in [83].
- Finally, in Section 4.6 we present a protection mechanism for MLN that aims at reducing the connection blocking probability, but keeping a high level of survivability. In the proposed scheme, SRLG-joint links are not totally removed from the graph when computing the backup path, but SRLG-disjoint links are still prioritized. We refer to this mechanism as *SRLG-scaled* protection scheme.

4.4 Link vs. SRLG disjointness dedicated path protection

The goal of this contribution is to compare two MLN protection scenarios in which the connections are protected satisfying either *link-* or *SRLG-disjointness* when computing the pair of working and backup paths. For the sake of simplicity, we call these two schemes: *Link Disjoint Scheme* (LDS) and *SRLG Disjoint Scheme* (SDS).

4.4.1 Description of LDS and SDS protection schemes

Both, the LDS and the SDS dedicated protection schemes rely on deploying a two-step Dijkstra algorithm, where after computing the working path, the backup path is computed. If either path cannot be established, the connection request is blocked. FA TE links are created dynamically on demand, where the underlying lower-layer LSP is triggered by an upper-layer LSP request (i.e., LSP nesting using the GMPLS hierarchy). Both working and backup paths can trigger the establishment of a new LSC LSP. Once the LSP is established, the newly induced FA TE link can be used for the establishment of both, a working and a backup L2SC path. In other words, the established lightpaths (i.e., FA TE links) are not dedicated to either L2SC working or backup paths. For the path computation of working path, we have used the same algorithm explained in the Section 3.6, which is proposed in [40], for both schemes (i.e., LDS and SDS). Once the working path is computed, both schemes compute the backup path. In the LDS scheme, all the working path TE links will be temporarily removed from the graph. Analogically, all the working path SRLG-joint links will be removed from the graph when SDS scheme is deployed. The pseudo code for the SDS scheme is depicted in Figure 28. We have used the same notion as in Section 3.7.2. WP and PP represent the working and the backup paths, respectively, and SRLG (WP) represents the union of the SRLGs of all the links that constitute the WP.

```
INPUT
    G, V, E, FA, E', s, d, req_band
OUTPUT
   WP, PP
STEP 1.
   Compute shortest path-based algorithm to find WP
   if WP is not found
      exit
STEP 2.
   for ∀e | e ∈ E'
      if SRLG(e) ∩ SRLG(WP) ≠ ∅
         remove the link e from the network graph

   Compute the shortest path-based algorithm in the modified
   network graph to find PP
   if PP is not found
      Restore all the removed links
      exit
STEP 3.
   Restore all the removed links
   end algorithm
```

Figure 28. Pseudo code for SDS protection scheme

4.4.2 Model assumptions

In our model, all the physical links are bidirectional, while the LSPs and, thus, the FA TE links in the network are unidirectional. Since we consider the 1+1 dedicated protection, as it is explained in the previous sections, both working and backup paths are established once a source node receives the connection request. Therefore, once a failure occurs, the destination node will immediately receive the traffic through the backup path. For that reason, when 1+1 protection scheme is used, it is not necessary to send a Failure Indication Message (FIM) [48]. However, in our model, for each connection that uses the damaged link, a FIM message is sent backwards to the corresponding source node, for two reasons: firstly, to notify the source node of each LSC LSP connection, so the associated FA TE link can be excluded from the network graph and, hence, it cannot be used for the future path computation until the failed link is repaired; secondly, we used a FIM message for statistical purposes.

In this dissertation, we model only link failures. A failed link is randomly chosen and its holding time is exponentially distributed. However, although we do not use the simultaneous failure model, it may happen that there are two (or more) links that are failed at the same time, due to the high dynamicity of the failures we used. Therefore, it may happen that both working and backup path are hit with two different failures, and thus, the 100% recovery cannot be achieved.

When the SDS scheme is deployed, we assume that each physical link belongs to one SRLG. Thereby, two physical links cannot belong to the same SRLG group. On the other hand, each FA TE link belongs to the union of the SRLGs of its underlying optical links.

It is worth mentioning that, in both schemes, if during the lifetime of a connection, there is always working and / or backup path that is not hit by a failure, we assume that the connection "survived", since the data transmission was never interrupted. On the other hand, if both paths, working and protecting, are affected at the same time by a set of contemporary failures, we assume that the connection cannot be protected (i.e., survivability is not 100%) and the resources are released.

4.4.3 Simulation parameters

The evaluation of the algorithm is performed in the OPNET simulation tool [4]. Two network topologies are considered: NSFNET [42] and the basic reference topology for the Pan-European network [43] (Section 3.7.3, Figure 16). All the nodes in the network are single switching capable nodes. Every OXC is connected with one CO-Ethernet switch with full grooming capacity. OXCs have a fixed number of wavelength converters, 10 per node in the NSFNET network and 20 per node in the Pan-European network. In NSFNET network, there are 21 links between OXCs and 14 links more, where each of them connects an Ethernet node with one OXC. In Pan-European network, there are 41 OXC-OXC links in the network and 28 Ethernet-OXC links that connect Ethernet nodes with the OXCs. A failed link is randomly chosen among all the physical links between any pair of OXCs. Therefore, we do not consider a failure of the links that connect CO-Ethernet and optical nodes. In practice, they are either very close (short link distances) or integrated in a single network element forming a hybrid node [6].

The Ethernet nodes generate connection requests dynamically, following a Poisson arrival process with the average inter-arrival time of 1s. The destination node is chosen uniformly, among the rest of the Ethernet switches. The LSP holding time is exponentially distributed and varies in different simulation sets from 400s to 1600s in NSFNET network, and from 400s to 3200s in Pan-European network.

Every optical link is bidirectional, and each of them supports 16 wavelengths in NSFNET and 32 wavelengths in the Pan-European network. The rate of every wavelength channel is 2.5Gb/s, while the requested bandwidth for every connection is

100Mb/s. We only consider a link propagation delay, which is the same for every link, and it is 3.5ms. As it is explained in the previous section, link failures happen following a Poisson point process, where a failed link is randomly chosen among all the links between any two OXCs. To stress the behavior of the network using LDS and SDS, average time between link failures is 200s in NSFNET, and 1000s in the Pan-European. In both networks, each link is, in average, 10s under failure, following an exponential distribution. Each simulation experiment is run with 100 000 L2SC LSP requests.

4.4.4 Numerical results

The considered performance metrics for the evaluation of the approaches are the survivability and the connection blocking probability.

Connection survivability in NSFNET network

Figure 29 shows the connection survivability as a function of the traffic load in the NSFNET network.

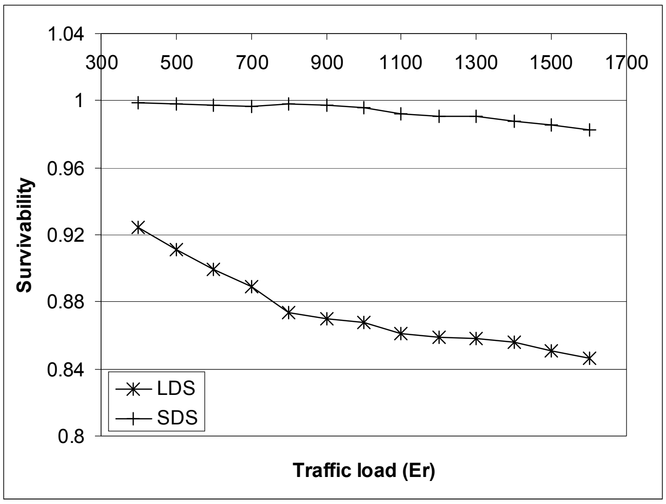


Figure 29. Connection survivability in NSFNET network

As it can be observed, for the given parameters, by using proposed SDS, the survivability is almost 100% at the low traffic load, while at the high traffic load is slightly less, around 98%. On the other hand, having LDS, the survivability is significantly lower ranging from ~8% at the low traffic load to ~15% at the high traffic load, when comparing to the survivability in SDS. This is because, in case where the L2SC connections are protected using LSD scheme, only TE links that are part of the working path are removed from the network graph. Therefore, it may happen that two upper-layer disjoint TE links (i.e., FA TE links) still share one or more physical links at the lower layer. Therefore, if one of these lower-layer links fails, both working and backup paths at the upper-layer will be affected. In such a situation, the survivability of the connection cannot be ensured, which in turn worsens the associated performance metric. On the other hand, using the proposed SDS strategy, both (working and backup) paths will fail only if there are two different simultaneous failures where one affects the working path and the other one affects the backup path. However, this double-failure situation is considered to occur rarely.

As the traffic load increases, connections tend to last longer. Thus, there is a higher probability that one LSP will be affected by the failure. That is why, with the higher traffic load, the survivability in both scenarios is decreasing.

Connection survivability in Pan-European network

Figure 30 shows the survivability in the Pan-European network as a function of traffic load. Similarly to NSFNET network, using the SDS scheme, the survivability attains almost 100% (specifically, ~99.9%) at low traffic load and ~99.7% at high traffic load in the Pan-European network. On the other hand, using LDS, the survivability is worsened up to ~3% at the low traffic load and ~7.5% when compared to SDS. Comparing the connection survivability of these two networks, we can observe that by using SDS in both networks, the survivability in Pan-European network decreases slightly slower than in NSFNET network. That is because in the scenario with NSFNET topology, link failures happen more often (with the mean of 200s) than in Pan-European network (~1000s). Finally, we can conclude that using SDS in both networks, we can achieve a survivability of almost 100%, even with very dynamic link failures.

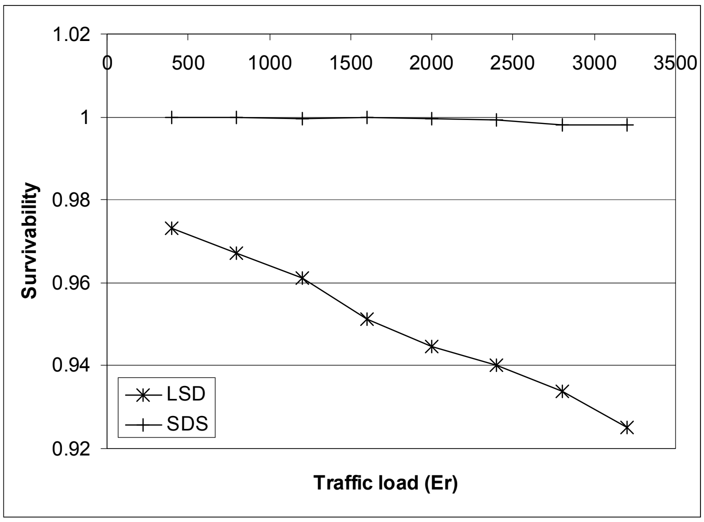


Figure 30. Connection survivability in Pan-European network

Connection blocking probability in NSFNET network

Figure 31 shows the connection blocking probability as a function of traffic load in the NSFNET network.

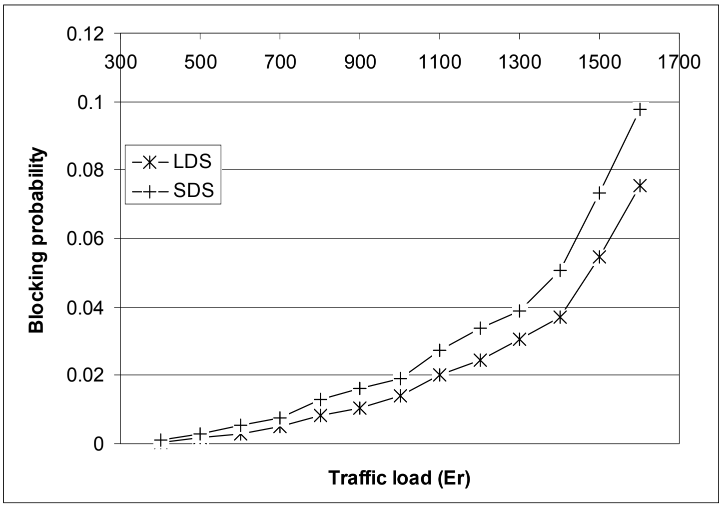


Figure 31. Connection blocking probability in NSFNET network

From the figure, it can be observed that SDS gives a higher connection blocking probability than LDS. That is because, by removing all the SRLG joint links from the network graph in the SDS, the probability to find a route for backup paths is lower than in the LDS scenario where only working upper-layer TE links are removed. Moreover, it can be seen that, as the traffic load is increased, the difference in blocking performance using the LDS and the SDS slightly increases (up to ~2.5% at high traffic load). That is because both schemes can easily accommodate the connections at the lower traffic load. Nevertheless, at the high traffic load, using SDS scheme it is more difficult to establish the connection since more links need to be excluded from the network graph when compared to LDS. Therefore, it is less probable to find a backup path and, thus, to establish a connection.

Connection blocking probability in Pan-European network

The connection blocking probability in the Pan-European network is shown in the Figure 32. From the figure, we can observe that the difference between the blocking probability in LDS and SDS is much more appreciable than in the NSFNET network. That is because the Pan-European network is bigger, and thus, the paths are longer in terms of number of hops. Therefore, there are more links to be excluded when computing a backup path, than in the NSFNET network. For that reason, it becomes more difficult to accommodate the connection requests on the links that are left. Moreover, similarly to the result in the NSFNET network, by increasing the traffic load, the difference between blocking probability using the LDS and SDS schemes is slightly increasing (up to ~8% at the high traffic load).

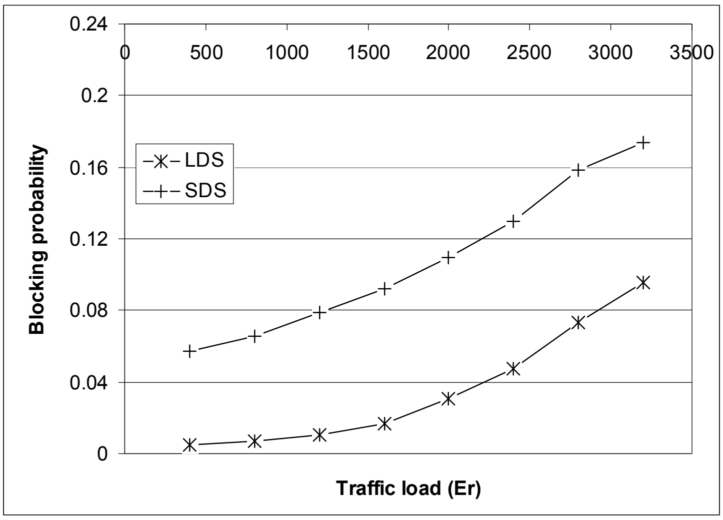


Figure 32. Connection blocking probability in Pan-European network

Finally, we can conclude that, using the SDS scheme, we get the survivability of almost 100% in both network topologies, NSFNET and Pan-European. However, by excluding SRLG disjoint links in a MLN, the probability that a path computation algorithm will not be able to find a path with the rest of the links increases. Moreover, the differences in the connection blocking probability become more relevant as traffic load grows up, since the connection duration is longer, and there are less available resources. Therefore, when choosing a protection scheme for a MLN between LDS and SDS, a trade-off between the connection blocking probability and the survivability should be taken into account.

4.4.5 Conclusion: Link vs. SRLG disjointness dedicate path protection

The main objective of this contribution was to compare two basic approaches for e2e dedicated protection in a MLN. Through the set of simulations using two reference network topologies, NSFNET and Pan-European, and using a given dynamic failure model, we have shown that SDS attains almost optimal results in the sense of connection survivability. However, since all the SRLG joint links used by the working path need to be excluded from the network graph when a computing the backup, SDS does increase the blocking probability with respect to LDS. Finally, we can conclude that, when the network disposes of enough resources, and when the connection survivability is the key objective, it is better to deploy SDS scheme since it assures the survivability of almost 100%. On the other hand, when it is more important to assure lower connection blocking probability, the LDS scheme is a better solution.

4.5 Link load dependent scheme for e2e dedicated protection

One of the issues in a protected MLN is to provide a high level of survivability, but keeping an efficient network resource usage at both layers. Herein, we propose a protection path scheme based on the two-step Dijkstra algorithm targeting the above challenge. The decision used for choosing the working and backup paths is based on two main criteria: The first one is the number of physical hops, and the second one is the current traffic load of the TE links. Thanks to the first criterion, the path with minimal number of physical hops is favored. By doing so, we reduce the probability of the creation of routing loops at the lower-layer (i.e., that two FA TE links constituting a path share the same physical link). Consequently, the possibility that the connection will be affected by a failure is reduced. However, because of the second criterion, the algorithm tends to allocate the requested traffic onto the least congested TE links. By doing so, the proposed algorithm balances the usage of the network resources at both involved layers. Furthermore, the proposed algorithm allows multiple changes of layer, as long as the switching capability constraints (i.e., layer adaptation) are satisfied. Finally, in order to ensure a high level of survivability, only working path SRLG-disjoint links are taken into account when computing a backup path [69]. The work has been published in [83].

4.5.1 Protection scheme

In order to describe the proposed path computation algorithm we use the following notation:

A connection request is specified as $r\ (s, d, bandw)$, where s is the source node, d is the destination node, and $bandw$ is the requested bandwidth by r;

E represents the set of the physical links in the network;

FA represents the set of all the currently established FA TE links;

$E' = (E \cup FA)$;

V represents the set of physical LSC or L2SC switch nodes;

$G = (V, E')$ represents the network graph G, with the set of the nodes V and the set of the links E';

$e = (u, v)$ is any link that belongs to E' between the nodes u and $v \in V$;

$W(e)$ presents the total number of wavelength channels supported by an optical link $e \in E$;

$W'(e)$ presents the number of occupied wavelengths channels of the link $e \in E$;

$max\_band(e)$ is the maximum bandwidth supported by a TE link $e \in FA$;

$occupied\_band(e)$ is the bandwidth of the link $e \in FA$ that is currently occupied;

cost (e) present a cost of any TE link in the network.

Path computation of working path

When a request arrives to the source node *s* with the parameters *r (s, d, bandw)*, every link $e \in E'$ that does not have enough bandwidth to accommodate the request is temporarily removed from the network graph *G*. Afterwards, the node *s* firstly checks if there is a direct upper-layer link between the nodes *s* and *d*. If there is, the working L2SC LSP will be established over such a link. If not, the node *s* will execute the shortest path algorithm.

The algorithm calculates the shortest path from the *s* to the *d*, taking into account the *cost(e)* across the path. For every $e \in E'$, the *cost (e)* is set as follows:

- For every optical link *e*, the cost is equal to one, if the number of occupied wavelengths is below a given threshold, that is, for a given parameter α = [0, 1], the condition $W'(e) \leq \alpha\ W(e)$ is satisfied. Otherwise, the cost of link *e* is assigned to two.
- For every FA TE link the *cost (e)* is equal to the number of underlying optical links constituting the associated LSC FA LSP, if the condition $occupied\_band(e) \leq \alpha\ max\_band(e)$ is satisfied. If this is not the case, the cost of the link *e* is equal to the number of underlying optical links plus one.

By assigning these values to the *cost (e)* we are achieving two goals: firstly, the paths with the minimal number of physical hops have priority, and, therefore, we minimize the probability of creating routing loops at the optical layer. In other words, the algorithm will avoid using two FA TE links that share a given physical optical link. And, secondly, whenever the link resources are occupied over a pre-defined limit determined by parameter *α*, we are avoiding using such a link by increasing its cost by +1. Therefore, the connections tend to be accommodated over other TE links less congested (i.e., lesser cost) considering both layers. By doing so, we achieve to maintain the balance of the network resource usage at both layers.

Apart from the explained metrics, the following constraints are also taken into account:

- Whenever it is possible, the algorithm tries to satisfy the WCC along the entire route / optical segments. When this constraint cannot be satisfied, the algorithm will check if there is an available wavelength converter at a given node. If either of these conditions is not satisfied, the request will be blocked.
- During a L2SC LSP establishment, it is possible to dynamically change the layer several times, taking into account the interface switching capability of the TE links. In this situation, it is necessary to satisfy the layer adaptation constraint among the different layers.

- If there are two or more paths that have the same cost, the algorithm will always choose the one that triggers the establishment of LSC LSP the least number of times, accommodating request over the upper-layer (i.e., L2SC).

Finally, if the path from the *s* to the *d* cannot be found, the request will be blocked. Otherwise, it is needed to compute the backup path for that request.

Path computation of backup path

Before computing a backup path, it is needed to temporarily remove from the network graph all the links that are SRLG joint with the working path. We consider a full cooperation among the layers, under the framework of GMPLS unified control plane. Therefore, we will remove all the links that have any physical resource occupied by the working path. In other words, the TE links that have at least one joint link considering both FA TE links and optical physical links with the computed working path will not be considered during the path computation. By doing so, we reduce the probability to zero that both working and backup paths will be affected by a single link failure. Finally, the same path computation algorithm used for computing the working path is executed to get the backup path.

If both, working and backup path are computed successfully, the source node *s* will start the signaling mechanism to reserve / allocate the resources through the computed working and backup routes, otherwise the request is blocked.

4.5.2 Performance evaluation

In order to evaluate the proposed path computation algorithm we compare it with the two following algorithms found in literature:

The first algorithm presents the *Policy 1* proposed in [17], with the assumption that the maximum number of hops at the upper-layer is unlimited and that all the metrics of the FA TE links are equal to one. Moreover, deploying this algorithm, a request will be accommodated either at the lower or at the upper-layer, exclusively (see Section 3.4). For the sake of simplicity, we call this as *minimal hops* algorithm, or simply, MH. The second algorithm is the algorithm proposed in [30], and it is called a *minimal physical hops* algorithm, or simply, MPH. It can be observed that our proposed algorithm is equal to the MPH algorithm for the values of $\alpha = 1$. The authors in [17] and [30] are taking into account an unprotected MLN network. However, both algorithms can be applied on a protected network as long as SRLG disjointness between a working and a backup path is guaranteed. In both, MH and MPH algorithms, similarly to the proposed algorithm, the WCC must be assured, and when this is not possible, both algorithms will check if there is a wavelength converter at the given optical node.

Finally, we call our proposed path computation algorithm as integrated multi-layer algorithm, or IML. For the sake of simplicity, we consider only three different values of the parameter: $\alpha = 0.25$, $\alpha = 0.5$ and $\alpha = 0.75$.

4.5.3 Model assumptions

In order to evaluate the network behavior when the proposed protection scheme is used, we use the same model assumptions like in Section 4.4.2. In summary, those are:

- We assume that each physical link belongs to a different SRLG group, while SRLG of every established FA TE link presents the union of SRLGs of the underlying optical links.
- We consider only a single link failure model.
- The dynamicity of failures is high, and, thus, it might happen that two links are hit by failures at the same time.
- We only consider the failures between two OXC, assuming that the distance between Ethernet nodes and OXC are either very small or together they constitute a hybrid node [6].

4.5.4 Simulation parameters

The algorithm is experimentally evaluated through the set of simulations [4], using NSFNET [42] and Pan-European [43] network topology. The nodes in the network are single switching capable nodes with either L2SC (i.e., CO-Ethernet switches) with full grooming capacity, or LSC (i.e., OXC) with the limited number of wavelength converters (10 per node in NSFNET and 20 per node in Pan-European network). All the optical physical links are bidirectional supporting 16 and 32 wavelengths per direction in the NSFNET and Pan-European topologies, respectively. The FA LSPs are unidirectional and they support 2.5Gb/s. We assume that the control link propagation delay is the same for every control link and it is 3.5ms. The Ethernet nodes generate LSP requests dynamically, following a Poisson arrival process with the average inter-arrival time of 1s and the LSP holding time is exponentially distributed. Destination is chosen uniformly, among the rest of the Ethernet nodes. The requested bandwidth for every connection is 100 Mb/s. Similar to Section 4.4.3, link failures arrive according to a Poisson process, where a failed link is randomly chosen among all the links between any pair of adjacent OXCs, with the average time between link failures set to 200 s in NSFNET and 1000 s in the Pan-European. In both networks, each link is, in average, 10 s under failure, following an exponential distribution. Each point is obtained having 90000 L2SC LSP requests.

4.5.5 Numerical results

The evaluation of the proposed protection scheme is carried out using two performance metrics: connection blocking probability and survivability, both as a function of traffic load.

Connection blocking probability in NSFNET network

Figure 33 shows the connection blocking probability as a function of the traffic load expressed in Erlangs (Er) in the NSFNET network.

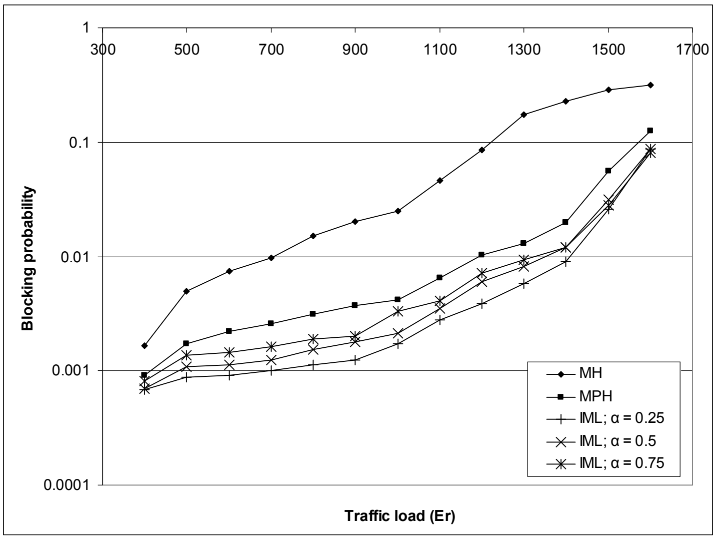


Figure 33. Connection blocking probability in NSFNET network

We observe that the highest blocking probability is obtained when the MH algorithm is used. The reason behind this is the following: since all the TE links in the network have the cost equal to one, the algorithm favors the routing over the upper-layer links. Therefore, it might cause routing loops, repeating the optical links which constitute the different FA LSPs forming the computed paths. Consequently, this leads to faster saturation of the network since the resources are much more rapidly exhausted, and thus, a very high blocking probability is attained.

The MPH algorithm outperforms (i.e., lower the blocking) the MH algorithm. This is because the MPH algorithm chooses the route with the minimal number of physical hops. Therefore, the possibility that an optical link that is a part of two different FA LSPs is repeated decreases. This, in turn, leads to attain less network saturation (i.e., exhaustion of the link resources) than when compared with the MH algorithm. On the other hand, the proposed IML algorithm still provides lower blocking probability for all three values of α (i.e., α = 0.25, α = 0.5 and α = 0.75). This is because, apart from the number of physical hops, this algorithm is "punishing" the links whose resource occupancy is greater than a predefined parameter α. Recall that this is applied to both, optical and FA TE links. Therefore, the IML algorithm aims at accommodating the traffic on less congested links. Consequently, both network layers achieve a more balanced resource utilization, which does lower the blocking probability. However, one can observe that, at the low traffic load, we obtain the worst results for α = 0.75. This is because the algorithm is avoiding using the links whose resources are occupied above 75% by adding +1 to the link costs. Nevertheless, at low traffic load, the number of links under such a condition is significantly small. Therefore, for higher values of α and at the low traffic load, the proposed algorithm behaves as the MPH algorithm, since the traffic load of a link rarely exceeds the predefined parameter. Finally, we can observe that, by increasing the traffic load, the differences of the blocking probability for the used values of α are becoming negligible.

Connection blocking probability in Pan-European network

The results in the Pan-European network topology are depicted in the Figure 34.

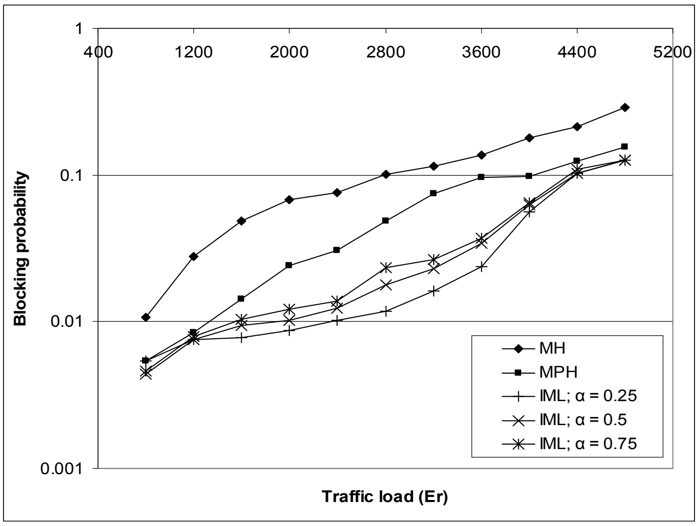


Figure 34. Connection blocking probability in Pan-European network

We can observe that the results are similar to those obtained in the NSFNET topology. Specifically, the proposed algorithm outperforms the MH and MPH for all the three values of α, and such a difference becomes more relevant when the traffic load grows. As it is explained above, the reason behind this is that, at the low traffic load, there are very few links whose resource occupancy exceeds even 25%. However, when the traffic load is increased, the difference becomes bigger, since the algorithm avoids using highly congested TE links. Finally, at the high traffic load, the connection blocking probability of all three algorithms converges since there are not enough resources to accommodate the connection requests.

Connection survivability in NSFNET network

Figure 35 depicts the connection survivability as a function of the traffic load. It can be observed that the connection survivability when the proposed algorithm is deployed outperforms the achieved by both algorithms MH and MPH, even for the different values of the parameter α. As it is explained before, the MH algorithm may lead to very long paths with the possibility of routing loops at the lower-layer. Therefore, the probability that both, working and backup paths will be affected by two different failures at the same time is higher. Deploying IML algorithm, we obtain almost 100% of survivability (i.e., ranging from ~99.9% to ~98.6 at low and high traffic loads, respectively) even with a very high dynamicity of link failures. This is because the algorithm accommodates the traffic onto the least congested TE links taking into account both layers. By doing so, it attains a good traffic load balance among all the network resources. As a result, fewer connections are interrupted by failures when compared to the routing performed by the MPH algorithm which does not permit the same traffic load balancing over the network resources. Analogously to the connection blocking probability, for the higher values of α, at the low traffic load, the network survivability attained by IML algorithm performs similarly to the network survivability of MPH algorithm.

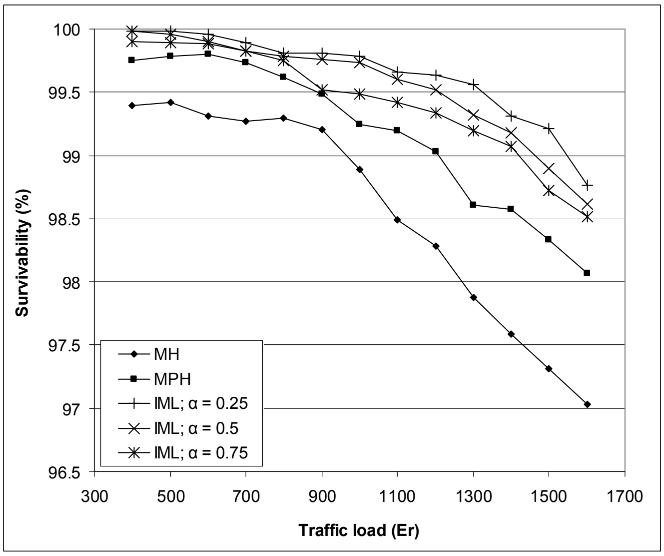


Figure 35. Connection survivability in NSFNET network

Connection survivability in Pan-European network

Finally, Figure 36 depicts the connection survivability as a function of traffic load in the Pan-European network. Similar to the obtained results in the NSFNET topology, the proposed algorithm outperforms the MH and MPH algorithms. Specifically, we can conclude that the proposed algorithm is able to ensure almost 100% of survivability (i.e., up to ~99%) even when both conditions high traffic load and very dynamic failure model are being applied.

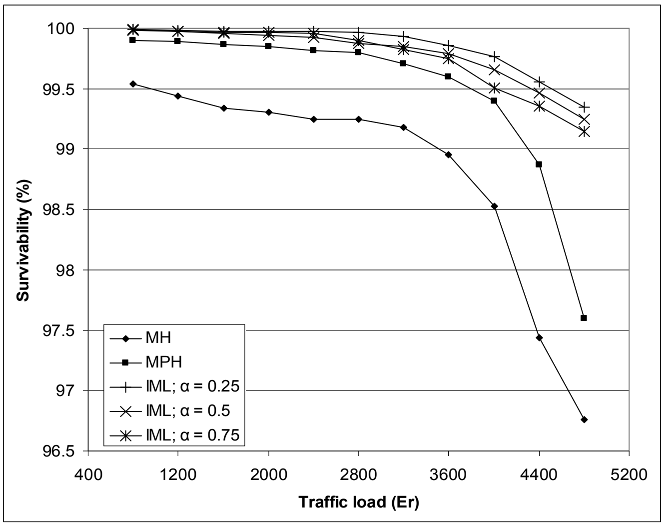


Figure 36. Connection survivability in Pan-European network

4.5.6 Conclusion: Link-load dependent scheme for e2e dedicated protection

The main goal of this contribution is to provide a protection scheme for a GMPLS-based controlled MLN, which would minimize the number of failed connections, attaining an adequate balance of the network resources at both involved layers. The algorithm is based on a two-step approach which computes the shortest path algorithm, taking into account the WCC and the regenerator allocation at the optical layer. Its objective function is to prioritize the paths with the least number of physical hops. Moreover, the algorithm balances the utilization of the network resources at both involved layers, avoiding the traffic accommodation over congested TE links. While computing a backup path, the algorithm ensures full SRLG disjointness with respect to the working path. We have used two network topologies to evaluate the algorithm: NSFNET and Pan-European.

Through the set of simulations, it is shown that the algorithm outperforms the MH and the MPH algorithms, regarding both, the connection blocking probability and the survivability. For a given dynamic failure model, we obtained almost 100% of connection survivability. Moreover, since the algorithm tends to accommodate the traffic over the least congested TE links, it attains much better performance regarding connection blocking probability, when comparing with the MH and MPH algorithms. However, we can conclude that, there is no universal optimal value of the parameter α, but it is needed to be adjusted according to the network parameters and resources that a network disposes.

4.6 SRLG-scaled dedicated protection scheme for MLN

In this section, we analyze the influence on the connection blocking probability when removing SRLG disjoint links from the graph at the time of computing the backup path. In general, a FA TE link in a MLN inherits the SRLGs from the corresponding underlying LSC LSP. For that reason, the more physical links constitute a FA TE link, the higher is the number of the SRLGs that the FA TE link belongs. Consequently, all the links in the network that belong to the SRLGs of the working path links (i.e., SRLG-joint links) will be removed from the network graph when calculating the backup route. The more links are excluded, the more difficult is to find a backup path, and, thus, to establish the connection. However, it is not always necessary to provide a full protection for every single connection in the network. Moreover, it is possible that a connection cannot be established due to SRLG constraints, even though a failure never occurs.

To the best of our knowledge, no works in the literature have addressed this problem in a MLN with dedicated path protection. We propose a scheme that maintains almost 100% survivability even in case of very high failure dynamicity, while keeping the blocking probability low. The scheme does not exclude all the SRLG-joint links from the path computation, but it still prioritizes the SRLG-disjoint links. Therefore, if a backup path cannot be found using exclusively the SRLG-disjoint working links, there is also a possibility of using the SRLG-joint working links. Through a set of simulations, it is shown that the proposed scheme provides a good balance between blocking probability and survivability performance metrics. Finally, we should stress that the scheme is flexible in the sense that, its special case does assure a full protection (i.e., 100% survivability in case of single failure). In the following section, the scheme will be explained in details.

4.6.1 Description of the path computation strategies

The proposed algorithm calculates the shortest path between the source and the destination node, for both, working and backup paths. The cost of each optical link is equal to one, while the cost of every established FA TE link is equal to the number of underlying optical links minus one, like in [84], aiming at favoring the use of FA TE links over a establishing new LSC LSP. If, during the path computation of either working or backup paths, there are two route solutions with the same cost, the algorithm always chooses the one that minimizes the number of inter-layer crossings, aiming at accommodating the LSP request as much as possible over multiple existing FA TE links. In the proposed scheme, all the LSC and L2SC LSPs, and, thus, the FA TE links, are set up dynamically and automatically. Both, working and backup paths can trigger the establishment of a LSC LSP, and the established lower-layer LSP can be used in the future by either working or backup paths of upper-layer connections, as long as sufficient unreserved bandwidth is available. The computation of the working and backup routes is done using the topology and resource information of all the involved layers gathered in the TED.

Path computation of working path

When a source node receives a LSP request, first, it will check whether there is a direct FA TE link that connects the source and the destination nodes with enough unreserved bandwidth. If this is the case, the connection request is established over the FA TE link. Otherwise, the source node computes a shortest path taking into account several constraints (i.e., WCC, switching capability adaptation constraints, etc) and using the TE link cost like in [84]. Once the route for the working path is found, the backup path computation is executed.

Path computation of backup path

Herein, we propose a protection scheme, where all the SRLG-disjoint links with respect to the computed working path are set to the same cost as for the working path computation [84]. The cost C of a link l is (i.e., $C(l)$), on the other hand, reconfigurable for those SRLG-joint links with respect to the computed working path, and it depends on the parameter α = [0, 1]:

$$C(l) \leftarrow \frac{1}{\alpha} C(l)$$

As it can be observed, for α = 1, the costs of the SRLG joint links is equal to the cost of the SRLG-disjoint links. In other words, in this case, the algorithm takes into account both equally, SRLG-joint and SRLG-disjoint links. On the other hand, setting the value of α to 0, the full SRLG-disjointness is totally guaranteed (i.e., 100% survivability for single failure occurrence). For the rest of α values, we achieve the protection scheme where the SRLG-joint links are not totally excluded from the backup path computation, but still the algorithm aims at favoring the usage of the SRLG-disjoint links. By doing so, the blocking probability due to SRLG diversity constraint is reduced, but the level of survivability still remains high since the SRLG-disjoint links have higher priority (i.e., lower link cost) compared to the SRLG-joint links. Consequently, adjusting the value α, with the proposed scheme we may state a tradeoff between the connection blocking probability and the connection survivability. Finally, it is noteworthy to outline that other constraints (i.e., WCC, switching capability, etc.) considered for working path computation are also applied for the backup path computation.

4.6.2 Simulation parameters

The proposed protection scheme is compared to two approaches (similarly to Section 4.5.2): the first one represents the *Policy 1* in [17]. This algorithm tries to accommodate a request over FA TE links, and if this is not possible, it establishes a direct LSC LSP between the source and destination nodes. The second scheme represents the algorithm proposed in [30], where, when establishing both working and backup paths, the cost of every FA TE link is equal to the number of underlying optical links, while the cost of each optical TE link is equal to one. This algorithm enables dynamically changing the layer at intermediate nodes as long as switching capability constraints are satisfied. However, since these two algorithms take into account an unprotected MLN, we assume that both schemes deploy the Suurballe-based algorithm for finding two disjoint paths. This is the main difference between these two algorithms and the algorithms used in Section 4.5.2. For both algorithms, we assume that a full SRLG-disjointness is guaranteed. For the sake of simplicity, we call these to schemes *upper-layer first scheme* (ULFS) and *multi-layer scheme* (MLS), respectively.

Finally, we call the proposed scheme as scalable multi-layer scheme, or simply, SMLS. In order to see the influence of parameter α in both the connection blocking and survivability figures of merit and for the sake of simplicity, we use two extreme values: $\alpha = 0.1$ and $\alpha=0.9$.

The evaluation of the protection scheme is carried out using the NSFNET topology [42]. Every OXC has 10 wavelength converters and optical links are bidirectional where each of them supports 16 wavelength channels. The dynamically induced FA TE links are unidirectional and they support 2.5 Gb/s, representing the rate of a wavelength. The requested bandwidth for each L2SC LSP is 100 Mb/s. The L2SC nodes generate connection requests dynamically, following a Poisson arrival process with the average inter-arrival time of 1s. Destination nodes are chosen uniformly, among the rest of the L2SC nodes. The LSP holding time is exponentially distributed, with a mean set to 1000s. A link failure occurs in average every 200s according to a Poisson point process. The duration time of such failures is exponentially distributed, and it varies allowing the evaluation of the network behavior under different dynamicity of failures. In different simulation cases, it is set from 120s to 300s. A failed link is randomly chosen among all the OXC-OXC links. Each point is obtained with 90000 L2SC LSP requests.

4.6.3 Suurballe-based algorithm

The proposed scheme can be applied using both algorithms for finding two disjoint paths, the two-step Dijkstra and the Suurballe algorithm. The main features of both algorithms are already explained in Section 4.2. However, we should remind that the main advantage of using the Suurballe algorithm is that it systematically avoids the so-called “trap topology”. However, using the two-step Dijkstra algorithm, it is obtained the minimal cost working paths and the complexity of this algorithm is much lower (Section 4.2).

In order to choose which one we will use for the proposed dedicated protection scheme, we have compared the total costs of working + backup path of these two algorithms. We have chosen this metric since the target is to minimize the signaling due to long paths (i.e., working and backup). In both algorithms, the two-step Dijkstra and the Suurballe, the cost of TE links is set like it is [84] (i.e., for each optical link the cost is equal to one, while the cost of every established FA TE link is equal to the number of underlying optical links minus one). In both cases, the algorithms satisfy the SRLG-disjointness.

For the sake of simplicity, we say that the Suurballe algorithm is composed of two parts (Section 4.2). The aim of the first part is to remove the links that can cause the trap topology problem (i.e., in Section 4.2, Figure 26. c). Once they are removed, a pair of two disjoint paths can be found, which represents the second part of the algorithm (Figure 26. d). We use a Suurballe-based algorithm adjusted to the proposed scheme. In the first part of the algorithm, we do take into account neither the WCC nor the SRLG-joint link cost. However, once the common links are removed and, thus, the trap topology is avoided [70][72], in order to find the pair of working and backup paths, we run a two-step Dijkstra algorithm, applying all the mentioned constraints and SRLG-joint link cost.

The total cost of working + backup path as a function of failure dynamicity is depicted in Figure 37. By the failure dynamicity, we assume the ratio between the holding time of a failure and its arrival time.

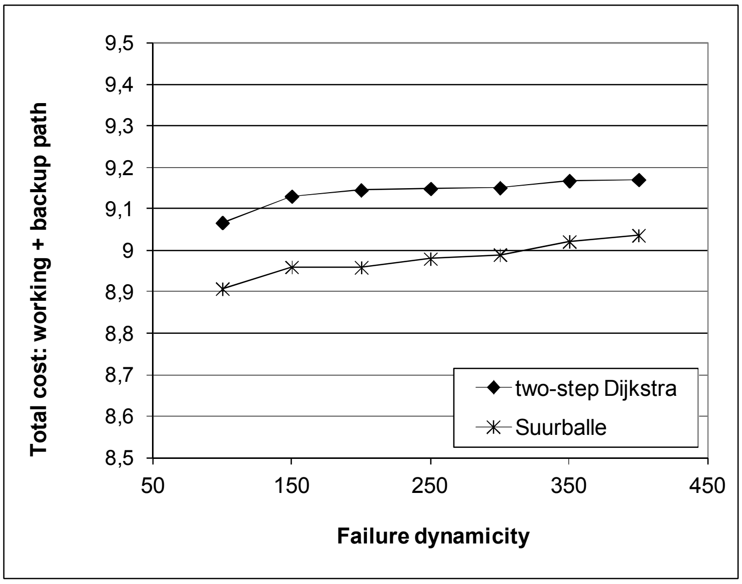


Figure 37. Total cost of working + backup path

From the figure, it can be observed that, with the Suurballe algorithm it is obtained a lower average cost of the pair: working + backup path. Therefore, in case of using the Suurballe algorithm, signaling messages traverse shorter distance and, thus, the control overhead is lower. Moreover, as it is explained before, the Suurballe algorithm successfully avoids the trap topology. For these reasons, we have implemented a Suurballe-based algorithm [70] in the developed simulator [4].

4.6.4 Numerical results

The used performance metrics are the blocking probability and the connection survivability, both as a function of the failure dynamicity under the same traffic load.

Connection blocking probability

Figure 38 shows the blocking probability as a function of the failure dynamicity. From the figure, it can be observed that the blocking probability is the highest when the ULFS scheme is used. This is because this scheme, first, tries to accommodate a request over the established FA TE links, and thus, it leads to faster saturation of the resources at the upper-layer. Moreover, this scheme does not allow the change of layer at intermediate nodes, so it is more difficult to establish a connection. On the other hand, using the MLS scheme it is possible to change the layer several times, as long as the switching capability constraints are satisfied. For this reason, the MLS is more flexible than the ULS scheme, and it gives better results in terms of blocking probability. However, both schemes exclude the SRLG-joint links when establishing a backup path. Therefore, it is more difficult to find a route for a backup path since there are less TE links that are taken into account for the path computation. Our proposed scheme, as mentioned above, aims at relaxing this problem through setting higher cost to conflicting (the SRLG-joint) links. Therefore, SMLS does lower the blocking probability compared to both ULS and MLS schemes for both considered values of α. Nevertheless, we can notice that, for $\alpha = 0.1$ we get higher values of blocking probability than when $\alpha = 0.9$. This is because, for the values of α that are closed to zero, the TE metrics of the SRLG-joint links are closer to infinite number. Therefore, for smaller values of α, the blocking probability of SMLS scheme converges to the blocking probability of MLS scheme. Finally, we can observe that, even setting α to 0.1, the blocking probability of SMLS scheme is from ~4 times (at the low failure dynamicity) to ~1.5 times (at the high dynamicity) lower than when MLS scheme is deployed.

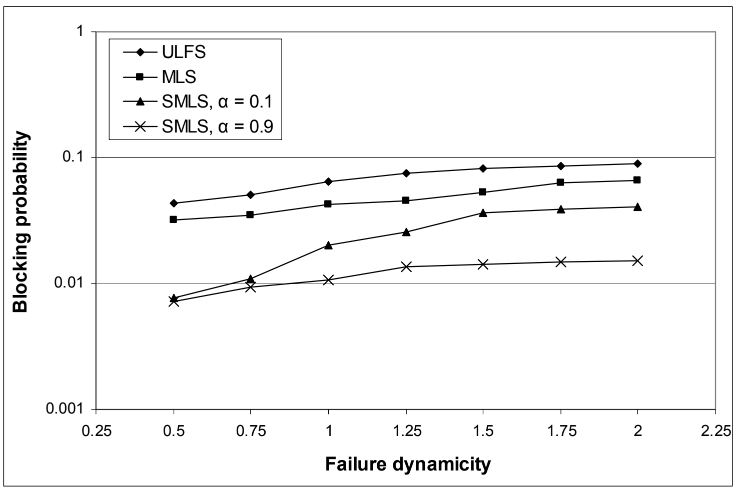


Figure 38. Connection blocking probability

Connection survivability

Figure 39 shows the connection survivability as a function of the failure dynamicity. The worst survivability is obtained deploying the ULFS scheme. This is because the ULFS scheme does not allow a path to change the layer, and, thus, there is no collaboration between the layers. Therefore, this algorithm may lead to very long paths, and, thus, the probability that both, working and backup path are affected by two failures is higher. The highest survivability is obtained using the MLS scheme. This is because this scheme assures a 100% survivability having a single link failure, satisfying SRLG-disjointness when computing the backup path. Finally, the proposed scheme outperforms the ULS scheme, but it still gives worse result comparing with MLS scheme. This is because this scheme is more flexible in terms of changing the layer at intermediate nodes, and, thus, it provides higher survivability than the ULS scheme. However, since the SMLS scheme does not fully exclude the SRLG-joint links from the network graph, it is very likely that a single link failure affects both the working and the backup paths of the same connection. Hence, the attained survivability by SMLS is worse comparing to the MLS scheme. Similarly to the blocking probability, the lower the value of α is, the survivability converges much faster to the survivability of MLS scheme. Nevertheless, one can still observe that even with a very high dynamicity, the difference between survivability of MLS and SMLS scheme is very small (i.e., from ~0.7% at low to ~2% at high failure dynamicity) for both, α = 0.1 and α = 0.9.

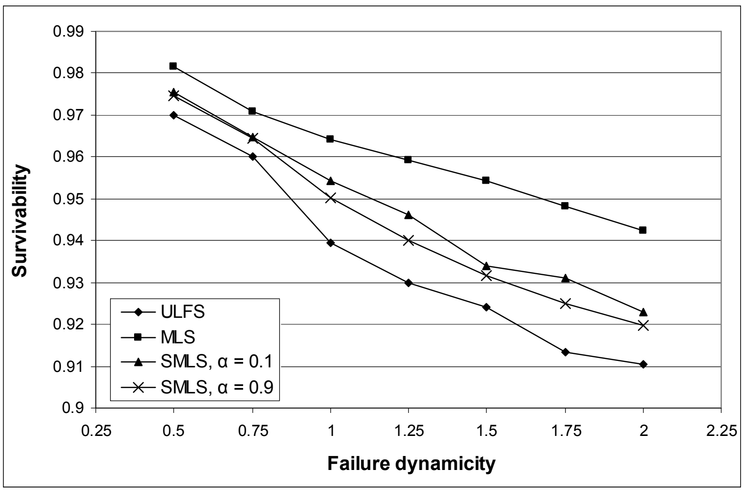


Figure 39. Connection survivability

4.6.5 Conclusion: SRLG-scaled dedicated protection scheme for MLN

The main goal of this contribution was to devise an e2e protection scheme that would decrease the blocking probability due to SRLG-disjointness constraints. Therefore, we have proposed a novel scheme based on the computation of working and backup path using the Suurballe-based algorithm, while also accounting for specific constraints of the optical technology and the deployment of sparse regenerators. The proposed scheme is more flexible by means of being able not to fully exclude all the SRLG-joint links, while still favoring the SRLG-disjoint links. It has been shown that, with the proposed SRLG-scaled multi-layer protection scheme significantly lower blocking probability can be obtained, comparing to the ULFS and MLS alternative schemes. On the other hand, it performs worse in terms of survivability when compared to the MLS scheme, at very high failure dynamicity. We have demonstrated that, deploying the proposed scheme, we achieve a good trade-off between the blocking probability and the survivability.

4.7 Chapter Summary and Conclusions

In this Chapter, we have first reviewed and classified the main approaches for recovery in a MLN. Then, we addressed the main challenges regarding e2e dedicated protection in a CO Ethernet over WSON controlled by GMPLS unified control plane. Finally, the thesis contributions regarding this topic are presented.

As a first contribution, we have compared two protection schemes where either link or SRLG disjointness is satisfied during the backup path computation. It has been shown that SDS scheme significantly improves the network survivability, but at the expense of achieving higher connection blocking probability.

The second contribution presents a scheme that aims at balancing the network resource usage at both involved layers, accommodating the connection requests over the least congested TE links. The scheme minimizes the number of connections affected by failures by ensuring a full SRLG-disjointness. Even with very high dynamicity of link failures, almost 100% of survivability is achieved, attaining low blocking probability.

The novelty in our third contribution regarding e2e dedicated protection is that it presents a MLN recovery strategy which successfully reduces the connection blocking probability due to the working SRLG joint links that are systematically excluded from the graph when computing the backup path. Even though the SRLG-joint links are not temporary excluded from the network graph, the scheme still favors the SRLG-disjoint links. Adjusting the value of α, the scheme achieves a nice trade-off between fully protected and unprotected MLN.

Chapter 5

Conclusions and Future Work

In this thesis dissertation, we have focused on concrete problems and challenges in the context of MLN constituted by two switching layer technologies, namely, CO-Ethernet and WSON, controlled by a GMPLS unified control plane. Specifically, we have studied two main issues. The first one deals with the problem of conceiving an efficient path computation algorithm that can consider the specifics of the underlying optical technology in order to improve the performance of the network with dynamic LSP provisioning, while the second one deals with the dedicated path protection schemes and mechanisms. In this section, we outline the main obtained results in this dissertation. Finally, we present some of the future work topics.

5.1 Conclusions

In Chapter 1, we have explained the necessity of moving towards a MLN composed of two switching layers, CO-Ethernet and WSON, in order to satisfy the ever-growing IP traffic demand and to overcome the scalability problems in the current transport network. Deploying such a high-capacity MLN, it is needed to have an intelligent mechanism which performs automated and rapid service provisioning, efficiently taking traffic engineering decisions and strategies. This issue has led to the development of GMPLS control plane that supports a variety of network technologies, offering dynamic cooperation among different switching layers. In this regard, in Chapter 1, we have presented the motivations for choosing this topic for the thesis dissertation. Finally, we have presented the thesis organization and the research contributions in terms of the produced publications in international, national and peer-reviewed journal papers.

After an overview of the main features and mechanisms supported by the GMPLS unified control plane handling the integration of both switching capabilities, provided in Chapter 2, we have concentrated on detailing the main contributions and achievements of the thesis. Specifically, this thesis faces up two challenges: the online path computation algorithms for dynamic service provisioning and the automatic LSP provisioning in a MLN infrastructure (Chapter 3), and the dedicated path protection schemes and mechanisms (Chapter 4).

Regarding online path computation algorithms, we have presented three contributions. The aim of the first contribution is to decrease the amount of signaling control processing due to the consecutive creations and deletion of a given FA TE link connecting the same pair of nodes. Through selected simulations, we have proved that the adoption of the *FA TE link timer* enables reducing the signaling control overhead caused by frequent LSC LSP establishment. Consequently, the average connection setup delay is slightly lower when the timer is deployed. Finally, we have confirmed that, increasing the timer value, the MLN converges to a network with statically configured VNT, since the established FA TE links remain in the network even if they are not being used by upper-layer (i.e., L2SC) connections.

The second contribution deals with different approaches for the VNT reconfiguration, namely: *dynamic*, *semi-dynamic* and *virtual*. The main objective of this contribution is to compare the three approaches regarding the connection blocking probability and setup delay. In this regard, we have proposed a dynamic Dijkstra-based path computation algorithm, where the VNT can be formed by both, virtual and FA TE links. It has been shown that the *dynamic* approach provides the lowest blocking probability, since the resources at both involved layers are occupied only when it is needed. In other words, when connections are torn down, associated resources are released. Nevertheless, due to the elevated dynamicity of the creation / deletion of FA TE links, this approach tends to perform the worst in terms of the average connection setup delay. However, from the obtained results, we could see that the difference between the average setup delays of the considered approaches is very small. The highest connection blocking probability is obtained applying the semi-dynamic VNT reconfiguration, since the resources are reserved for pre-established FA TE links even if they were not optimally selected. Finally, in light of the results, we can conclude that a trade-off between the connection blocking and setup delay should be taken into account when choosing a VNT configuration approach.

Finally, as the third contribution, we have presented an online path computation algorithm that chooses a route according to the state and the occupancy of the network elements (i.e., links and nodes). The algorithm tries to accommodate a requested connection over the least congested links and nodes. Moreover, some constraints were taken into account such as the WCC on the optical path segments as well as the adequate switching capability adaptation between the layers along the computed routes. It has been shown that the proposed algorithm provides a good and efficient usage of the network resources at both layers, attaining a lower blocking probability when compared with well-known MLN routing algorithms found in the literature.

Chapter 4 is dedicated to recovery methods and mechanisms. Concretely, we have focused on dedicated path protection within MLN. In this context, we have outlined the challenges of this kind of protection. This task allowed us to clearly identify and classify the proposed MLN path protection solutions found in the literature. Consequently, we have devised novel recovery path computation algorithms and solutions which are exhaustively simulated and benchmarked. The results are obtained considering a very high dynamicity of the link failure events in order to stress the network behavior when these two schemes are deployed.

Firstly, through a set of simulations, we have shown the advantages and drawbacks of using *link-* and *SRLG- disjoint* path protection schemes in MLN.

Although with the *link-disjoint* protection scheme we have obtained a lower connection blocking probability, this type of protection still provides worse survivability, since it does not assure the full SRLG-disjointness, but only link-disjointness. Therefore, we can conclude the scheme should be chosen taking into account the trade-off between the survivability and the connection blocking probability.

Then, we have proposed a protection scheme which favors the usage of the least congested links (i.e., larger amount of unused resources) when establishing both, working and backup path. Moreover, taking into account the number of traversed physical hops, we have reduced the potential creation of routing loops (i.e., two FA TE links share at least one physical link). It has been shown that the proposed scheme attains a high level of survivability, maintaining the connection blocking probability low.

Finally, the Section 4.6 addresses the problem of the increased blocking probability due to the constraint of finding routes guaranteeing the SRLG-disjointness when computing the backup paths. The proposed SRLG-scaled scheme does not remove the SRLG-joint links from the network graph when computing the backup path for a specific working path, but it still favors the utilization of SRLG-disjoint links. By doing so, the proposed scheme does lower the connection blocking probability, but at the expense of slightly increasing the survivability, comparing with the algorithm that eliminates the SRLG-joint links between the working and the backup paths.

5.2 Future work

There are many further issues regarding GMPLS unified control of CO Ethernet over WSON that have remained open. In that sense, one of the possible research lines for future work includes the shared path protection schemes and mechanisms, since this recovery type is more capacity-efficient comparing with the dedicated path protection schemes. Moreover, the research would contain different protection schemes applied on the upper-layer connections having different priority types. One of the possible ideas is to conceive and develop a protection scheme in which, the high-level priority connections are protected deploying a dedicated protection scheme, while the backup paths of low-level priority connection are shared among the working paths of other low-level priority connections (i.e., as long as the corresponding working paths are SRLG disjoint). The goal is, thus, to conceive a scheme / mechanism which would combine the fast recovery time of dedicated path protection with capacity efficiency of shared path protection. Another possible challenge is to develop efficient and fast protection schemes for a MLN, providing recovery from multiple network failures.

Appendix

In this Appendix, we present the design and implementation details of the selected functions of the multi-layer GMPLS control plane (signaling and routing) simulator used in this PhD thesis. As stated before, the simulator is deployed in the OPNET Modeler [4]. The source code is written in programming languages C and C++.

A. Introduction to OPNET modeling domains

The OPNET Modeler is a network simulator tool for analyzing and designing communication networks, devices, protocols, and applications for researcher and development. It provides an elevated abstraction level, ranging from GUI representation of a network topology, nodes and links, to C and C++ programming code for the implementation of hardware and software subsystems, communication protocols, algorithms, etc. There are three modeling domains in OPNET:

- The *network modeling domain*. This domain defines the network topology that is composed of *node instances* (e.g., controlled packet and optical switches) and links (e.g., Ethernet links, optical fibers). Within a network, there can be many *node instances* belonging to the different or the same *node model*. Network topologies are developed in the *Project Editor* that can provide different maps and/or administrative system segmentation (i.e., the world, a specific country, city, etc.), where the network elements will be placed.
- The *node modeling domain*. The capabilities of every network node are defined within this domain. The *node models* are developed in the *Node Editor*. Each *node modeling domain* is composed of smaller building blocks called *node modules*. There are two types of node modules: the ones that can only be configured through a set of built-in parameters (such as receivers and transmitters) through OPNET GUI, and the ones that are user-programmable. The former ones allow a node to be attached to communication links in the network domain. The latter ones, called *processors*, define the behavior of the node. There can be one or more processors within a node, belonging to a *process modeling domain* of that node.
- The *process modeling domain*. Within a *process modeling domain*, the processes are created and destroyed dynamically, during the simulation. The processes are developed in the *Process Editor.* Within each process, there are various blocks: Header Block, Function Block, Temporary Variable Block, State Variable Block, etc. In the blocks are stored the declarations of the variables and the functions needed for that process or

the processes of other nodes if so defined. Moreover, within the *process modeling domain* we define statistics we want to gather during a simulation run. The process may enter several *states*. Each state has *Enter* and *Exit state*, which are *Executives* that specify series of actions that a process implements when a node enters / exits the corresponding *state*. The executives are written in C or C++ programming languages. Typical actions implemented within a *state* include modifying state information, creating, sending and receiving messages, updating the messages contents, updating statistics, etc. Below every state, it is written how many C/C++ code lines are in the corresponding *Enter* and *Exit state*.

The relations between these three modeling domains are shown in the Figure 40.

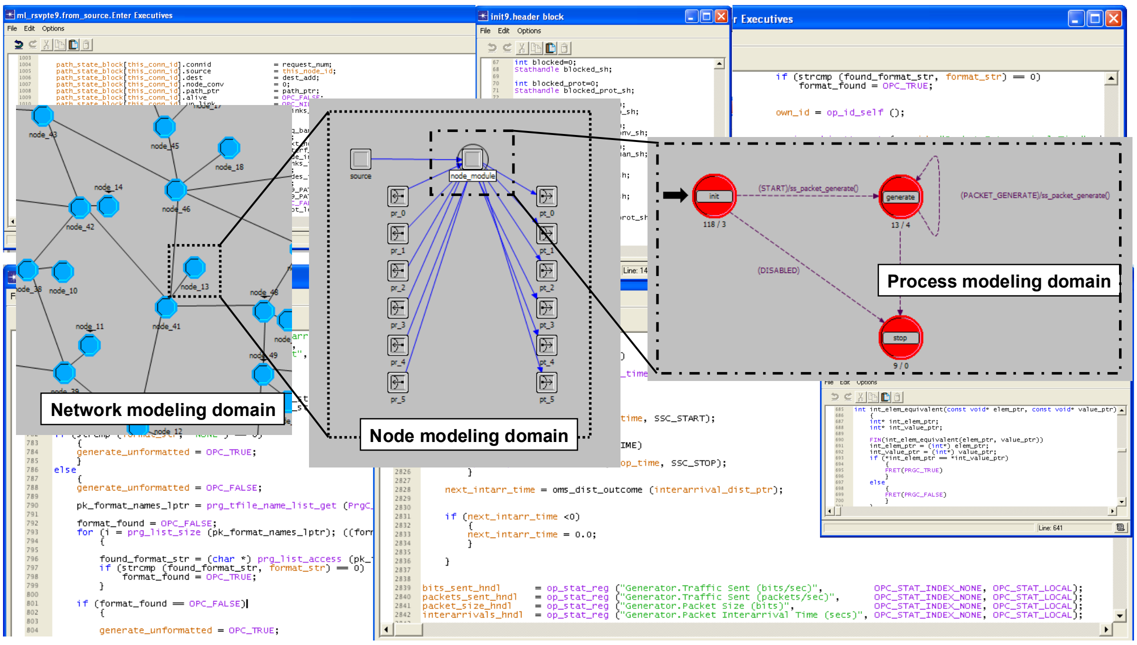


Figure 40. OPNET network modeling domains

B. Model overview

In the simulator, we have implemented some of the GMPLS RSVP-TE protocol functionalities for the provisioning of CO Ethernet services over WSON networks. The following RSVP-TE messages have been implemented: Path, Resv, PathTear, PathError and Notification messages [5]. In each of them, simplified objects are used. The routing protocol functionalities are not implemented, and it is assumed that the information is updated automatically once a network change (e.g., LSP establishment/tear down) occurs. We assume that the control plane and the data plane are congruent, that is, having the same network topology, where a GMPLS controller is responsible for every network node (i.e., either CO Ethernet or OXC).

We have entirely implemented two objects needed for the realization of the research objectives: the *initialization node* and the *network nodes*. For the same purposes, we have adapted the existing *network link* model. Figure 41 shows one example of a simple network containing these three objects. The following section provides a brief description of each of them.

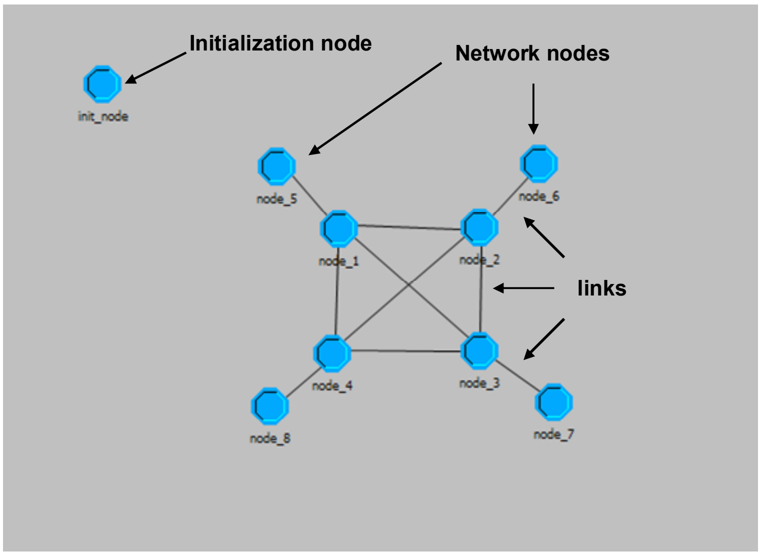


Figure 41. OPNET network objects

Initialization node

The initialization node is disconnected from the rest of network links, and it contains a module with a single *state process model*. The process defined in this *process model* is invoked only once, immediately after a simulation is executed.

In the *Header Block,* the global variables are declared (i.e., number of optical links, supported wavelengths, switching capabilities of the link interfaces, number of network nodes, etc.). These variables can be also used by the processes of the *network nodes*. The initialization node *Enter Executive* contains the code to be run at the simulation startup to initialize the global variables. Moreover, the following tasks are performed:

- Network elements discovery (i.e., network nodes and links), connectivity and network topology distribution.
- Initialization of the global attributes which will be accessed by the modules of the network nodes (e.g., mean connection inter-arrival time, mean connection duration, etc.).
- Initialization of the array of the TE links (i.e., *links_in_net*). The array of elements are structures, where an element represents simplified TE link attributes (i.e., link type, link ID, local interface ID, remote interface ID, TE link metric, available bandwidth, etc).

- Building an *NxNxM* array (i.e., *topology_array*) representing the node connectivity (i.e., network graph). In the array, *N* is the number of the network nodes, and *M* represents a positive integer presenting the number of links between the same pair of nodes. The array is tridimensional, since two nodes can be connected with more than one link (e.g., in the MLN context, several FA TE links). If two nodes are connected, the array element has value 1, otherwise it is equal to infinity. The array is updated every time a new FA TE link is established / eliminated.
- Building the *NxNxM* array (i.e., *TE_metrics_array*), where is stored the link TE metric (link cost). The array is updated every time a new FA TE link is established / eliminated.
- Creating the handles for the statistics we want to calculate (e.g., connection blocking probability, network survivability, wavelength channel usage, setup delay, etc.).
- Initialization of the connection list where the information about the active LSPs is stored (i.e., connection ID, source node ID, destination node ID, requested bandwidth, etc).

Network node

The *network nodes* we consider represent single switching capable nodes, constituting a MLN / MRN. That is, they can be either CO Ethernet switches or OXCs. In these nodes, the RSVP-TE signaling protocol is executed in order to dynamically establish and tear down either L2SC (Ethernet) or LSC (optical) end-to-end connections.

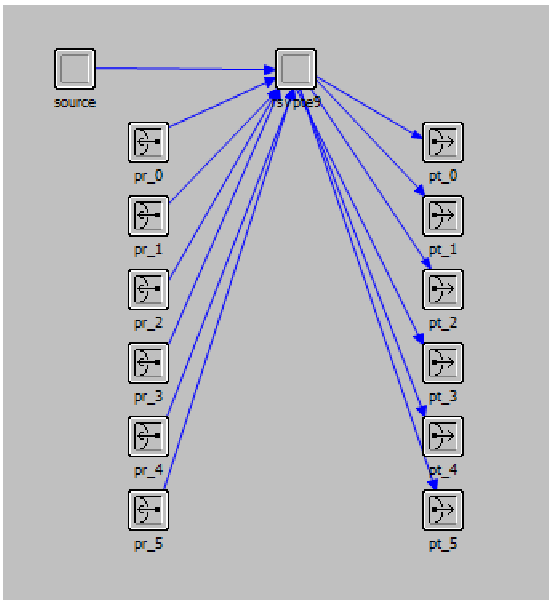


Figure 42. Network node structure

Node attributes (i.e., node ID, number of wavelength converters for OXC nodes, number of ports for CO Ethernet nodes, etc) are defined within the *model attributes* of the *node modeling domain*. A network node is constituted of several *node modules*, which are shown in the Figure 42. Apart of the six receivers and six transmitters that enable the communication of a network node with up to six neighboring nodes, there are two *node modules* of *processor* type: *source* and *RSVP-TE*. We have adapted the *source* OPNET module for our purposes, while we have developed the entire *RSVP-TE module*, including its functions and states.

In a *source processor*, connection requests are generated according to a Poisson process and with exponentially distributed duration. Moreover, a destination node is selected among the rest of the nodes following a uniform distribution. When a connection request is generated, the relevant information (e.g., selected destination node, holding time, etc.) is transferred to the *RSVP-TE module*. In case of protected MLN, the *source processor* also generates a failure of a physical link. The failed link is randomly chosen among the links between two OXCs, as it is explained in Section 4.4.2

The *RSVP-TE module* models the GMPLS signaling protocol. There are five states within this module, and they are shown in the Figure 43. As explained before, the numbers below each *state* represent the number of C/C++ code lines we have written, which define the series of actions that a process implements when a node enters / exits the corresponding *state*.

The initialization state (i.e., *init*) is executed only once during a simulation run for each node instance in the network. The processes performed within this state invoke right after the *initialization node* processes. In this state, the following actions are performed:

- Initialization of the global attributes for every network node instance (e.g., node ID, number of placed wavelength converters for the corresponding OXC, switching capability of the link interfaces that starts / terminates at that node, etc.);
- Initialization of the list of Path State Blocks (PSBs) maintained in each network node instance. Each PSB holds path state for a particular RSVP-TE session, storing the information about the corresponding connection (i.e., source node ID, destination node ID, incoming label, outgoing label, etc);
- Initialization of the destination nodes in case of pre-defined / pre-established virtual or FA TE links (Section 3.6);
- In case of the semi-dynamic approach (discussed in Section 3.6) where there are some pre-established FA TE links, this state triggers the LSC LSP establishment. To this end, firstly the path computation algorithm is executed to find a feasible route to the destination node. Next, the signaling objects carried in the RSVP-TE Path message used to set up the connection are conveniently set. Such information is collected in the corresponding new node PSB to store the status and manage the status of

each established/requested connection traversing a specific node. Finally, the state will initialize the signaling.

- In case of the virtual approach (Section 3.6), where there is a set of pre-defined but not established virtual links, at this state the path that the virtual link has to follow during the establishment of corresponding lower-layer LSP is determined. Moreover, the *links_in_net* array is updated including a new virtual link. Note that, in case of semi-dynamic approach, this action is performed once a FA TE link is established.

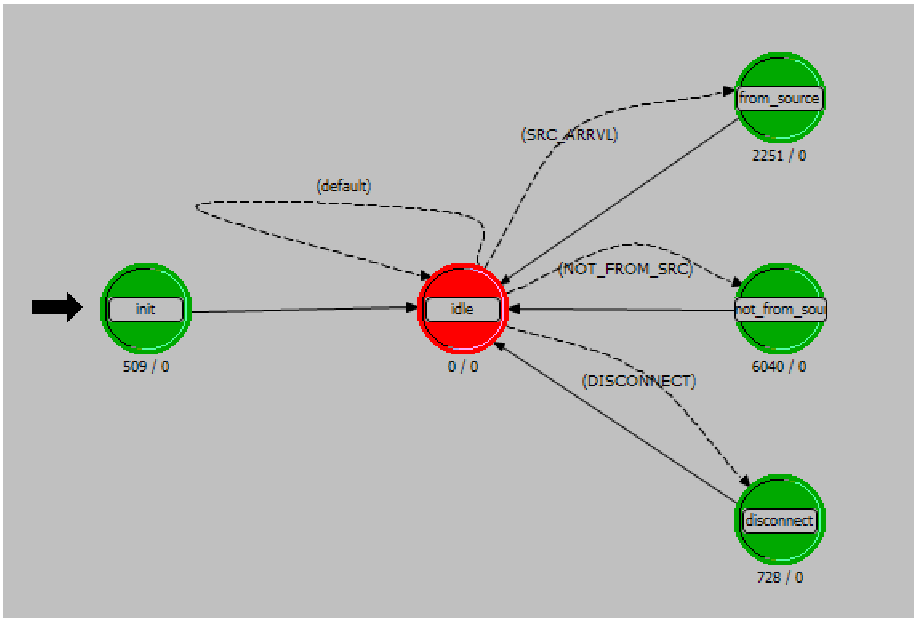


Figure 43. RSVP-TE module architecture

After the initialization state, the *RSVP-TE process* enters the *idle* state, where it is waiting for an event so it could either change to *from_source*, *not_from_source* or *disconnect states* (see Figure 43.).

From_source state

If a connection request from the Ethernet layer is generated and origins / comes from the *source processor*, the *from_source state* is reached. In this state, a path computation algorithm is executed aiming at finding a path between the current node in which the *process* is executed (i.e., the source node for that L2SC connection) and the destination node defined by the *source processor*. In case of dedicated protection, the path computation algorithm for the required backup path is also triggered. After finding the route, the objects (e.g., explicit route object, ERO) carried in the RSVP-TE Path message for that session are set. As described above, a new PSB is the created and added to the node's PSB list storing the associated information/attributes for that session (i.e., Session ID, source node ID, destination node ID, etc.).

The *from_source state* will then perform some control operations to check and validate that the computed path for the demanded connection request can be actually established. In this regard, it checks the remote interface switching capability of the outgoing link forming the computed route. If it is optical switching capability (LSC), the *process* will stop the establishment of the current session for L2SC LSP, and it will start a new one in order to establish the LSC LSP. The signaling procedure for LSC LSP establishment is explained in Section 2.2.6.

If either the path computation algorithm cannot find a route or the outgoing link does not have enough resources (e.g., unused bandwidth) to accommodate the requested connection, the L2SC LSP establishment will fail, and the number of blocked requests is increased for calculating the blocking probability. Otherwise, the *from_source state* will send the Path message towards the next node along the computed route.

Not_from_source state

Not_from_source state is the most complex state within the *RSVP-TE processor*. The RSVP-TE process reaches this state when a node gets a signaling control packet from some other node (i.e., not from the *source processor*). The behavior of this state depends on the RSVP-TE message type that this node has received.

If the node gets a **Path message**, the attributes carried into the message are stored in a newly created PSB for that connection. If the node is at the Ethernet layer, the *not_from_source state* checks if there is a sufficient unreserved bandwidth on the outgoing link. If there is no sufficient available bandwidth, a PathErr message is created and sent backwards to the previous node informing about the raised error. On the other hand, if the node is at the optical layer, the set of the acceptable wavelengths specified in the carried Label Set Object will be updated according to those one available on the next outgoing link. This is done in order to guarantee the wavelength continuity constraint from the source to the destination node. If the Label Set becomes an empty set (i.e., no continuous available wavelengths from the source), the request is blocked. Then, a PathErr message will be sent backwards to the source node. Conversely (i.e., Label Set is not empty), the Path message is transmitted to the next node according to the ERO object.

Once the Path message reaches the destination node, if there are enough resources at the upstream link to accommodate demanded connection, the state will perform the resource reservation (i.e., allocation of the requested bandwidth at the Ethernet layer or occupancy of the wavelength channel at the optical layer). Then, the state will create the RSVP-TE Resv message which will be sent upstream towards the source node. Otherwise, if there are no enough resources, a PathErr message is sent backwards and the connection will be closed, releasing the occupied resources.

If, at some point, a node detects that is at the boundary between two switching capabilities, the current session for the Ethernet layer is put “on hold”. Next, a new RSVP session is constructed to set up the underlying LSC LSP. To do this, the LSP_TUNNEL_INTERFACE_ID object, with the Node and the allocated Interface ID for the higher-layer FA TE link being created, is added to the newly created Path message, identifying the head-end of the created FA-LSP. After checking the available wavelengths on the outgoing link, the Path message is sent to the next node in order to establish the lower-layer LSP.

If the *not_from_source state* receives the RSVP-TE **Resv message**, it will update the node PSB associated to that connection with the information carried in the message. Then, the state checks whether there are still available resources on the upstream link. In case of LSC LSP establishment, the state will check if the selected wavelength is still available. This verification is essential in optical networks without wavelength converters, where the WCC needs to be ensured. In any case, if the wavelength is still available, the Resv message is propagated along the reverse path. In the contrary, the state creates and sends the PathErr message backwards and, a PathTear message is created and forwarded to release the allocated resources by the connection that is being established. If the Resv message successfully reaches the source, this node schedules the connection duration timer according to the computed holding time in order to tear down the connection. Moreover, the list of established L2SC LSPs, as well as the corresponding PSB, will be updated.

In case of the lower-layer LSP establishment, the LSC LSP is advertised as a new virtual FA TE link and, thus, the *links_in_net* array will be updated. Later on, the remaining available bandwidth on such a FA TE link can be used for the establishment of LSPs in the upper-layer. In this scenario, if the creation of a TE link was triggered dynamically, it will be released as soon as it is not used by any higher-layer connection. On the other hand, the statically established higher-layer TE links will behave as a permanent Ethernet links, and they will not be released even when are not being used. Finally, the session of the corresponding L2SC LSP that has been put on hold and triggered the establishment of the LSC LSP, will be again invoked following the standard GMPLS signaling procedure within MLN (Section 2.2.6).

If the *not_from_source state* receives a **PathTear message**, the resources occupied by the corresponding LSP will be released, and the RSVP-TE message will be sent downstream until it reaches the destination. If the *not_from_source state* receives a PathErr message, it will send upstream until it reaches the source node. If either a PathTear message reaches the destination or a **PathErr message** reaches the source node, the process is closed, releasing the occupied memory.

Apart from Path, Resv, PathErr and PathTear messages, there are two types of Notification messages: Failure Indication Message (FIM) and Repair Notification Message (RNM). When a link is hit by a failure, two nodes that are connected by the damaged bidirectional optical link will create as many as **FIMs** as LSC LSPs are established over such a link. This is resolved by checking the corresponding nodes' PSBs. Then, the nodes will send the created FIMs to the source node of every affected LSC connection. When a source node of each LSC connection gets the FIM message, it will notify to all the L2SC LSPs that use the affected FA TE links induced over the disrupted LSC LSP about the failure, sending newly created FIM messages to the source nodes of such L2SC connections. Finally, when a source node of a L2SC connection receives a FIM message, it needs to check if both working and backup paths are affected by two different failures. If this is the case, the connection is closed, the statistics are updated and memory for that connection is released (i.e., PSB).

Analogously, once a link is repaired, **RNM** messages are sent back to the source nodes of all the affected LSC and L2SC LSPs, notifying that the damaged link has been repaired. The RNM message is sent backwards in order to switch the traffic to the previously damaged working path and for the statistical purposes (i.e., to calculate how many connections have been repaired).

In order to better understand the *not_from_source* state of the *RSVP-TE process*, in Figure 44 are shown the simplified operations for the connection provisioning. The *border node* represent the boundary node between the CO Ethernet and WSON layer. For the sake of simplicity, the operations are shown in case of unprotected MLN. Therefore, FIM and RNM messages are not depicted.

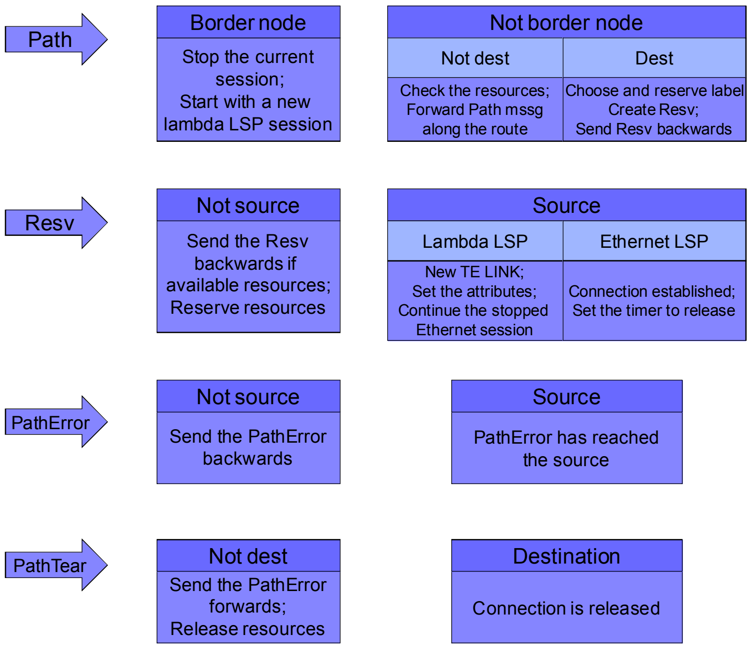


Figure 44. Simplified operations of *not_from_source* state

Disconnect state

When the Resv message reaches the source node, it schedules the connection duration timer in order to tear down the connection. When the timer expires at the connection source node, the *disconnect state* will be invoked (see Figure 43). The role of this state is to release the corresponding connection once the holding time expires. Therefore, the state will create a PathTear message, which will be sent towards the destination node in order to release the occupied resources across the path. Once the PathTear reaches the destination node, the process for that connection is closed.

It is worth mentioning that, every time a PathTear message needs to be sent to the next node using a FA TE link, the current node checks whether that FA TE link is being used by some other connections. If this link disposes the whole maximum bandwidth (i.e., no occupied resources), the FA TE link and the resources will be released, unless the link is established statically. Moreover, in case when the *FA TE link timer* is used (Section 3.5), the state will schedule the timer in order to postpone the release of such a TE link. Because of the simplicity of the Figure 44, that procedure is not shown within the PathTear message.

Network links

In the simulator, we assume that the physical links between two network nodes are bidirectional with interface switching capability of either L2SC or LSC. However, FA TE links and virtual links are unidirectional. In a protected MLN network, we take into account only link failure between two OXCs, as it is explained in Chapter 4.

We assume that the control links through which signaling messages are exchanged between teach pair of adjacent GMPLS controllers are fully reliable links, that is, without transmission errors. The propagation delays of such control links are considered with respect to the link distance.

C. Discrete Event Simulations

The evaluation of the network performance can be analyzed through Discrete Event Simulations (DES). We have created four reference network topologies in the *Project Editor* used in this thesis dissertation (Figure 5, Figure 12 and Figure 16).

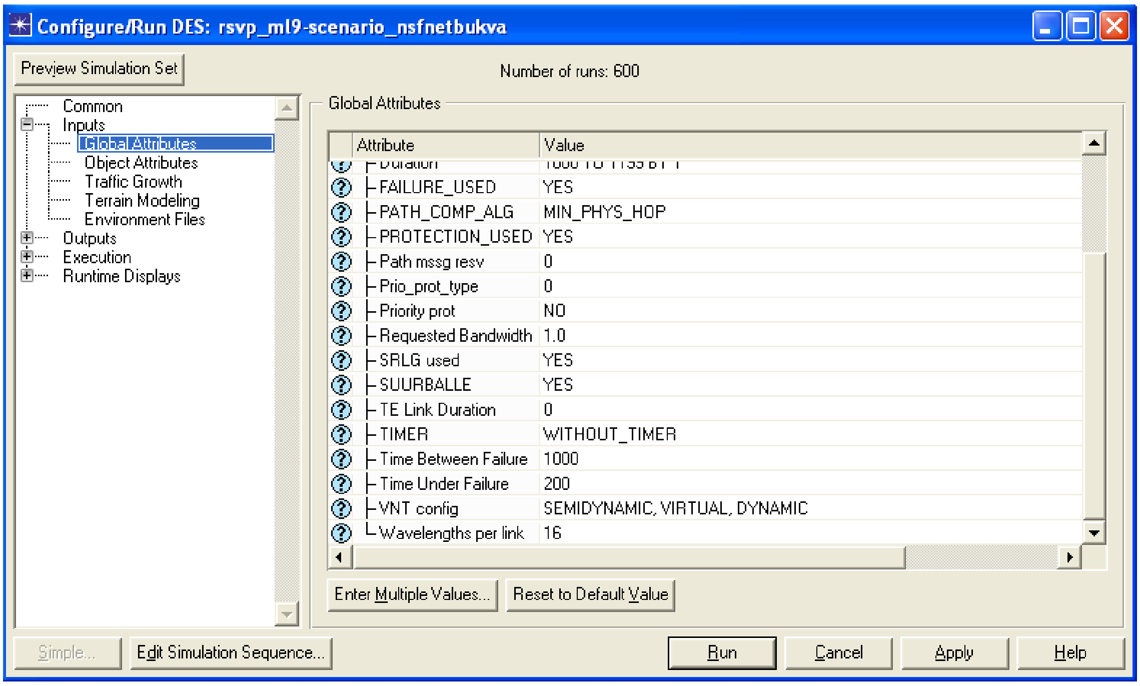


Figure 45. Example of DES configuration

After a network scenario (i.e., topology and network resources) and the targeted statistics to be collected are set up, the next step is to configure and to run the simulation. In the Figure 45, an example of the DES configuration is shown. In the DES window, we configure all the attributes required for the targeted network, initialization node process models (e.g., the path computation algorithm, the mean connection request holding time, mean connection request inter-arrival time, requested bandwidth, etc.). As the simulation runs, the requested statistics will be collected. Depending on the assigned attribute values (e.g., holding time attribute value, the type of the path computation algorithm used in the simulation run, etc.), a simulation can have one or more simulation runs.

Finally, once the simulation finishes, the results can be seen and displayed in the *Results Browser* that supports different collection methods and statistics types such as graphs, *OPNET parametric studies*, etc.

References

[1] IEEE 802.1ay "Provider Backbone Bridge Traffic Engineering"

[2] B. Niven-Jenkins, D. Brungard, M. Betts, N. Sprecher and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2012

[3] E. Mannie, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004

[4] www.opnet.com

[5] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003

[6] K. Shiomoto, D. Papadimitriou, J.-L. Le Roux, M. Vigoureux, and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, July 2008

[7] IEEE 802.1ah "Provider Backbone Bridges"

[8] IEEE 802.1ad, "Provider Bridges"

[9] D. Awduche, et. al "Requirements for traffic engineering over MPLS", IETF RFC 2702, Sept. 1999.

[10] S. Bryant and P- Pate, "Pseudo wire emulation edge-to-edge (PWE3) architecture", IETF RFC 3985, March 2005.

[11] E. Rosen, A. Viswanathan and R. Callon, "Multiprotocol Label Switching Architecture", IETF RFC 3031, Jan. 2001.

[12] P. Ashwood-Smith and L. Berger, "Generalized MPLS - CR-LDP Extensions", RFC 3472, January 2003

[13] L. Andersson, G. Swallow, "The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signaling protocols", RFC 3468, February 2003

[14] K. Kompella and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005

[15] H. Smit and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004

[16] J. Lang, "Link Management Protocol (LMP)", RFC 4204, October 2005.

[17] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku, and Y. Takigawa, "Dynamic multilayer routing schemes in GMPLS-based IP+optical networks," IEEE Commun. Mag., vol. 43, no. 1, pp. 108–114, Jan. 2005.

[18] JL. Le Roux and D. Papadimitriou, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 5339, September 2008

[19] R. Ramaswami and K. N. Sivarajan, “Design of logical topologies for wavelength-routed optical networks,” in Proc. IEEE J. Selected Areas in Communication,, vol. 14, pp. 840 - 851, June 1996.

[20] B. Mukherjee, D. Banerjee, S. Ramamurthy, and A. Mukherjee, “Some principles for designing a wide-area WDM optical network,” in Proc. IEEE/ACM Trans. Networking, vol. 4, pp. 684 - 696, Oct. 1996.

[21] Y. Bréhon, D. Kofman, M. Pióro, and M. Diallo, “Optimal Virtual Topology Design using Bus-Label Switched Paths”, in Proc. IEEE J. on Selected Areas in Communications, vol. 25, no. 5, pp. 1001-1010, June 2007.

[22] P. Pavon-Marino, R. Aparicio-Pardo, B. Garcia-Manrubia, and N. Skorin-Kapov, “Virtual topology design and flow routing in optical networks under multihour traffic demand”, in Proc. Photonic Network Communications, Vol. 19, Issue 1, pp 42-54, February 2010.

[23] K. Manousakis, K. Christodoulopoulos, E. Kamitsas, I. Tomkos, and E. A. Varvarigos, “Offline Impairment-Aware Routing and Wavelength Assignment Algorithms in Translucent WDM Optical Networks”, in Proc. Journal of Lightwave Technology, vol. 27, no. 12, pp. 1866 – 1877, June 2009.

[24] R. Dutta and G. N. Rouskas, “A Survey of Virtual Topology Design Algorithms for Wavelength Routed Optical Networks,” in Proc. Optical Net., vol. 1, no. 1, Jan. 2000, pp. 73–89.

[25] G. Agrawal and D. Medhi, “Lightpath Topology Configuration for Wavelength-routed IP/MPLS Newtworks with Time-Dependent Traffic”, in Proc. of IEEE Globecom 2006, pp. 1-5., Nov-Dec. 2006.

[26] A. Farrel, J.-P. Vasseur, and J. Ash, “A path computation element (PCE)-based architecture,” RFC 4655, Aug. 2006.

[27] F. Cugini, A. Giorgetti, N. Andriolli, F. Paolucci, L. Valcarenghi, and P. Castoldi, “Multiple path computation element (PCE) cooperation for multi-layer traffic engineering,” in Proc. OFC 2007, Mar. 2007, OWK5.

[28] N. Andriolli, F. Cugini, L. Valcarenghi, P. Castoldi, and A. Welin, “Virtual network topology manager (VNTM) and path computation element (PCE) cooperation in multi-layer GMPLS networks,” in Proc. of the OFC/NFOEC, 2009

[29] Y. Zhao, J. Zhang, H. Zhang, W. Gu, and Y. Ji, “Novel path computation element-based traffic grooming strategy in Internet protocol over wavelength division multiplexing networks,” IET Commun., vol. 5, no. 8, pp. 1138–1146, 2011.

[30] P. Fodor, G. Enyedi, G. Rétvári, and T. Cinkler, “Layer-preference policies in multi-layer GMPLS networks,” Photonic Network Commun., vol. 18, no. 3, pp. 300–313, Feb. 2009.

[31] W. Colitti, K. Steenhaut, D. Colle, M. Pickavet, J. Lemeire, and A. Nowé, “Integrated routing in GMPLS-based IP/WDM networks,” Photonic Network Commun., vol. 21, pp. 238–252, Oct. 2010.

[32] K. Zhu and B. Mukherjee, “On-line approaches for provisioning connections of different bandwidth granularities in WDM mesh networks,” in Proc. OFC 2002, Mar. 2002, pp. 549–551.

[33] M. Kordialam and T. V. Lakshman, "Integrated dynamic IP and wavelength routing in IP over WDM networks," in Proc. INFOCOM 2001, 2001, pp. 358–366.

[34] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," IEEE/ACM Trans. Netw., vol. 11, pp. 285–299, Apr. 2003.

[35] N. Andriolli, F. Cugini, P. Iovanna, G. Bottari, A. Bogoni, L. Valcarenghi, and P. Castoldi, "Introducing TE metrics to account for transponder and grooming resources in GMPLS multi-layer network," in Proc. of the IPR/PS, 2010.

[36] K. Shiomoto, E. Oki,W. Imajuku, S. Okamoto, and N. Yamanaka, "Distributed virtual network topology control mechanism in GMPLS-based multiregion networks," IEEE J. Sel. Areas Commun., vol. 21, pp. 1254–1262, Oct. 2003.

[37] D. Shimazaki, E. Oki, K. Shiomoto, and S. Urushidani, "Traffic-driven virtual network topology reconfiguration for GMPLS network," in Proc. of the IEICE General Conf., 2006.

[38] E. Oki, T. Takeda, J. L. Le Roux, and A. Farrel, "Framework for PCE-based inter-layer MPLS and GMPLS traffic engineering," RFC 5623, Sept. 2009.

[39] A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "Enhanced Dynamic VNT configuration in GMPLS Controlled Ethernet over WSON with timer-based lightpath holding time", in Proc. of ICTON 2011, June 2011

[40] A. Bukva, R. Casellas, R. Martínez, and R. Muñoz, "A dynamic on-line path computation algorithm for VNT configuration in GMPLS controlled multi-layer (Ethernet/WSON) network," in Proc. of HPSR 2011, July 2011, pp. 209–214.

[41] A. Bukva, R. Casellas, R. Martínez, R. Muñoz, "A Dynamic Path-Computation Algorithm for a GMPLS-Enabled Multi-layer Network", in Proc. of OSA Journal of Optical Comm. and Networking (JOCN), Vol. 4, No. 6, pp. 436-448, June 2012

[42] R. Hulsermann, S. Bodamer, M. Barry, A. Betket, C. Gauger, M. Jager, M. Kohn and J. Spath, "A set of typical transport network scenarios for network modeling", in Proc. of the 5th ITG - Workshop on Photonic Networks, May 2004.

[43] S. de Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz, "Pan-European optical transport networks: An availability-based comparison," Photonic Network Commun., vol. 5, no. 3, pp. 203–225, 2003.

[44] S. Ramamurthy and B. Mukherjee, "Survivable WDM Mesh Networks, Part I – Protection", in Proc. IEEE INFOCOM 1999, pp. 744-751

[45] A. Fumagalli and L. Valcarenghi, "Ip restoration vs. WDM protection: is there an optimal choice?" IEEE Network, vol. 14, no. 6, pp. 34-41, Nov./Dec. 2000.

[46] S. Ramamurthy, L. Sahasrabuddhe, and B. Mukherjee, "Survivable WDM Mesh Networks", in Proc. of IEEE Journal of Lightwave Technology, Vol. 21. no. 4. April 2003, pp. 870-882.

[47] O. Gerstel and R. Ramaswami, "Optical layer survivability - an implementation perspective", IEEE J. Sel. Areas in Commun., vol. 18, no. 10, pp. 1885-1899, Oct. 2000.

[48] J. Lang, B. Rajagopalan and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Recovery Functional Specification", RFC 4426, March 2006

[49] D. Papadimitriou and E. Mannie, "Analysis of Generalized Multi-Protocol Label Switching (GMPLS)-based Recovery Mechanisms (including Protection and Restoration)", RFC 4428, March 2006

[50] Y. Ye et al., "A Simple Dynamic Integrated Provisioning/Protection Scheme in IP over WDM Networks", IEEE Commun. Mag., November 2001.

[51] M. Pickavet, P. Demeester, D. Colle, D. Staessens, B. Puype, L. Depré and I. Lievens, "Recovery in Multilayer Optical Networks", in Proc. of IEEE Journal of Lightwave Technology, Vol. 24. no. 1. January 2006, pp. 122-134.

[52] P. Chołda and A. Jajszczyk, "Recovery and Its Quality in Multilayer Networks", in Proc. of IEEE Journal of Lightwave Technology, Vol. 28. no. 4. January 2010, pp. 372-389.

[53] Y. Qin, L. Mason and K. Jia, "Study on a Joint Multiple Layer Restoration Scheme for IP over WDM Networks", in Proc. of IEEE Network Magazine, pp. 43-48, March / April 2003.

[54] T. Cinkler, P. Hegyi, G. Geleji, J. Szigeti and A. Ladanyi, "Adaptive Multi-Layer Traffic Engineering with Shared Risk Group Protection", in Proc. of IEEE ICC, pp. 5367 – 5371, May 2008

[55] J. Segovia, E. Calle and P. Vila, "Availability Analysis of GMPLS Connections based on Physical Network Topology", in Proc of IEEE ONDM 2008, March 2008

[56] Y. Yamada, H. Hasegawa and K. Sato, "Survivable Hierarchical Optical Path Network Design With Dedicated Wavelength Path Protection", in Proc. of IEEE Journal of Lightwave Technology, Vol. 29. no. 21. Nov. 2011, pp. 3196 - 3209.

[57] C. Ou, K. Zhu, H. Zang, J. Zhang, H. Zhu, L. H. Sahasrabuddhe and B. Mukherjee, "Traffic Grooming for Survivable WDM Networks – Dedicated Protection", in Proc. of OSA Journal of Optical Comm. and Networking (JOCN), Vol. 3, No. 1, pp. 50-74, January 2004

[58] M. Tornatore, G. Maier and A. Pattavina, "Variable Aggregation in the ILP Design of WDM Networks with Dedicated Protection", in Proc. of OSA Journal of Optical Comm. and Networking (JOCN), Vol. 9, No. 4, pp. 419-427, December 2007

[59] R. Martínez, R. Muñoz, R. Casellas, J. Comellas and G. Junyent, "Experimental shared path protection algorithms in distributed all-optical GMPLS-based networks", in Proc. of Design and Reliable Communication Networks (DRCN) 2007, pp. 1-7, Oct 2007.

[60] Q. Guo, P. Ho, A. Haque and H. T. Mouftah, "Availability-Constrained Shared Backup Path Protection (SBPP) for GMPLS-Based Spare Capacity Reprovisioning", in Proc. of IEEE ICC 2007, pp. 2186 – 2191, June 2007

[61] L. Velasco, S. Spadaro, J. Comellas and G. Junyent, "Wavelength Selective Switches time requirements for shared path protection in ASON/GMPLS rings", in Proc. of ECOC 2008, Sept. 2008

[62] L. Velasco, F. Agraz, R. Martínez, R. Casellas, S. Spadaro, R. Muñoz and G. Junyent, "GMPLS-based Multidomain Restoration: Analysis, Strategies, Policies and Experimental

Assessment", ", in Proc. of OSA Journal of Optical Comm. and Networking (JOCN), Vol. 2, No. 7, pp. 427 - 441, July 2010

[63] I. Cerutti, A. Giorgetti, N. Andriolli, P. Castoldi, R. Muñoz, S. Ruepp, R. Casellas, R. Martínez, A.V. Manolova, "Restoration in GMPLS-based Wavelength Switched Optical Networks with Limited Wavelength Converters", in Proc. of Photonics in Switching, Sept. 2009.

[64] X. Cui, J. Wang, X. Yao, W. Liu, H. Xie and Y. Li, "Optimization of Multilayer Restoration and Routing in IP-over-WDM Networks", In: Proc. of OFC/NFOEC, pp. 1–10, 2008

[65] P. Han-Ho and H. T. Mouftah, "A framework for service-guaranteed shared protection in WDM mesh networks", IEEE Commun. Mag., vol. 40, no. 2, pp. 97-103, Feb. 2002.

[66] L. Ruan and Z. Liu, "Upstream Node Initiated Fast Restoration in MPLS Networks", in Proc. IEEE ICC 2005, vol. 2, pp. 959 – 964, May 2005

[67] W. Tavernier, D. Colle, M. Pickavet and P. Demeester, "GMPLS-controlled Ethernet segment protection", in Proc . of IEEE ECOC 2007, September 2007

[68] N. Sambo, I. Cerutti, F. Cugini, A. Giorgetti, L. Valcarenghi and P. Castoldi, "Segment Restoration Scheme with QoT-guarantees in GMPLS-controlled Translucent Networks", in Proc. of ECOC 2008, Sept. 2008.

[69] E. Oki, N. Matsuura, K. Shiomoto and N Yamanaka, "A Disjoint Path Selection Scheme with SRLG in GMPLS Networks", in Proc. of IEEE HPSR 2002, pp. 88-92.

[70] J.W. Suurballe, "Disjoint Paths in a Network," Networks, Vol.4, pp. 125-145, June 1974.

[71] Dunn, W. Grover, and M. MacGregor, "Comparison of k-shortest paths and maximum flow routing for network facility restoration," IEEE J. Selected Areas in Communications, vol. 12, pp. 88–99, Jan 1994

[72] M. German, A. Castro, X. Masip-Bruin, M. Yannuzzi, R. Martínez, R. Casellas and R. Muñoz, "On the challenges of finding two link-disjoint lightpaths of minimum total weight across an optical network", in Proc. of 14th European Conference on Networks and Optical Communications (NOC 2009), June 2009

[73] M. H. MacGregor and W. D. Grover, "Optimized k-shortest-paths Algorithm for Facility Restoration", Software—practice and experience, vol. 24(9), pp. 823–834, September 1994

[74] G. Mohan and C. Siva Ram Murthy. "Lightpath Restoration in WDM Optical Networks", IEEE Network Magazine. vol. 14. no. 6, November / December 2000, pp. 24-32.

[75] V. Anand and C. Qiao, "Dynamic Establishment of Protection Paths in WDM Networks Part-I", Proc. IEEE ICCCN, October 2000, pp. 198- 204.

[76] Kar. M. Kodialam. and T. V. Lakshman, "Routing Restorable Bandwidth Guaranteed Connections Using Maximum 2-Route Flows'', in Proc. IEEE INFOCOM 2002, pp 113-121.

[77] Y. Ye, S. Dixit and M. Ali, "On joint protection/restoration in IP-centric DWDM based optical transport networks", in Proc. IEEE Communications Magazine, Vol. 38, pp. 174 – 183, Jun 2000

[78] Weidong C., I. Stoica and Randy H. Katz, "Backup Path Allocation Based On A Correlated Link Failure Probability Model In Overlay Networks", in Proc. IEEE ICNP'02, 2002

[79] Q. Zheng and G. Mohan, "Multi-layer protection in IP-over-WDM networks with and with no backup lightpath sharing", Computer Networks, 2006, pp. 301-316

[80] A. Urra, E. Calle and Jose L. Marzo, "Enhanced Multi-Layer Protection in Multi-Service GMPLS Networks", in Proc. IEEE Globecom 2005.

[81] B. Puype, J. P. Vasseur, A. Groebbens, S. De Maesschalck, D. Colle, I. Lievens, M. Pickavet, and P. Demeester, "Benefits of GMPLS for Multilayer Recovery", in Proc. of IEEE Comm. Magazine, pp. 51-59, July 2005.

[82] A. Bukva, R. Casellas, R. Martínez and R. Muñoz, "A dynamic path computation algorithm for e2e dedicated protection in a GMPLS controlled multilayer (Ethernet/WSON) network", in Proc. of International Workshop on Reliable Networks Design and Modeling (RNDM 2011), October 2011

[83] A. Bukva, R. Casellas, R. Martínez and R. Muñoz, "An On-Line Path Computation Algorithm for a Protected GMPLS Enabled Multi-Layer Network", in Proc. of 17th European Conference on Networks and Optical Communications (NOC 2012), June 2012

[84] K.Kompella, Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005