ADVERTIMENT. La consulta d’aquesta tesi queda condicionada a l’acceptació de les següents condicions d'ús: La difusió d’aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d’investigació i docència. No s’autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d’un lloc aliè al servei TDX. No s’autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you’re accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it’s obliged to indicate the name of the author

On the resource abstraction, partitioning and composition for virtual GMPLS-controlled multi-layer optical networks

TESI
per a l'obtenció del títol de
Doctor per la Universitat Politècnica de Catalunya

per
Ricard Vilalta Cañellas

Director: Dr. Raül Muñoz González

Centre Tecnològic de Telecomunicacions de Catalunya

Ponent: Dr. Gabriel Junyent Giralt

Universitat Politècnica de Catalunya
Departament de Teoria del Senyal i Comunicacions

Composició del Tribunal

President: Dr. Pablo Pavón Mariño
Secretari: Dr. Salvatore Spadaro
Vocal: Dr. Shuping Peng

Per a la Joana i l'Aina

Resum

La gestió de xarxes òptiques virtuals permet la provisió dinàmica de xarxes dedicades a sobre la mateixa infraestructura de xarxa i ha cridat molt l'atenció als proveïdors de xarxes. Els requisits de xarxa (per exemple la qualitat de servei, els acords de nivell de servei o la dinamicitat) són cada cop més astringents per a les aplicacions emergents d'elevat ample de banda i dinàmiques, que inclouen per exemple la reproducció en temps real de vídeo d'alta definició (telepresència, televisió, telemedicina) i serveis d'informàtica en núvol (còpies de seguretat en temps real, escriptori remot). Aquests requisits poden ser assolits a través del desplegament de serveis de infraestructura dinàmics per construir xarxes òptiques virtuals (VON, en anglès), fet que és conegut com a infraestructura com a servei (IaaS). La internet del futur hauria de suportar dos entitats diferenciades: els proveïdors d'infraestructures (responsables de gestionar la infraestructura física), i els proveïdors de serveis (responsables dels protocols de xarxa i d'oferir els serveis finals). D'aquesta forma els proveïdors de serveis podrien sol·licitar i gestionar en funció de les necessitats xarxes òptiques virtuals dedicades i específiques per les aplicacions.

Les tecnologies de virtualització de xarxes òptiques virtuals permeten la partició i composició de infraestructura de xarxa (nodes i enllaços òptics) en recursos virtuals independents que adopten les mateixes funcionalitats que els recursos físics. La composició d'aquests recursos virtuals (nodes i enllaços òptics virtuals) permet el desplegament de múltiples VONs. Una VON no sols està composada per un pla de transport virtual, sinó també per un pla de control virtual, amb l'objectiu d'incorporar les funcionalitats necessàries a la VON(provisió de connexions automàtiques i recuperació (protecció/restauració), enginyeria de tràfic, etc.).

Aquesta tesis es centra en la virtualització de xarxes òptiques amb tres objectius principals. El primer objectiu consisteix en el disseny, implementació i avaluació de l'arquitectura i els protocols i interfícies necessaris per la virtualització de xarxes encaminades a través de la longitud d'ona i controlades per GMPLS. També inclou la introducció d'un gestor de recursos per desplegar xarxes òptiques virtuals de forma dinàmica. La introducció d'aquest gestor de recursos implica la necessitat d'una gestió dels recursos virtuals i d'algortimes d'assignació de recursos per a la utilització òptima dels recursos físics. A més el gestor de recursos ha de ser capaç del desplegament dels recursos assignats, incloent un pla de control GMPLS virtual independent per a cada VON desplegada. Finalment, aquest objectiu inclou la introducció de mecanismes de virtualització per a xarxes elàstiques òptiques (EON, en anglès).

El segon objectiu és el disseny, la implementació i l'avaluació experimental d'una arquitectura de sistema per oferir xarxes MPLS-TP virtuals controlades per GMPLS sobre una infraestructura

WSON compartida. Per això, aquesta tesis també es centra en el disseny i desenvolupament d'un node MPLS-TP que ha estat desplegat al demostrador ADRENALINE, al CTTC.

Finalment, el tercer objectiu és la composició de múltiples xarxes òptiques virtuals en dominis de control heterogenis (GMPLS i OpenFlow). Un gestor de recursos multi-domini ha estat dissenyat, implementat i avaluat.

Resumen

La gestión de redes ópticas virtuales permite la provisión dinámica de redes dedicadas encima la misma infraestructura de red y ha llamado mucho la atención a los proveedores de redes. Los requisitos de red (por ejemplo la calidad de servicio, los acuerdos de nivel de servicio o la dinamicidad) son cada vez más estringentes para las aplicaciones emergentes de elevado ancho de banda y dinámicas, que incluyen por ejemplo la reproducción en tiempo real de vídeo de alta definición (telepresència, televisión, telemedicina) y servicios de computación en la nube (copias de seguridad en tiempo real, escritorio remoto). Estos requisitos pueden ser logrados a través del despliegue de servicios de infraestructura dinámicos para construir redes ópticas virtuales (VON, en inglés), hecho que es conocido como infraestructura como servicio (IaaS). La internet del futuro tendría que soportar dos entidades diferenciadas: los proveedores de infraestructuras (responsables de gestionar la infraestructura física), y los proveedores de servicios (responsables de los protocolos de red y de ofrecer los servicios finales). De esta forma los proveedores de servicios podrían solicitar y gestionar en función de las necesitados redes ópticas virtuales dedicadas y específicas por las aplicaciones.

Las tecnologías de virtualización de redes ópticas virtuales permiten la partición y composición de infraestructura de red (nodos y enlaces ópticos) en recursos virtuales independientes que adoptan las mismas funcionalidades que los recursos físicos. La composición de estos recursos virtuales (nodos y enlaces ópticos virtuales) permite el despliegue de múltiplos VONs. Una VON no sólo está composada por un plan de transporte virtual, sino también por un plan de control virtual, con el objetivo de incorporar las funcionalidades necesarias a la VON(provisión de conexiones automáticas y recuperación (protección/restauración), ingeniería de tráfico, etc.).

Esta tesis se centra en la virtualización de redes ópticas con tres objetivos principales. El primer objetivo consiste en el diseño, implementación y evaluación de la arquitectura y los protocolos e interfaces necesarios por la virtualización de redes encaminadas a través de la longitud de ola y controladas por GMPLS. También incluye la introducción de un gestor de recursos para desplegar redes ópticas virtuales de forma dinámica. La introducción de este gestor de recursos implica la necesidad de una gestión de los recursos virtuales y de algortimes de asignación de recursos para la utilización óptima de los recursos físicos. Además el gestor de recursos tiene que ser capaz del despliegue de los recursos asignados, incluyendo un plan de control GMPLS virtual independiente para cada VON desplegada. Finalmente, este objetivo incluye la introducción de mecanismos de virtualización para redes elásticas ópticas (EON, en inglés).

El segundo objetivo es el diseño, la implementación y la evaluación experimental de una arquitectura de sistema para ofrecer redes MPLS-TP virtuales controladas por GMPLS sobre una infraestructura WSON compartida. Por eso, esta tesis también se centra en el diseño y desarrollo de un nodo MPLS-TP que ha sido desplegado al demostrador ADRENALINE, en el CTTC.

Finalmente, el tercer objetivo es la composición de múltiples redes ópticas virtuales en dominios de control heterogéneos (GMPLS y OpenFlow). Un gestor de recursos multi-dominio ha sido diseñado, implementado y evaluado.

Summary

Virtual optical networking supports the dynamic provisioning of dedicated networks over the same network infrastructure, which has received a lot of attention by network providers. The stringent network requirements (e.g., Quality of Service -QoS-, Service Level Agreement -SLA-, dynamicity) of the emerging high bandwidth and dynamic applications such as high-definition video streaming (e.g., telepresence, television, remote surgery, etc.), and cloud computing (e.g., real-time data backup, remote desktop, etc.) can be supported by the deployment of dynamic infrastructure services to build ad-hoc Virtual Optical Networks (VON), which is known as Infrastructure as a Service (IaaS). Future Internet should support two separate entities: infrastructure providers (who manage the physical infrastructure) and service providers (who deploy network protocols and offer end-to-end services). Thus, network service providers shall request, on a per-need basis, a dedicated and application-specific VON and have full control over it.

Optical network virtualization technologies allow the partitioning/composition of the network infrastructure (i.e., physical optical nodes and links) into independent virtual resources, adopting the same functionality as the physical resource. The composition of these virtual resources (i.e., virtual optical nodes and links) allows the deployment of multiple VONs. A VON must be composed of not only a virtual transport plane but also of a virtual control plane, with the purpose of providing the required independent and full control functionalities (i.e., automated connection provisioning and recovery (protection/restauration), traffic engineering (e.g., QoS, SLA), etc.).

This PhD Thesis focuses on optical network virtualization, with three main objectives. The first objective consists on the design, implementation and evaluation of an architecture and the necessary protocols and interfaces for the virtualization of a Generalized Multi-Protocol Label Switching (GMPLS) controlled Wavelength Switched Optical Network (WSON) and the introduction of a resource broker for dynamic virtual GMPLS-controlled WSON infrastructure services, whose task is to dynamically deploy VONs from service provider requests. The introduction of a resource broker implies the need for virtual resource management and allocation algorithms for optimal usage of the shared physical infrastructure. Also, the deployment of independent virtual GMPLS control plane on top of each VON shall be performed by the resource broker. This objective also includes the introduction of optical network virtualization for Elastic Optical Networks (EON).

The second objective is to design, implement and experimentally evaluate a system architecture for deploying virtual GMPLS-controlled Multi-Protocol Label Switching Transport Profile (MPLS-TP) networks over a shared WSON. With this purpose, this PhD Thesis also focuses on the design and

development of MPLS-TP nodes which are deployed on the WSON of the ADRENALINE Testbed at CTTC premises.

Finally, the third objective is the composition of multiple virtual optical networks with heterogeneous control domains (e.g., GMPLS, OpenFlow). A multi-domain resource broker has been designed, implemented and evaluated.

This PhD Thesis is structured in six parts. Part I a background, motivation and objectives for this PhD Thesis. Parts II, III and IV focus on solving the three proposed objectives. Part V is devoted to dissemination and exploitation results. Finally, Part VI concludes and describes possible future work.

As explained, Part I provides a background, motivation and objectives for this PhD Thesis. Chapter 1 provides an overview on optical transport networks, providing an evolution from current transport infrastructure of IP over Point-to-Point Dense Wavelength Division Multiplexing (DWDM), towards IP offloading to WSON. Later, IP traffic offloading to multi-layer MPLS-TP over WSON is presented and finally, a rationale for EONs is provided.

Chapter 2 introduces different network control technologies, such as GMPLS and OpenFlow, and describes how they have been used as network control technologies in WSONs.

Chapter 3 provides a state of the art on optical network virtualization. A motivation for VONs is provided, providing clear distinction between VONs and Layer 1 Virtual Private Networks (L1VPN). The key concepts of abstraction, partitioning and aggregation are presented. Then, one of the key research areas related with VONs is introduced as the VON resource allocation problem. Later, a resource broker taxonomy is discussed and several examples of current resource brokers are provided. Current limitations are highlighted, in order to provide a strong motivation for this PhD Thesis objectives, which are described in Chapter 5.

Part II focuses on GMPLS-controlled VONs. Chapter 6 presents the deployment of GMPLS-controlled WSON as a Service. To this end, an architecture for virtualizing both the data and control planes is presented. Finally, the development of a virtualizable GMPLS-controlled WSON platform in the ADRENALINE Testbed is described and performance evaluation results are provided in terms of provisioning and recovery delay.

Chapter 7 introduces two architectures for a VON resource broker with the purpose of deploying dynamic WSON as a service. The first architecture deploys VON connected subgraphs of the underlying physical infrastructure, meaning that virtual links belong to the set of physical links. The second presented architecture deploys VON virtual network topologies, whose virtual links do not need to be directly mapped into the underlying physical links. Both architectures have been designed and developed in the framework of the virtualizable GMPLS-controlled WSON platform in the ADRENALINE Testbed. Performance evaluation results are provided in terms of virtual GMPLS-controlled WSON setup and tear down delay.

Chapter 8 is the first of four chapters which deals with the VON resource allocation problem. In this chapter, VON resource allocation algorithms for dynamic VON transport plane allocation of VON connected subgraphs. The proposed algorithms are Transparent Wavelength Allocation First-Fit algorithm (TWA-FF) and Transparent Wavelength Allocation Random algorithm (TWA-RND). Later, Wavelength Converters (WC) are introduced with the purpose of bypassing the Wavelength Continuity Constraint (WCC). Two algorithms which take into account the use of WCs are presented: Wavelength Allocation with Wavelength Converters First-Fit and Random algorithms (WAWC-FF/WAWC-RND). Performance results are provided in terms of VON request blocking rate in three different scenarios (14-node European, NSFNet and full-meshed topologies).

Chapter 9 studies four VON resource allocation algorithms for VON virtual network topologies: Shortest Path Unreserved Bandwidth First-Fit (SPUB-FF), Shortest Path Common Wavelength Allocation First-Fit (SPCWA-FF), Shortest Path Common Wavelength Allocation Random (SPCWA-RND) and Wavelength Tree Minimum Hop Shortest Path (WTMHSP). These algorithms take advantage of Global Concurrent Optimization (GCO), which is performed at the Path Computation Element (PCE). The performance of the presented algorithms is evaluated in terms of VON request blocking rate and VON setup delay. WTMHSP algorithm is the less blocking algorithm, with a small penalty of higher setup delays.

Chapter 10 focuses in Virtual Elastic Optical Networks (VEON) and details the required modifications for a resource broker in order to provide EONs as a Service. This chapter also describes a VON resource allocation algorithm which allocates the requested transport resources in a VEON. An experimental performance evaluation is performed in terms of VEON request blocking rate in a 14-node spanish topology scenario.

Finally, concluding Part II, Chapter 11 focuses on VON control resource allocation. This Chapter analyzes the virtualization of the GMPLS control plane, and provides a resource allocation algorithm, which assigns the required virtual control resources to the requested VON. Later, the impact of the virtualization of the GMPLS control plane is analyzed, by providing performance results on the blocking probability and setup delays of the deployed VONs.

Part III focuses on the design and evaluation of a resource broker for deploying virtual GMPLS-controlled MPLS-TP networks over a shared WSON. Firstly, Chapter 12 presents the architectural design and performance evaluation of a MPLS-TP node with tunable 10Gbps transponders. The node has been implemented using commercial off-the-shelf hardware and the forwarding engine has been implemented using open source software. An evaluation of the node is provided by means of analyzing the obtained throughput and CPU usage in different evaluation scenarios with different traffic grooming and traffic aggregation strategies. Finally, a description on how to virtualize the presented MPLS-TP node is provided.

Chapter 13 proposed the deployment of virtual GMPLS-controlled MPLS-TP networks over a Shared WSON to overcome the limitations of the current Data Center interconnection infrastructure. The design, implementation and evaluation of a resource broker for deploying virtual GMPLS-controlled MPLS-TP networks are provided. Experimental assessment carried out in the ADRENALINE testbed has shown the feasibility of deploying independent instances of GMPLS control plane for each deployed virtual MPLS-TP network under dynamic virtual network requests, providing low setup delays.

Part IV, Chapter 14 introduces the concept of a multi-domain resource broker for the composition of multiple VONs with heterogeneous control domains (i.e., GMPLS and OpenFlow). The design, implementation and evaluation of the proposed multi-domain resource broker are detailed, and an experimental assessment is provided in terms of VON request setup delay.

Part V is devoted to dissemination and exploitation results. Chapter 15 presents the scientific publications published or accepted at the time or writing this PhD thesis. Publications are classified in Journals, International conferences, National conferences and Collaborations. 3 Journals have been published or accepted: one OSA Optics Express (2012), one Elsevier Journal of Computer Networks (2012) and one OSA/IEEE Journal of Optical Communication Networks (2013). 9 International conferences: European Conference and Exhibition on Optical Communication (ECOC 2013), International Conference on Transparent Optical Networks (ICTON 2013), European Conference and Exhibition on Optical Communication (ECOC 2012), European Conference on Network and Optical Communications (NOC 2012), International Conference on Optical Networking Design and Modeling (ONDM 2012), Optical Fiber Communication Conference and Exposition and The National Fiber

Optic Engineers Conference (OFC/NFOEC 2012), Photonics in Switching (PS 2012), IEEE Conference on High Performance Switching and Routing (HPSR 2011) and Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC 2011). 1 National Conference and 14 Collaborations.

Chapter 16 presents the International, European and National R&D projects in optical networking in which some experimental results obtained in this PhD thesis are used. A summary of the objectives is given for FP7 EU-Japan ICT STRAUSS, FP7 ICT IDEALIST, FP7 ICT STRONGEST, FP7 ICT OFELIA and National MINECO DORADO and FARO projects.

Finally, Chapter 17 in Part VI concludes and describes possible future work on optical network virtualization.

Keywords: Optical Network Virtualization, GMPLS/PCE, Testbed, resource allocation algorithms, MPLS-TP/WSON networks, OpenFlow, virtualization orchestration.

Contents

List of Figures

List of Tables

Abbreviations

API	Application Programming Interface
ADC	Analog to Digital Converter
AoD	Architecture on Demand
AWG	Arbitrary Waveform Generator
BER	Bit Error Ratio
BVT	Bandwidth Variable Transponder / Transceiver
CAPEX	Capital Expenditures
CFO	Carrier Frequency Offset
CW	Control Word
DAC	Digital to Analog Converter
DC	Data Center
DEMUX	Demultiplexer
DD	Direct Detection
DMT	Discrete MultiTone
DP	Dual Polarization
DSP	Digital Signal Processing
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbium Doped Fibre Amplifier
EN	Edge Node
ERO	Explicit Route Object
FCS	Frame Check Sequence
FFT	Fast Fourier Transform
FPGA	Field Programmable Gate Array
FEC	Forward Error Correction
GMPLS	Generalized Multi Protocol Label Switching
HT	Holding Time
HW	HardWare
IaaS	Infrastructure as a Service
IAT	Inter-Arrival Time
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IM	Intensity Modulation
IPR	Intellectual Property Rights

IQ	In-phase Quadrature
ITU	International Telecommunications Union
JP	Japan
L2	Layer 2
LAN	Local Area Network
LICL	Logical Infrastructure Composition Layer
MEF	Metro Ethernet Forum
MEMS	MicroElectroMechanical Systems
MSA	Multi-Source Agreement
MTU	Maximum Transmission Unit
MUX	Multiplexer
MZM	Mach-Zehnder Modulator
NHLF	Next Hop Label Forwarding
OCS	Optical Circuit Switching
ODU	OTN Data Unit
OEO	Optic-Electric-Optic
OF	OpenFlow
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OIF	Optical Internetworking Forum
ONF	Open Networking Foundation
OPAN	Optical Packet Access Network
OPEX	Operational Expenditures
OPM	Optical Performance Monitor
OPS	Optical Packet Switching
OSNR	Optical Signal to Noise Ratio
OTDM	Optical Time Domain Multiplexing
OTN	Optical Transport Network
OXC	Optical Cross Connect
PIN	Positive Intrinsic Negative
PM	Person Month
PSK	Phase Shift Keying
PWE	PseudoWire Emulation Edge to Edge
PXC	Photonics Cross-Connects
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
ROADM	Reconfigurable Optical Add/Drop Module
RRO	Record Route Object
RZ	Return to Zero
SDH	Synchronous Digital Hierarchy
SDN	Software Defined Networking
SDO	Software Defined Optics
SDO	Standards Defining Organization
SLA	Service Level Agreement
SMP	Symmetric MultiProcessing
TE	Traffic Engineering
VC	Virtualization Composer
VHDL	Very high speed integrated circuit Hardware Description Language
VON	Virtual Optical Network

VP	Virtualization Partitioner
VPN	Virtual Private Network
VV	Virtualization Visor
WSON	Wavelength Switched Optical Networks
WSS	Wavelengths Selective Switches

Acknowledgements

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

This thesis would not have been possible without the help, support and patience of my PhD Advisor, Dr. Raul Muñoz. It would also have not been possible with the help and mentorship of my Optical Networks and Systems Department (former Optical Networking Area) colleguees: Dr. Ramon Casellas, Dr. Ricardo Martínez, Dr. Michela Svaluto, Dr. Josep M. Fàbrega and Javier Vílchez.

I would like to acknowlege Dr. Gabriel Junyent for his help and guidance in Universitat Politècnica de Catalunya (UPC).

I would like to acknowledge the financial, academic and technical support of the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), and its Director, Professor Miguel Ángel Lagunas. I would also like to thank my colleagues and friends at the CTTC.

Above all, I would like to thank my parents and sister, and Montse for the personal support and great patience at all times.

Barcelona, Spain - August 30, 2013 Ricard Vilalta

Part I

Background and motivation

Chapter 1

Optical transport networks

This chapter describes the State of the Art (SoA), giving an overview on different optical transport technologies . Firstly, current transport architecture is described. The evolution towards Wavelength Switched Optical Networks (WSON) is introduced. Multi-layer networks are presented. Later, Elastic Optical Networks (EON) are discussed. Finally, Ethernet transport services are detailed.

1.1 Current transport architecture

Historically, transport network architectures have been built on Synchronous Optical NETwork / Synchronous Digital Hierarchy (SONET/SDH) technology. SONET/SDH is based on Time Division Multiplexing (TDM) technology, that is, different digital signals can be efficiently multiplexed (in time) within a single frame. It has been designed with specific characteristics such as strictly connection oriented (i.e., coarse-bandwidth, long-lived and manually-provisioned connections), high level of protection and availability, Quality of Service (QoS), and extended Operations, Administration and Manteinance (OAM) capabilities. This provides network operators (carriers) with a high benchmark for reliability and operational simplicity.

In the current IP over Dense Wavelength Division Multiplexing (DWDM) architecture, IP packets are encapsulated over MultiProtocol Label Switching (MPLS) packets for traffic engineering (TE) and QoS purposes, which are transported into SONET/SDH frames or Optical Transport Network (OTN)

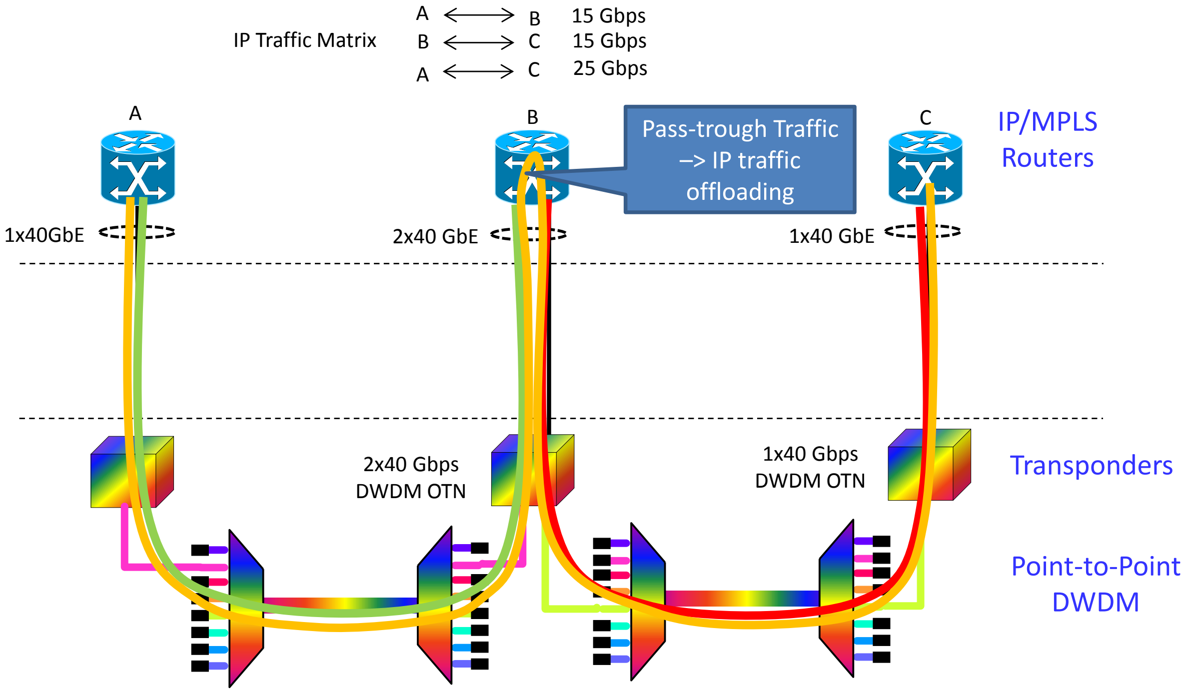


Figure 1.1: Current transport architecture: IP over Point-to-Point WDM

for extended OAM capabilities, and then carried on a dedicated wavelength through the point-to-point DWDM technology (see Fig. 1.1). IP packets are sent hop-by-hop from the ingress router to the egress router at the IP layer [1]. This solution allows the progressive introduction of point-to-point optical DWDM links, with IP routers connected to (or with integrated) DWDM-based transponders transmitting IP traffic over dedicated wavelengths. A main drawback of this solution is the costly electronic data switching capacity and the scalability of the IP routers.

In Figure 1.1 three IP traffic flows are defined of 15, 15 and 25 Gb/s. It can be observed how IP packets are sent hop-by-hop from the ingress router to the egress router at the IP layer. Point-to-point optical DWDM links are introduced, with IP routers connected to (or with integrated) DWDM-based transponders transmitting IP traffic over dedicated wavelengths.

1.2 Wavelength Switched Optical Networks

The accelerating growth and dynamicity of IP traffic (e.g., due to Internet video and cloud-based applications) has driven network operators to seek for more efficient transport alternatives and solutions for IP traffic offloading to the transport layer (i.e., IP packets are directly sent from the ingress router to the egress router using a transport connection at layer 1) in order to achieve cost reduction, simplified operations and increased scalability [2].

The last advances in wavelength switching technologies are driving the migration towards WSON [3]. This eliminates the costly Optical-Electronic-Optical (OEO) transponders and introduces Optical

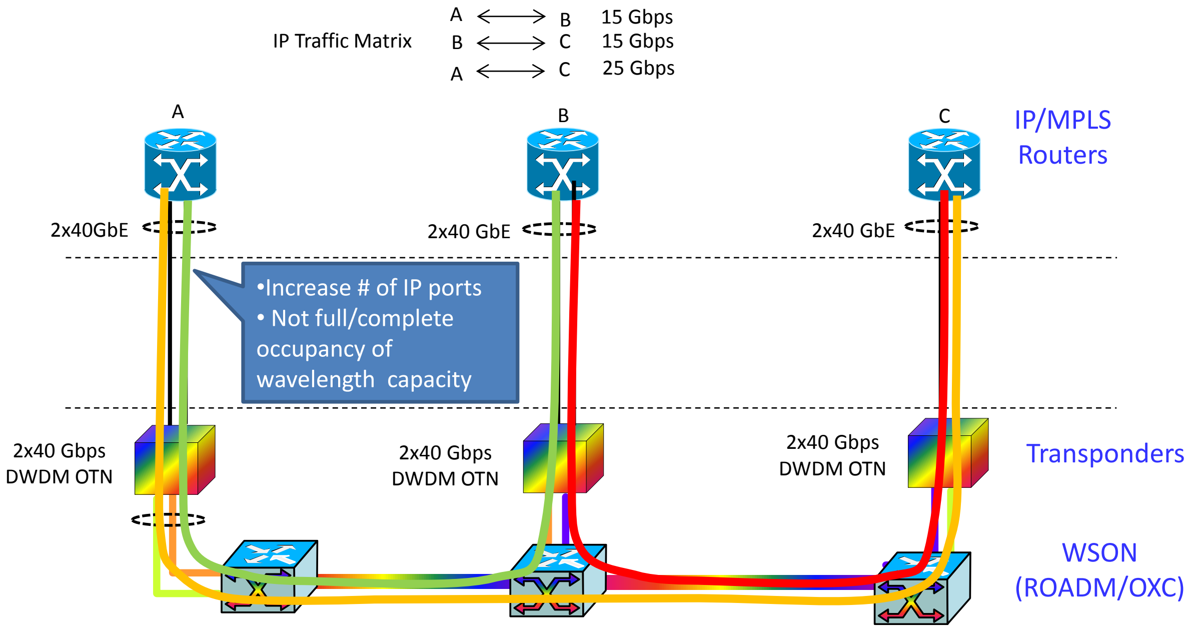


Figure 1.2: IP traffic offloading to WSON

Cross-Connects (OXC) and Reconfigurable Optical Add Drop Multiplexers (ROADM), that is, bypassing the IP/MPLS layer. The primary benefit of all-optical devices may be their greater scalability compared with OEO ones.

An OXC is able to switch DWDM channels among different fibers [4], demultiplexing the optical channels from any incoming fiber and routing them to the required output port [5]. Each output port is wavelength-multiplexed before the signal is sent to the outgoing fiber. A ROADM is able to drop DWDM channels from the incoming fiber and add DWDM channels to the outgoing fiber. The combination of both technologies allows the establishment of lightpaths between a source node and a destination node [6]. Current WSON are based on a fixed ITU-T DWDM wavelength grid, for a fixed set of channel spacing options (e.g., 50 GHz or 100 GHz).

Thanks to these emerging optical systems, in combination with the adoption of advanced modulation formats and tuneable transceivers, optical connections can be automatically switched entirely within the optical domain between source and destination nodes, optically bypassing the IP/MPLS routers.

Figure 1.2 shows IP traffic offloading to WSON, following the previously defined traffic matrix. Traffic from IP router A to IP router C is directly offloaded to WSON, without the need of electronic switching at IP router B. This architecture results in an increase of the number of IP ports, and has the drawback of not full occupancy of the wavelength capacity.

1.3 Multi-layer MPLS-TP/WSON

A challenging approach to fill wavelength capacity is a dynamic multi-layer transport network architecture integrating both a connection-oriented packet transport network (e.g., MultiProtocol Label

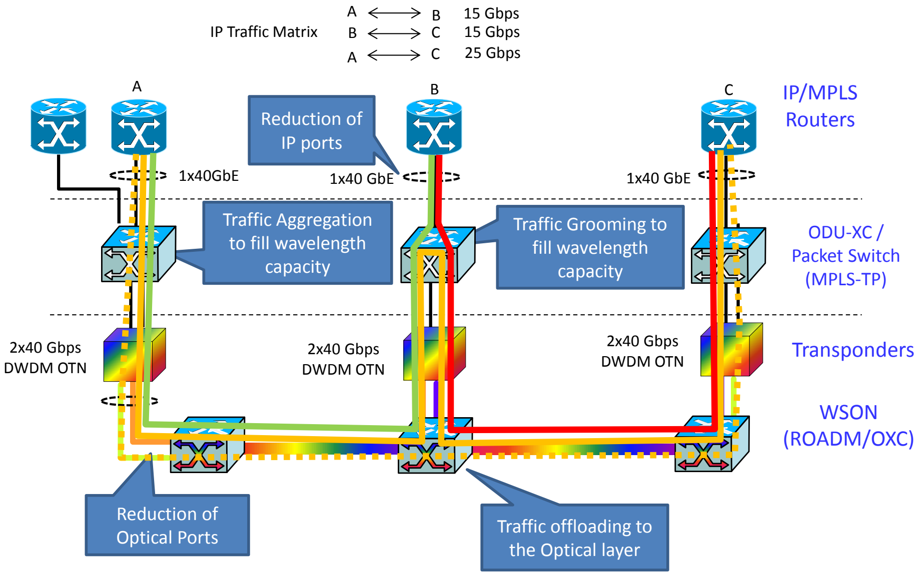


Figure 1.3: IP traffic offloading to multi-layer MPLS-TP over WSON

Switching Transport Profile -MPLS-TP-) and a WSON. The combination of MPLS-TP and WSON technologies leverages the high-bandwidth transport capacity and deterministic performance provided by the optical circuit switched technology and the efficient traffic aggregation/grooming and the statistical multiplexing provided by packet switched networks.

Reconfigurable and high-bandwidth optical connections can be set up entirely within the optical domain to bypass the transit traffic in the core IP/MPLS routers. Consequently, the processing at higher layers is avoided at the intermediate routers, left only at the points where header processing is needed (e.g., the network access edge). Thus, allowing IP traffic offloading to multi-layer MPLS-TP switches over WSON (Fig. 1.3).

MPLS-TP is defined as a profile of MPLS with the fundamental idea of expanding MPLS where necessary with OAM tools that are widely applied in existing transport network technologies [7]. A Network Management System (NMS) can be used to configure connections in a MPLS-TP network. Connection provisioning management and restoration functions, however, can alternatively be provided utilizing the Generalized MPLS (GMPLS) control plane protocols which are also applicable to the MPLS-TP data plane [8]. MPLS-TP is strictly connection oriented and client-agnostic (can carry L3, L2, L1 services). It is also, physical layer agnostic (i.e., can run over IEEE Ethernet PHYs, SONET/SDH, OTN, WDM, etc.).

1.4 Elastic Optical Networks

As described in section 1.2, current WSON are based on a fixed ITU-T DWDM wavelength grid, for a fixed set of channel spacing options (e.g., 50 GHz or 100 GHz). This rigid grid-based approach does not seem adapted to new data rates beyond 100 Gbps, and it is particularly inefficient when a whole wavelength is assigned to a lower speed optical path (e.g., 10 Gbps) that does not fill the entire wavelength capacity.

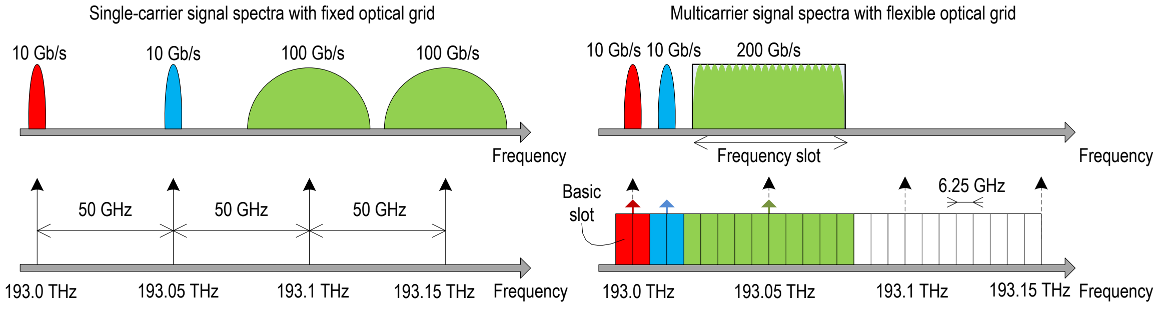


Figure 1.4: Spectrum-sliced elastic optical path network

WSON is designed with the premise that all channels in a network have the same spectrum needs (e.g. a 50 GHz channel spacing is the most common approach for 100 Gbps transmission, requiring 37.5 GHz with DP-QPSK). In fact, the ITU-T has normalized the allowed channels by specifying the DWDM grid. However, this rigid grid-based approach is not adapted to the spectrum requirements of the signals that are best candidates for long-reach transmission and high-speed data rates of 400Gbps and beyond (e.g. 400 Gbps using DP-QPSK does not fit in the 50 GHz DWDM grid). One conservative approach would be to increase the channel spacing (e.g. 150 GHz for 400 Gbps using DP-QPSK), or to demultiplex 400 Gbps into four 100 Gbps in order to fit in the 50 GHz DWDM grid. However, it is particularly inefficient in terms of spectrum efficiency in both cases [1].

A more disruptive approach is to replace the fixed DWDM grid by a flexible grid, in which the optical spectrum is characterized by a frequency grid having nominal central frequencies and the granularity (e.g. 6.25 GHz) [9]. The required amount of optical bandwidth/spectrum (the so called ”frequency slot”) for an optical channel is determined by the signal modulation format and its data rate, and it can be dynamically and adaptively allocated to an (elastic) optical connection by selecting a nominal central frequency and a slot width (see Fig. 1.4).

The key elements of the Elastic Optical Networks are the programmable (Bandwidth/spectrum)-Variable Transponders (BVT), which generate optical signals while supporting multiple modulation formats and bit-rates, and the Bandwidth-Variable OXC (BV-OXC), which are able to switch an optical signal based on a frequency slot rather than on a fixed wavelength. Figure 1.5 details how IP traffic is offloaded to Elastic Optical Networks, with the help of BVT and BV-OXCs.

The problem of computing a path/route and allocate spectrum resources (i.e., contiguous frequency slots) is known as Routing and Spectrum Allocation (RSA). Recent studies have considered both the offline RSA planning and the online dynamic RSA [10]. As for dynamic RSA algorithms, they can be further sorted as centralized or distributed routing control.

In centralized routing control, a single entity (e.g., a network management system with a stateful Path Computation Element (PCE) [11]) maintains a complete, global and unique view of the network topology, spectrum resource availability (i.e., frequency slots), and physical impairment parameters (e.g., OSNR, CD, etc).

1.5 Ethernet transport

Ethernet protocol has been massively deployed, due to its simplicity and ease of maintenance, its reliability and the low cost of installation and upgrade. IETF has introduced the Pseudowire Edge-

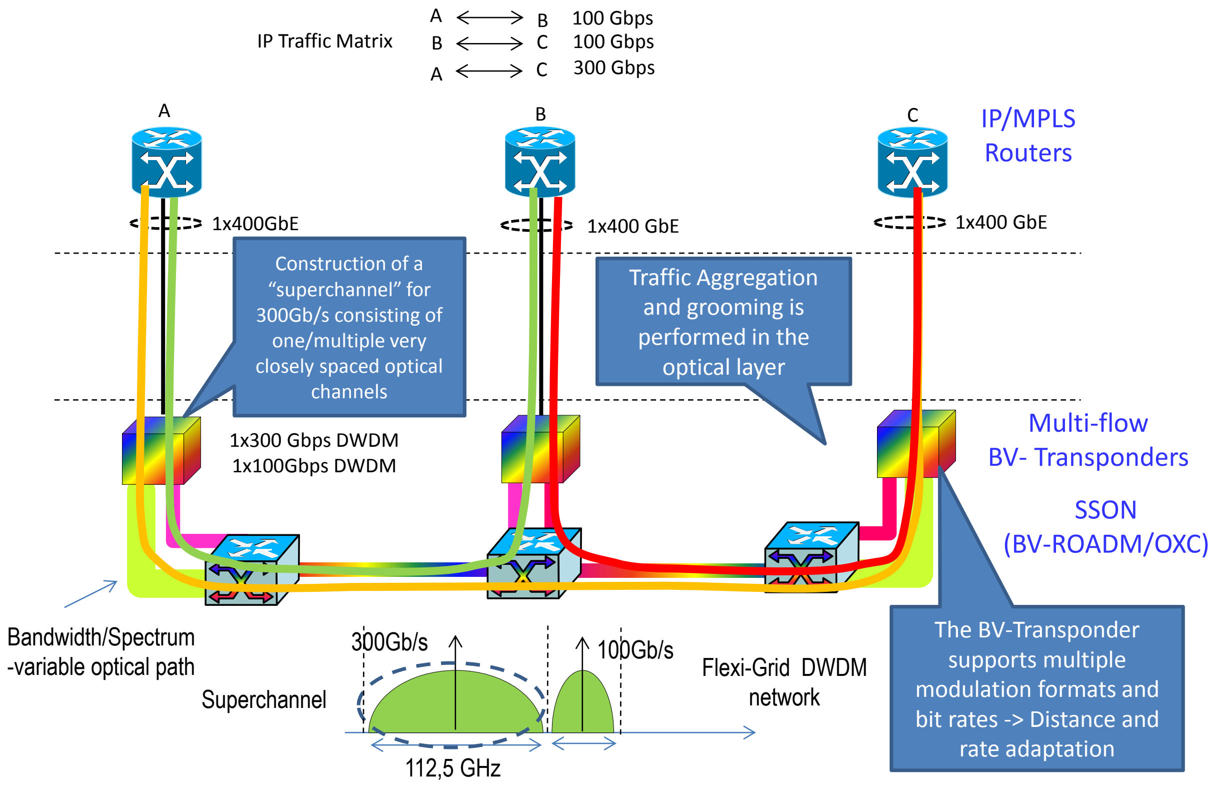
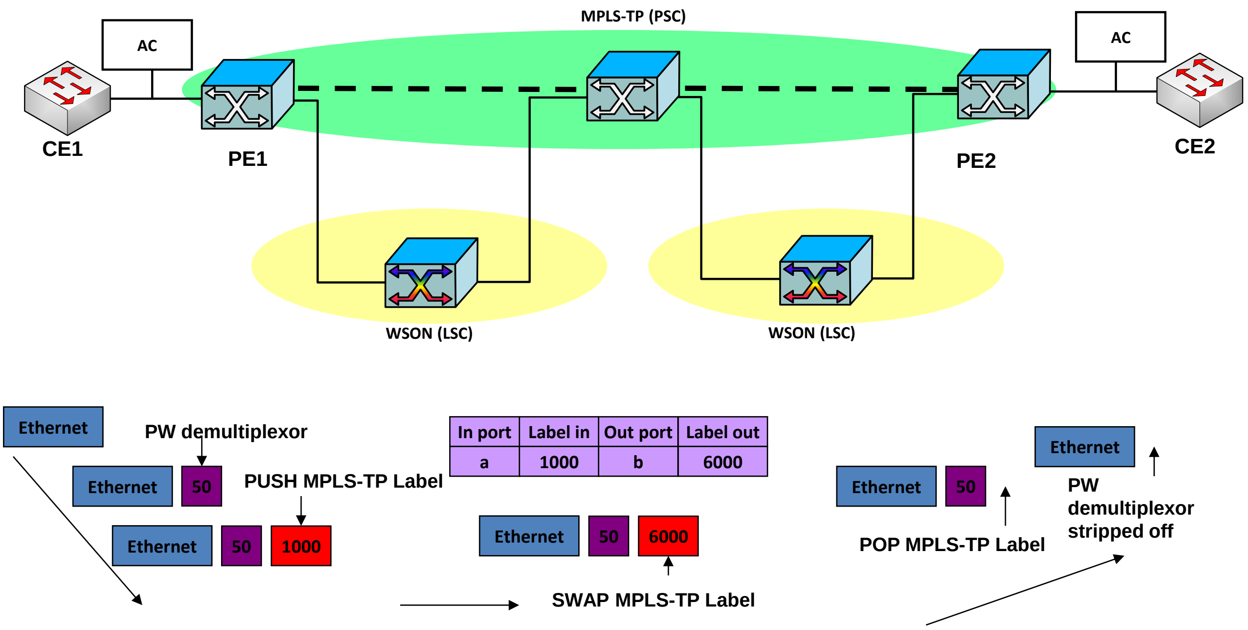


Figure 1.5: IP traffic offloading to Elastic Optical Networks



In port	Label in	Out port	Label out
a	1000	b	6000

Figure 1.6: Pseudowire Edge-to-Edge (PWE3) Ethernet transport services

to-Edge (PWE3) [12], which is an emulation of a native service (e.g., Ethernet transport service) over a Packet Switched Network (PSN), which is the presented MPLS-TP Network.

In Figure 1.6, it is shown an emulated Ethernet service for a point-to-point connection between two customer edges (i.e., CE1 and CE2). Provider Edges 1 and 2 (i.e., PE1 and PE2) are the ingress and egress nodes, respectively, and they are responsible to create the PWE3 service. An Attachment Circuit (AC) is the physical or virtual circuit attaching the CE to a PE. For the Ethernet service, the AC is an Ethernet port. Indeed, the function of the PE1 is to receive the (Ethernet) PDU coming from CE1 and to encapsulate and transport it to PE2, where it is decapsulated and transmitted out on the attachment circuit (AC) of the PE2. The pseudowire starts at a logical port within the PE (PE1). This port provides an Ethernet MAC service that will deliver each Ethernet frame to the other end (PE2) of the PW.

An important function of the Provider Edge is the Native Service Processing (NSP) that includes the frame processing that is required for the Ethernet frames that are forwarded to the pseudowire termination point. The ingress NSP strips the preamble and the Frame Check Sequence (FCS) from the Ethernet frame and transports the frame in its entirety across the PW. Then the control word, if necessary, is prepended to the resulting frame. The function of the forwarder is to select the PW based on the incoming AC, the contents of the payload or some statically and/or dynamically configured forwarding information. Next, the proper PW demultiplexer (PW label) is prepended to the resulting packet. Since the PW is transported over an MPLS-TP PSN, an inner MPLS-TP label is used to provide the PW demultiplexing function. The proper tunnel encapsulation (MPLS-TP label) is prepended to the packet, and the packet is transmitted.

When the packet arrives over a PW, the tunnel encapsulation and PW demultiplexer are stripped off. If the control word is present, it is processed and stripped off. The egress NSP function receives the Ethernet frame from the PW and regenerates the preamble or FCS before forwarding the frame to the attachment circuit.

Chapter 2

Network control technologies

This chapter introduces two network control technologies: GMPLS [13] and OpenFlow [14]. Although GMPLS network control will be the main network control technology for the development of the proposed PhD Thesis, optical network virtualization in heterogeneous control domains is also discussed, so OpenFlow is discussed as a network control technology.

2.1 Generalized Multiprotocol Label Switching (GMPLS)

In the previous section, a set of different transport network topologies has been introduced (i.e., WSON, MPLS-TP, etc.). [15] defines a transport network as "the functional resources of the network which conveys user information between locations". The term user information provides a clear layering structure on the network. GMPLS provides a common control plane for managing different network technologies and enables high-function services across the network. The GMPLS control plane consists on a set of routing and signaling protocols. The signaling protocols are responsible for the establishment of transport plane paths, while the routing protocols are responsible for dynamically distributing connectivity and reachability information, and the Traffic Engineering (TE) resource attributes. The Link Management Protocol (LMP) is added to the protocol family in order to support link discovery and verification in GMPLS networks [16].

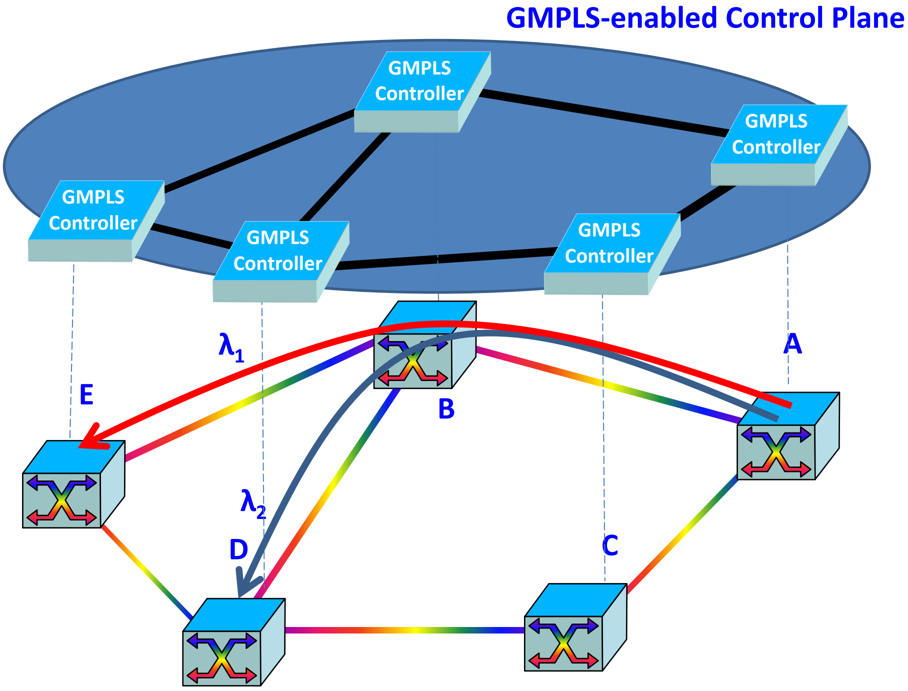


Figure 2.1: GMPLS-controlled WSON

The GMPLS protocol suite extends MPLS to manage further classes of interfaces and switching technologies, such as Time Division Multiplex (TDM) or wavelength switching, among others. GMPLS is based on the concept of Generalized Labels, which are abstract labels which can represent either a single fiber in a bundle, a single waveband within fiber, a single wavelength within a waveband (or fiber), or a set of time-slots within a wavelength (or fiber). GMPLS is composed of three main protocols:

- Resource Reservation Protocol with Traffic Engineering extensions (RSVP-TE) signaling protocol [17].
- Open Shortest Path First with Traffic Engineering extensions (OSPF-TE) routing protocol [18].
- Link Management Protocol (LMP) [16].

The provisioning of end-to-end connections requires horizontal coordination among the nodes, performed by the RSVP-TE signaling protocol and employing a hop-by-hop mechanism from the source to the destination node. A Label Switched Path (LSP) is the path set up by the signalling protocol. The OSPF-TE routing protocol is responsible for disseminating any change occurring in the network state, allowing the nodes to update their local Traffic Engineering Databases (TED).

The GMPLS Control Plane is responsible for the actual resource and connection management within a WSON. It consists of a series of optical connection controllers (i.e., GMPLS Controllers), interconnected via Network to Network Interfaces (NNI). The described GMPLS Controllers include the following functions:

- Network topology discovery (resource discovery)

- Signaling, routing, address assignment
- Connection set-up/tear-down
- Connection protection and restoration
- Traffic Engineering
- Path Computation

Figure 2.1 shows a GMPLS-controlled WSON, consisting of OXCs and ROADMs. Each OXC or ROADM is controlled by its own GMPLS controller. It can be observed that GMPLS Control Plane GMPLS controllers interconnection can be established via a Dedicated Control Network (DCN), based on IP Control Channels (IPCC).

2.1.1 Path Computation Element

A PCE is an entity (component, application or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints [19]. It involves a (standard) functional formalization of a network architecture and a communications protocol interface (PCE Protocol -PCEP-). Additionally, new collaborative approaches and hierarchical models are being proposed for advanced (multi-domain and multi-layer) path computation solutions [20].

For large optical networks, the path computation process may become computationally complex and will require dedicated resources. To this end, the PCE provides advanced path computations serving requests from network nodes (i.e., GMPLS controllers) through a standard interface, considering also the TED as input.

In the context of WSON, providing reconfigurable lightpaths, the PCE is responsible for computing paths (i.e., nodes, links, allocated 3R regenerators, and used wavelength channels) for each connection request [21], fulfilling the requirement of adequate quality of transmission and minimizing the number of the used network resources. This is known as the Routing and Wavelength Assignment (RWA) problem which requires extending the current routing protocols to flood information concerning physical impairments, 3R regenerators, node wavelength converters and wavelength channel availability, besides the standard TE attributes (Fig. 2.2).

It is worth mentioning that PCE will be a key element of the proposed GMPLS-controlled WSON virtualization architecture.

2.1.2 Unified Control Plane

The automation of multi-layer MPLS-TP/WSON transport networks can be efficiently encompassed through the deployment of a distributed Unified Control Plane (UCP) based on the GMPLS architecture [22]. A distributed control plane provides a common set of interconnection mechanisms (e.g., signaling protocol for resource reservation and routing protocol for topology and resource dissemination) and functions such as automated connection provisioning and recovery, traffic engineering or QoS [23].

Focusing on the UCP based on GMPLS, a controller's TED contains the information relative to both the optical circuit switched and packet switched layers (i.e., Lambda Switched Capable - LSC TE links and Packet Switched Capable - PSC TE links, respectively). Therefore, a path can be computed

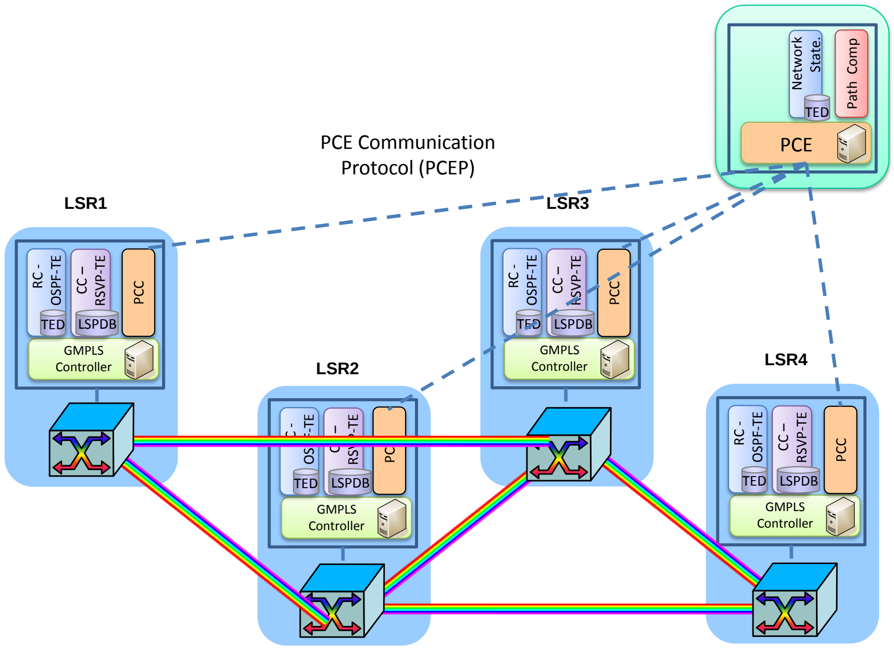


Figure 2.2: Functional architecture of the Path Computation Element

traversing multiple layers. To this end, GMPLS defines the concept of Forwarding Adjacency (FA). That is, assuming a local control policy, a GMPLS node may advertise an LSC connection as a PSC TE link. Such a link is referred to as FA. The corresponding LSP is thus referred to as a FA LSP (see Fig. 2.3).

The goal of using FAs is that, after the routing protocol floods the TE attributes (i.e., end-point LSRs, bandwidth, etc.) of a particular FA, the packet nodes (e.g., MPLS-TP) may use the FAs as a regular PSC TE link for path computation purposes. Additionally, if the user service is Ethernet, according to the PWE3 architecture, a Pseudo-Wire tunnel is set up through the just created PSC LSP, by means of Target Label Discovery Protocol (T-LDP) or an external Network Management System (NMS).

2.2 OpenFlow

OpenFlow is an open standard that enables researchers to run experimental protocols in ethernet networks. OpenFlow is added as a feature to commercial Ethernet switches, routers and wireless access points, and provides a standardized hook to allow researchers to run experiments, without requiring vendors to expose the internal workings of their network devices [14]. OpenFlow enables the determination of the path of network L2 frames by software running on a separate server. OpenFlow is a key enabler for Software Defined Networking (SDN), which is defined as the physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices [24]. In SDN, network intelligence and state are logically centralized, and the underlying network infrastructure is abstracted from the applications.

The OpenFlow protocol abstracts data plane switches as a flow table, in which a flow is defined as any combination of L2, L3 and L4 packet headers, providing the control plane with a common

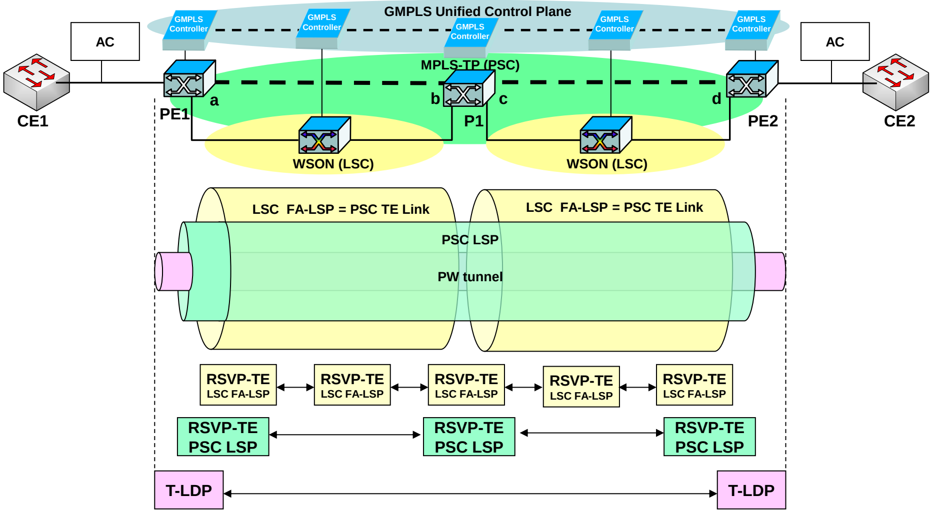


Figure 2.3: Setting up of a PSC LSP to transport Ethernet services over a WSON using a GMPLS unified control plane

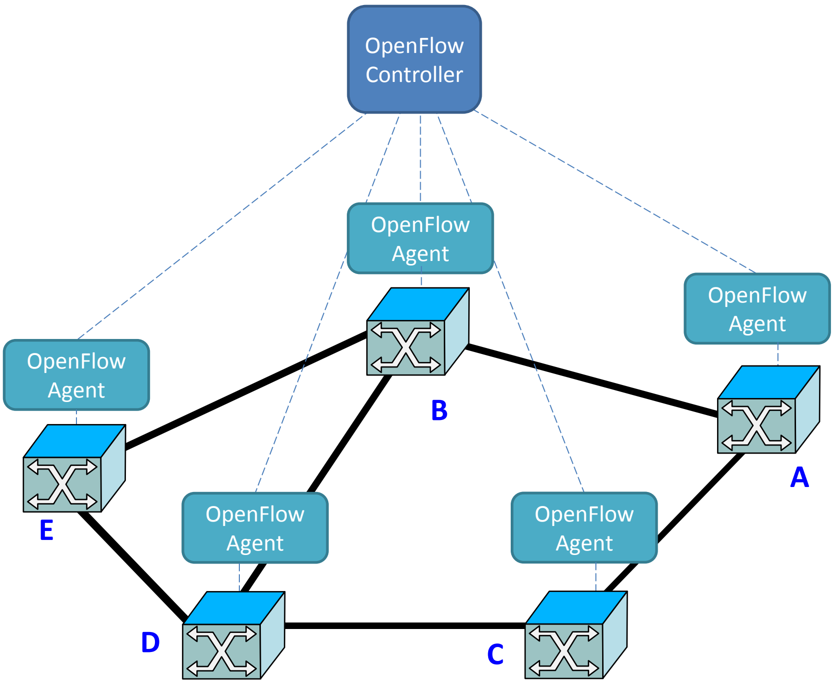


Figure 2.4: OpenFlow Network

hardware abstraction through a standardized open interface (i.e., the OpenFlow protocol) to program the flow table (see Fig. 2.4).

In general, OpenFlow assumes a logically centralized control plane, with a OpenFlow Controller running a network-wide operating system (e.g., NOX [25], FloodLight [26] or OpenDayLight [27] are well-known OF Controllers) that controls the data plane using the aforementioned OpenFlow protocol (see Fig. 2.4). The network-wide operating system is usually composed of a single software controller, and a single network view database, that is kept in a database running on one of the servers.

For example, the NOX's network view can store the switch-level topology, the locations of users, hosts, middleboxes and other network elements, and the services (e.g., HTTP or NFS being offered). The network view may also include all bindings between names and addresses, but it does not include the current state of the network traffic to minimize scalability issues.

Different applications can run on the NOX which using this centralized network view make management decisions such as deciding how each flow is routed (i.e., routing), which flows are admitted (i.e., control access), where they are replicated (i.e., multicast), configuring the switches in a vertical manner, etc. Such a network-wide operating system allows these management applications to be written in as centralized programs. That is, they are written as if the entire network were present in a single machine. To this end, NOX offers a standardized API used by the management applications to create new functionalities and services to the network.

Since many vendors are yet to embrace OF, an in-house built OF agent placed on the node is used to provide OF abstractions. A modular openflow agent might provide hardware abstractions to the controller via an OpenFlow interface. This agent utilizes the network equipment management interface (e.g., SNMP, TL-1) to communicate with the data plane.

2.2.1 Optical OpenFlow

More recently, the OpenFlow architecture has also been proposed to control multi-layer networks in a unified way [28]. An experimental OpenFlow UCP has been proposed [29] for packet and circuit switched networks. A simple proof-of-concept testbed has been demonstrated, where a bidirectional wavelength circuit is dynamically created to transport a TCP flow. In [30] a unified control plane for Packet, Fixed and Flexible DWDM Grid Technologies has proposed the following protocol extensions for different optical networks:

- Fixed DWDM grid.
- Flexible grid.
- Multi-domain.

Although the efforts in OpenFlow circuit switching, the proposed protocol extensions are not supported by latest openflow specifications, meaning that OpenFlow still remains packet oriented.

In Figure 2.5, it can be observed a single OpenFlow Controller responsible for a packet and circuit switched network. An OpenFlow agent is running on top of each packet switch and ROADM/OXC and is responsible for the translation of the OpenFlow protocol commands into specific vendor equipment management interface [11].

In the following subsections, the proposed protocol extensions shown in Figure 2.6 are described.

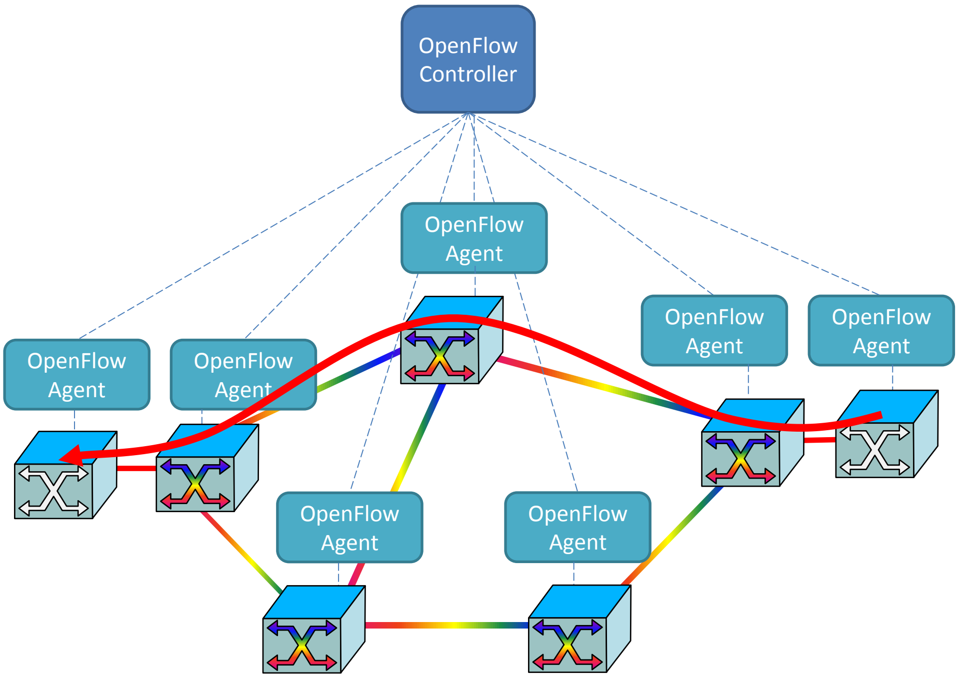


Figure 2.5: Unified control of Packet and Circuit Switched OpenFlow Network

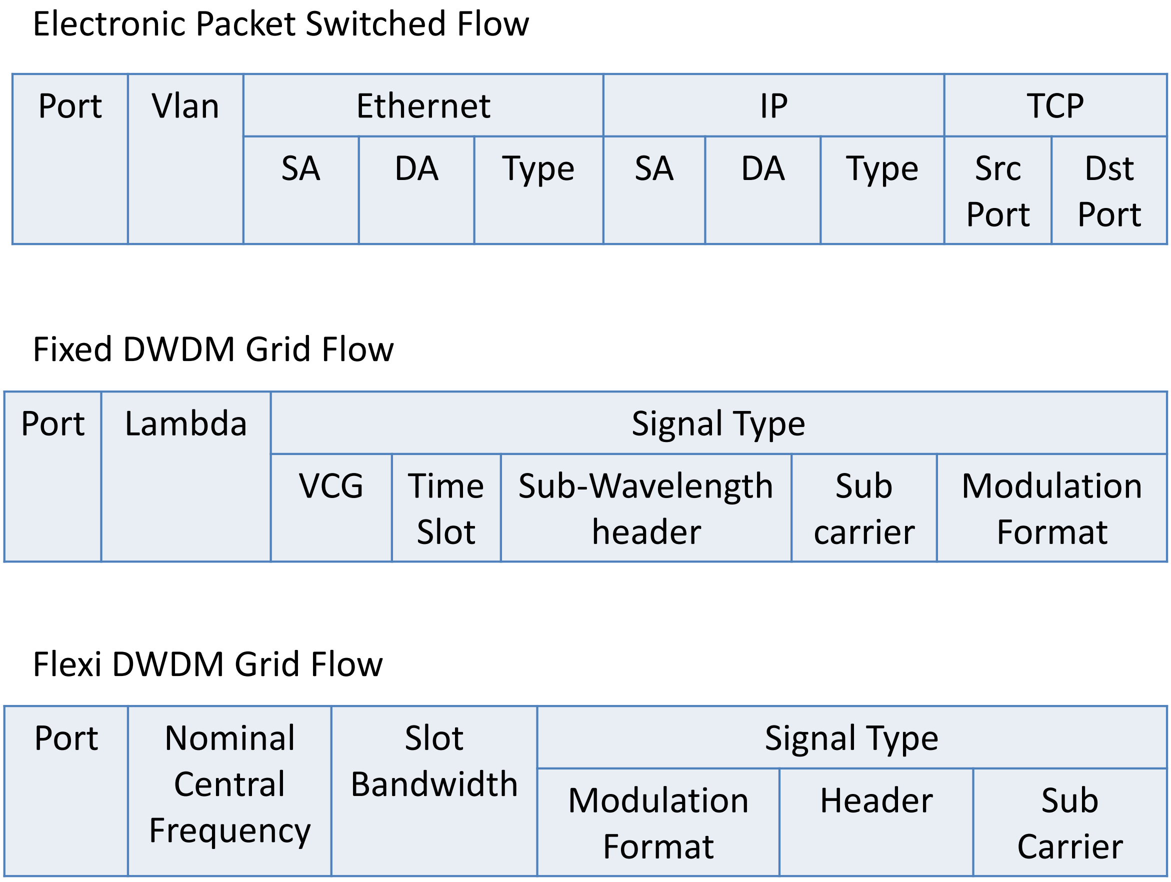
2.2.1.1 Fixed DWDM grid

Figure 2.6 shows the fixed DWDM grid flow definition, which includes port, wavelength, and signal type. The controller utilizes the switch feature requests/replies ($Switch\_Features$ messages) to construct the network topology and $CFlow\_Mod$ (flow specification) message to control the cross connections in the optical switches. Switching constraints are sent to the controller using a generic vendor extensions messages. The $CFlow\_Mod$ message is used to establish or teardown the cross-connections in each Network Equipment (NE). A specialized path computation module (OF_PCE) can be implemented and integrated in the extended controller. This module is responsible to compute a lightpath inside the optical domain with proper consideration for switching constraints.

These extensions rely upon OF agent software, which provides a novel hybrid switch abstraction to the OF controller beyond v0.3 circuit addendum [28]. The fixed grid optical agents extend the existing OF circuit extension by introducing switching constraint and power equalization messages. The agents resource model calculates the constraints from the capabilities information and exchanges it with OF message using the OF channel. Similarly the power equalization Of message is injected upon cross connection initiation via the OF controller.

2.2.1.2 Flexi-grid

Traditional fixed-grid network uses fixed size optical spectrum frequency ranges with channel spacing of 100 or 50 GHz but in flexi-grid optical networks, slot width of the frequency ranges assigned to different channels are flexible. This allows efficient and flexibly configurable, manageable optical spectral bandwidths for high bit-rate systems. The authors in [31] proposed extensions to the OF protocol and the design of a NOX based OF controller to extend the circuit specification to introduce flexi-grid functionalities for seamless path computation between different optical domains.



Electronic Packet Switched Flow

Port	Vlan	Ethernet			IP			TCP	
		SA	DA	Type	SA	DA	Type	Src Port	Dst Port

Fixed DWDM Grid Flow

Port	Lambda	Signal Type				
		VCG	Time Slot	Sub-Wavelength header	Sub carrier	Modulation Format

Flexi DWDM Grid Flow

Port	Nominal Central Frequency	Slot Bandwidth	Signal Type		
			Modulation Format	Header	Sub Carrier

Figure 2.6: OpenFlow protocol extensions

For the forwarding functions the agent processes flows based on OF specifications. Figure 2.6 shows the flow specification for flexi-grid optical networks. In a flexi-grid optical network a flow is identified by a flexible flow identifier comprising port, Nominal Central Frequency (NCF), slot bandwidth and type of signal fields associated with that switch.

For flexible DWDM grid, allowed frequency slots have a Nominal Central Frequency (NCF) defined by Equation 2.1 which is used to calculate the central frequency of a Frequency Slot (FS). n is an integer (either positive, negative or zero). Equation 2.2 describes the assigned slot bandwidth, where m is a positive integer. The allowable granularity of m and n for flexi-grid equipment can be determined using the $Switch\_Features$ messages.

$$NCF = 193.1 + n * 0.00625 (THz) \tag{2.1}$$

$$BW = 12.5 * m(GHz) \tag{2.2}$$

The $Switch\_Features$ message extension supports optical Network Equipment capabilities including: NCF, slot width and signal type (including modulation format). To control a BVT or BV OXC the flexi-grid DWDM flow definition values are exchanged between the controller and a Network Equipment (NE) via $CFlow\_Mod$ messages.

2.2.1.3 Multi-domain OpenFlow

For multi-domain aspects the controller is made aware of domain constraints by utilizing intra-domain and inter-domain flow tables. An intra-domain flow table holds flow identifiers and associated actions for each network element (NE) within a particular domain. For optical NEs, an action is defined as a cross connection associated with one or more flow identifiers. Inter-domain flow tables hold flow identifiers and associated actions for NEs that interconnect between neighboring domains. Actions in flow tables that are associated with two heterogeneous technology domains must comply with multi-domain mapping rules.

2.2.2 Optical FlowVisor

OpenFlow FlowVisor [32] is an OpenFlow switch virtualization approach (for packet switching domain) in which, the same hardware forwarding plane can be shared among multiple virtual networks. FlowVisor is implemented as an OpenFlow proxy that intercepts messages between OpenFlow-enabled switches and OpenFlow controllers.

The authors in [33] have proposed the concept of Optical FlowVisor, in order to provide Virtual Optical Networks (VON). Because optical FlowVisor defines an optical slice as a set of optical flows, and two slices can be isolated by making sure their flowspaces do not overlap anywhere in the topology.

Chapter 3

Optical network virtualization

This chapter provides a state of the art on optical network virtualization. First, a motivation and use cases for VONs is provided. The VON resource allocation problem is analyzed and several VON resource broker are detailed. This chapter provides the necessary concepts in order to understand how this PhD Thesis objectives fulfill the identified research challenges.

3.1 Introduction

Cloud and Grid Computing have introduced a new application paradigm where storage and server infrastructures are hosted and shared. These new paradigm has allowed cost reduction and innovation in services and applications. These new innovations are based on sliced servers, introducing a pay-per-use model. In this context, the concept of Platform as a Service (PaaS) is introduced. PaaS is a new service model for networking the future Internet. Today's Internet Service Providers (ISPs) serve two roles:

- Managing their network infrastructure.

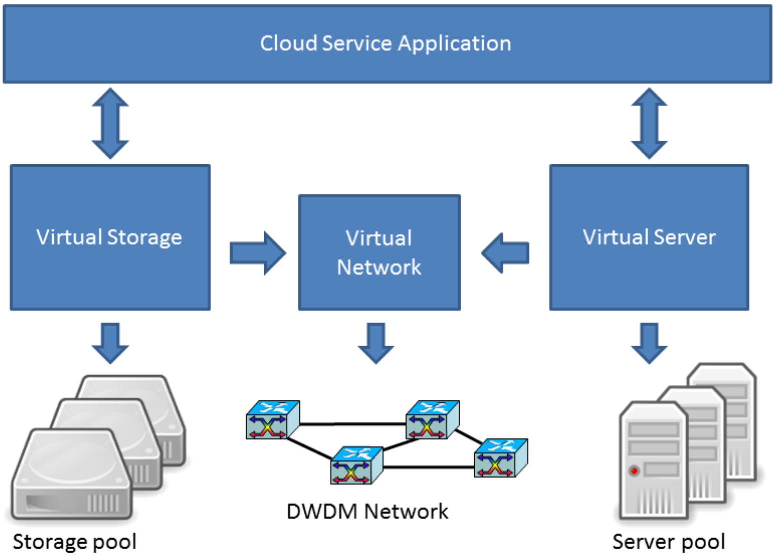


Figure 3.1: VONs for Grid/Cloud Computing

- Providing, arguably limited, services to end users.

The authors in [34] argue that coupling these roles impedes the deployment of new protocols and architectures. Instead, the future Internet should support two separate entities:

- Infrastructure providers, who manage the physical infrastructure.
- Service providers, who deploy network protocols and offer end-to-end services.

Optical network virtualization has been proposed for addressing this challenge, thus, providing better services to the end-users, which their requirements can be supported by the deployment of dynamic infrastructure services to build ad-hoc VON services, which can be offered on-demand. Thus, network service providers can request, on a per-need basis, a dedicated VON for each application and have full control over it. VON support dynamic provisioning of dedicated networks (a.k.a., slices) over the same network infrastructure. Thus, the stringent network requirements of the emerging high bandwidth and dynamic applications such as high-definition video streaming (e.g., telepresence, television, remote surgery, etc.), and cloud computing (e.g., real-time data backup, remote desktop, etc.) can be supported.

Optical network services have been proposed for Cloud and Grid Computing applications. This is the case of the Large Hadron Collider (LHC) and its data processing project MONARC [35]. Figure 3.1 shows a use case for a VON which can communicate different processing centers and storage centers over a shared DWDM optical physical infrastructure.

Network virtualization has been widely studied, an a comprehensive survey has been published [36]. Network virtualization takes into account node virtualization, link virtualization and generic data formatting (see Fig. 3.2) [37].

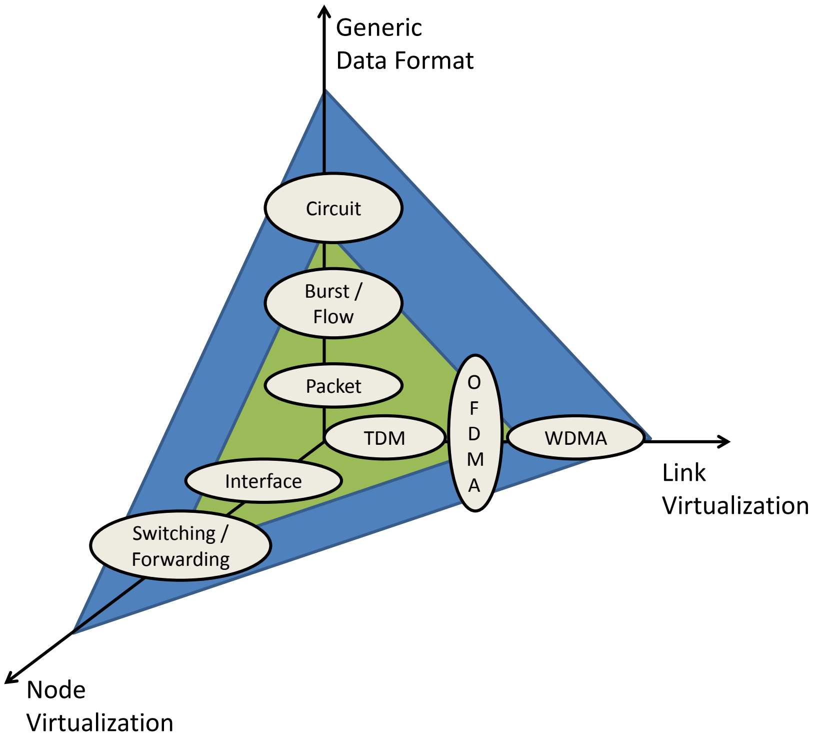


Figure 3.2: The 3 Edges of Virtualization

Node virtualization is considered the virtualization of switches, routers, and servers into logical or virtual routers, with different virtual interfaces (addresses) where data can be forwarded.

Link virtualization separates flows from different slices across their data path on both levels of bandwidth resource sharing schemes (e.g., wavelength, subwavelength, super-channel, waveband, multi-core or multi-fiber levels). As an example wavelength level is addressed through wavelength division multiple access (WDMA) via lightpath-based logical links that connect programmable routers.

Specifically, optical node virtualization is the creation of a virtual representation of an optical network node (i.e., ROADM or OXC), achieved by partitioning and/or aggregation [38].

Within a VON, a virtual optical link is defined as a connection between one port of a virtual network node to a port of another virtual network node. A physical optical link can be partitioned based on a TDMA, OFDMA or WDMA basis. Finally, a VON is the composition of isolated virtual optical resources, which are existing simultaneously over the same physical optical network infrastructure [38].

Optical network virtualization takes advantage of three key concepts:

- Abstraction: hides the technology specific details of the physical network resources and enables the physical infrastructure providers while representing them as a set of necessary uniform attributes, characteristics and functionalities.
- Partitioning: consists on the partition the physical resources into multiple independent slices (virtual resources) with each virtual resource exactly mimicking functionality and performance of the real physical resource slices.

- Aggregation: enables multiple inter-connected physical resources represented as a single virtual resource.

The authors in [39] define Layer 1 Virtual Private Networks (L1VPN). An L1VPN is defined as a VPN whose data plane operates at layer one. A connection between Customer Edges (CE) located in different sites of an L1VPN is called a L1VPN connection. L1VPN connections enable inter-site CE-CE links.

As an example, the authors in [40] introduced an optical network virtualization framework (UCLPv2) upon which communities of users can build their own services or applications without dealing with the complexities of the underlying network technologies and still can maintain the functionality that the network provides. The system has been designed as a service-oriented architecture where Web services and Web services workflows are the basic building blocks. Articulated L1VPNs are presented as the first services built upon the UCLPv2 network virtualization middleware.

Although they have proven useful, L1VPN have the following limitations [41]:

- L1VPN users do not have visibility into the carrier networks, and completely depend on the carriers to provision connections between carriers edge node ports
- An L1VPN can only interact with a user network based on a client-server relationship, where the L1VPN functions like a virtual node or link
- Contentions on time-sharing resources need to be solved by the carrier administrative policy

VONs overcome the presented limitations, by providing a more flexible and scalable solution. The Logical Infrastructure Composition Layer (LICL) has been introduced as a layer which decouples the physical infrastructure from its associated control plane by means of infrastructure virtualization [42]. The LICL enables the partitioning of physical hybrid infrastructures, that is, infrastructures composed either by IT resources, or by network resources, or by both at the same time, by abstracting the underlying resources and offering them as services. Moreover, the LICL facilitates the capability of composing resources in order to build and compose virtual infrastructures.

The LICL virtualization architecture is data plane agnostic, thus needing to deploy agents on top of each optical equipment, and requiring the development of several plugins for each particular node vendor. One of the objectives of this PhD Thesis is the design, implementation and performance evaluation of a multi-vendor GMPLS-based optical virtualization architecture.

3.2 Data plane virtualization

Optical node virtualization is the creation of a virtual representation of an optical network node (i.e., ROADM or OXC) with the same functionalities, based on an abstract model that is often achieved by partitioning and/or aggregation. A virtual optical link is defined as a connection between a port of a virtual optical network node to a port of another virtual optical network node (Fig. 3.3). A physical optical link can be partitioned based on a TDMA, OFDMA or WDMA basis. Finally, a VON is the composition of isolated virtual optical resources (i.e., virtual nodes and links), which exist simultaneously over the same physical optical network infrastructure.

In this section, an analysis of the several methods for partitioning the physical optical network and composition the isolated virtual optical networks is provided. These methods for partitioning and compositing the VONs are generally referred as resource brokers. Later, an introduction to VON resource broker taxonomy is discussed, and finally several literature resource brokers are analyzed.

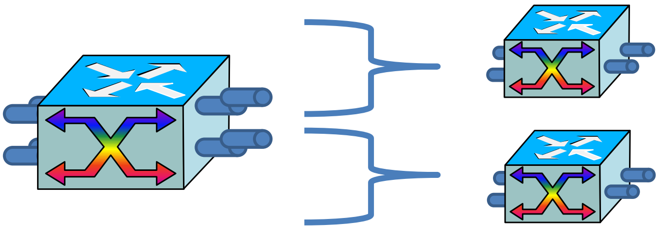


Figure 3.3: GMPLS-controlled optical switch partitioning

3.2.1 The VON Resource Allocation Problem

The Virtual Optical Network Resource Allocation (VON RA) problem differs from the classical RWA problem, because VONs are treated as single entities, instead as the grouping of lightpaths that can be independently requested in the RWA problem. A VON request is served only if all the considered virtual optical links can be allocated.

Resource allocation is described as a research challenge, where efficient allocation and scheduling of physical network resources among multiple virtual network requests is essential to maximize the number of coexisting independent virtual networks [36]. This challenge is divided into two different versions, namely offline and online. We refer to offline virtual network embedding when all virtual network requests are known in advance, while online virtual network embedding refers to the dynamic reception of incoming virtual network requests.

The introduction of optical network virtualization has introduced new challenges for optical networking researchers in order to apply VON RA strategies into the different system architectures that virtualize the physical optical infrastructure and offer it as a service.

The authors in [43] present an Integer Linear Programming (ILP) formulation to optimally allocate offline VON requests over a transparent physical optical infrastructure. Different ILP formulations are proposed for transparent VON RA, where the Wavelength Continuity Constraint (WCC) is imposed and no 3Rs are introduced, and for opaque VON RA, where sparse 3Rs are introduced into the network. The WCC requires that the same set of assigned wavelengths is allocated on all of the requested virtual optical links. Wavelength Converters (WC), such as 3Rs, might be introduced to bypass the WCC. The authors expanded the work in transparent VON RA, for opaque VON RA in [44].

Physical optical channels (i.e., wavelengths in DWDM networks) in the same optical link interfere with each other; due to the Physical Layer Impairments (PLI). The authors in [45] introduce a Mixed ILP (MILP) formulation which takes into account PLI for offline VON requests. This work also enhances [46], where online routing algorithms where recursively applied taking into account PLI.

The authors in [47] focus on the cost-efficient embedding of on-demand VON requests for interconnecting geographically distributed data centers. A MILP formulation is presented introducing flexibility in the virtual-physical node mapping in order to optimize the usage of the underlying physical resources.

One of the main focusses of this PhD Thesis is the Resource Allocation for VONs, and it is addressed in Part II.

3.2.2 VON resource broker taxonomy

Resource brokers can be classified on several properties depending on how they provide the solution to the problem of deploying VONs on top of a common shared optical infrastructure.

One of the most important properties for a resource broker is the form it processes incoming requests, either in a static or dynamic way. Static request processing implies pre-planning of the resources, while dynamic request processing consists on the resource allocation depending on the state of the available optical infrastructure.

Another important property for the resource broker is the ownership of the optical resources, i.e. exclusive or shared. Exclusive optical resources shall be assigned to a single user/request, while shared resources can be allocated to several users/requests. Examples of possible shared optical resources are optical nodes and optical fibers, while wavelength channels within a link shall be considered exclusive.

Several interfaces to the Resource Broker can be observed, being the most common an interface for requesting virtual optical networks based on XML requests. These interfaces have not been standardized and a possible general solution would be the usage of Network Description Language (NDL), which has been successfully applied to optical networks [48].

The last important characteristic of the resource broker is the allocation of resources to deploy a control plane for each of the different virtual optical networks. The requests for VONs shall include key control plane configuration parameters.

3.2.3 Studied resource brokers

The authors in [35] propose a VON service composition framework which focuses on shared VON provisioning across multiple domains for grid applications. The presented resource management focuses on shared VON resources allowing dynamic explicit lightpath provisioning by user. The solution proposed for a single domain VON service consists of a layer acting as middleware, for which the Service Access Interface (SAI) is introduced. The SAI interacts with the Physical Optical Network Traffic Enginery Database (PON-TED) and the VON-TED. This single domain service is responsible to interact with the control plane of the infrastructure provider.

In [49], is proposed an impairment-aware VON Composition Mechanism, which is composed of an optical physical infrastructure layer (i.e. optical network resources), an optical network composition layer (i.e. where the physical resources are first abstracted and then partitioned/aggregated into virtual resources) and an optical network control and management layer. The focus of this paper is on the resource allocation of the data plane optical resources taking into account PLIs. A similar architecture has been proposed in [50] for provisioning Virtual Elastic Optical Networks (VEON).

In the studied resource brokers, the virtualization of optical networks only includes the virtualization of the optical data plane, not taking into account the necessary control plane for each deployed virtual optical network. Neither of the presented VON Resource Broker architectures take into account the necessity of a control plane for the deployed VON. This PhD Thesis proposes GMPLS-controlled virtual optical network resource broker takes into account the necessity of the provisioning of a virtual control plane, responsible of the VON, while it is also able to provision the virtual transport plane.

3.3 Control plane virtualization

A VON must be composed of not only a virtual transport plane but also of a virtual control plane, with the purpose of providing the required independent and full control functionalities (i.e., optical connection provisioning, traffic engineering, protection/restauration, etc.). There are few references in the State of the Art about the virtualization of network resources for deploying a virtualized control plane.

The GEYSERS architecture [51] proposes a Virtual IT Manager (VITM). The VITM is in charge of the end-to-end IT services management and the virtual IT resource configuration. The network resources can deploy a Network Control Plane, which performs all control and management functions necessary to operate the virtual network resources within the virtual infrastructure. The enhanced Network Control Plane proposed in GEYSERS architecture offers integrated mechanisms for Network + IT Provisioning Services (NIPS) through a service-to-network interface, named NIPS UNI during the entire VON service life cycle.

This PhD thesis will introduce a resource broker able to provision an independent virtual GMPLS control plane for each VON, in order to allow the configuration of specific internal routing policies or network addressing schemes at each virtual GMPLS-controlled WSON network. Moreover, this PhD Thesis will analyze the effect of virtualizing the GMPLS control plane, and the opportunities and drawbacks of running the control plane in virtualization servers.

3.4 Optical network virtualization with heterogeneous control domains

The authors in [30] have presented an SDN application, including several algorithms designed for different scenarios (i.e. single/ multiple fixed-grid, single/multiple flexi-grid, mixed fixed and flexi-grid), which is running on top of the OF controller to compose virtual network (VN) slices over flexi- and fixed-grid domains. The proposed SDN application is able to allocate VNs in heterogeneous networks, due to the fact of the inclusion of a extended OpenFlow Controller which integrates the OpenFlow and GMPLS control planes [52]. The OpenFlow Controller, based on the destination address of a request, identifies the end points of the lightpath in the circuit switched domain and requests a lightpath from GMPLS control plane using UNI. Once the lightpath is established the extended controller then updates the flow table of the ingress and egress switches, finalizing the establishment of an end-to-end optical flow path. The extended OpenFlow controller also maintains a list of established lightpaths in terms of lightpath Identifiers. This approach is based on optical OpenFlow virtualization, and only abstracts a GMPLS network as a single OpenFlow node.

This PhD Thesis proposes a more complex solution, allowing heterogeneous control domains (e.g., GMPLS and OpenFlow), in order to interconnect multi-site Data Centers, through multi-tenant networks using different control domains. A multi-domain resource broker will be introduced with this purpose.

Chapter 4

The ADRENALINE Testbed

Demonstration is performed by means of the ADRENALINE Testbed, a test platform designed and developed by the CTTC Optical Networking Area for experimental research on high-performance and large-scale intelligent optical transport networks. This chapter describes its main platforms.

4.1 GMPLS/PCE enabled Ethernet over WSON platform of the ADRENALINE Testbed

The ADRENALINE testbed is composed of an all-optical Dense Wavelength Division Multiplexing (DWDM) mesh network (Fig. 4.1) with two colour-less ROADM nodes and two OXC nodes, providing reconfigurable (in space and in frequency) end-to-end lightpaths, transparent to the format and payload of client signals (e.g., SONET/SDH, Gigabit Ethernet). Each optical node has two DWDM transceivers up to 2.5 Gb/s and one at 12.5 Gb/s with fully tuneable laser sources. Arrays of power meters and Variable Optical Attenuators (VOAs) are used for optical power equalization at output fibers. ADRENALINE deploys a total of 610 km of G.652 and G.655 optical fiber divided in 5 bidirectional links, in which optical amplifiers (Erbium-Doped Fiber Amplifiers -EDFAs-) are allocated to compensate power losses during optical transmission and switching at C-band. ADRENALINE transport plane also includes non-intrusive Optical Performance Monitors (OPM) to obtain spectral information tapping a 5% of all the input and output fibers, namely, channel and in-band Optical Signal to Noise Ratio (OSNR), channel and aggregate optical power, and wavelength drift.

Each optical node is equipped with aGMPLS controller for implementing a distributed GMPLS-based distributed control plane [IETF RFC3945]. Such a control plane is responsible for handling, dynamically and in real-time, the resources of the optical node in order to manage the automatic provisioning and survivability of lightpaths (using the Resource Reservation Protocol – Traffic Engineering (RSVP-TE) signaling protocol for wavelength reservation, and the Open Shortest Path First

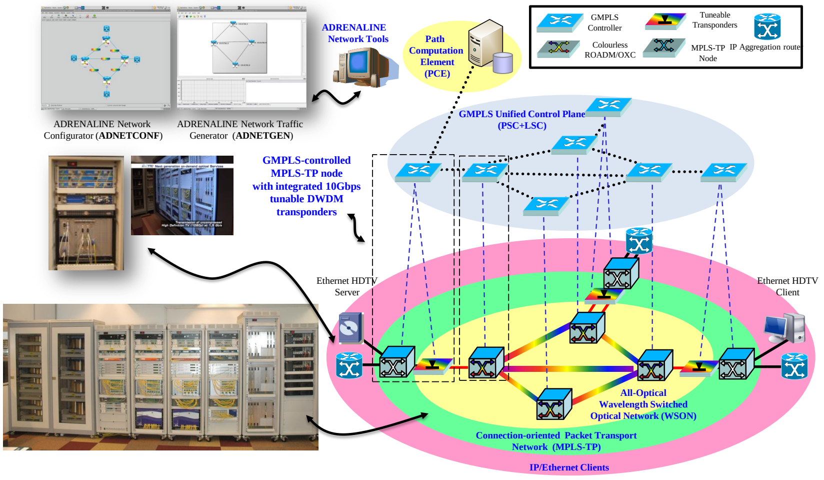


Figure 4.1: GMPLS/PCE enabled Ethernet over WSON platform of the ADRENALINE Testbed

– Traffic Engineering (OSPF-TE) routing protocol for topology and optical resource dissemination), allowing traffic engineering algorithms with Quality of Service (QoS). The system running each GMPLS controller is based on a Linux-based router with an Intel Core 2 Duo E6550 2,33 GHz processor. The DCN is based on IPCC carried at 1310 nm and C-band at the optical fiber with a line rate of 100 Mb/s using point-to-point links.

This PhD Thesis includes the evolution of the ADRENALINE Testbed from a single optical switching layer to a dual-layer architecture through the addition of MPLS-TP/PWE3 nodes with 10 Gbps tunable transceivers, which is described in chapter 12.

The ADRENALINE network includes a PCE, which is a dedicated network entity responsible for doing advanced path computations. The PCE serves requests from PCCs, and computes constrained EROs over the topology that constitutes the optical transport layer. The selected PCE deployment model is based on deploying a single PCE per OSPF-TE area, co-located in a GMPLS-enabled controller node and coupled to a Routing Controller. The preferred synchronization mechanism, by which the PCE constructs a local copy of the TED is non-intrusive: by sniffing OSPF-TE traffic, it constructs a dedicated (i.e. not shared) database using stateful inspection of the TE Link State Advertisements (LSAs) contained within the OSPF-TE Link State (LS) update messages, thus passively reusing the OSPF-TE dissemination mechanism, and not requiring the creation of an additional listener adjacency.

4.2 GMPLS/PCE control plane platform of the ADRENALINE Testbed

The GMPLS/PCE control plane platform of the ADRENALINE Testbed consists on 74 GMPLS-enabled controllers (Fig. 4.2) without associated optical hardware (i.e., the optical hardware is emulated). This set of GMPLS controllers introduces a new degree of flexibility in topology configuration, without restrictions regarding the targeted optical network topology or regarding the resources per link (e.g., number of available wavelengths, fibers, etc). Thus, the GMPLS controllers can be interconnected following any devised topology, by means of Ethernet point-to-point channels carried over emulated optical links.

The proposed solution allows the specification of control link parameters for realistic QoS constraints (fixed and variable packet delays, packet losses, bandwidth limitations, etc.). In particular, in

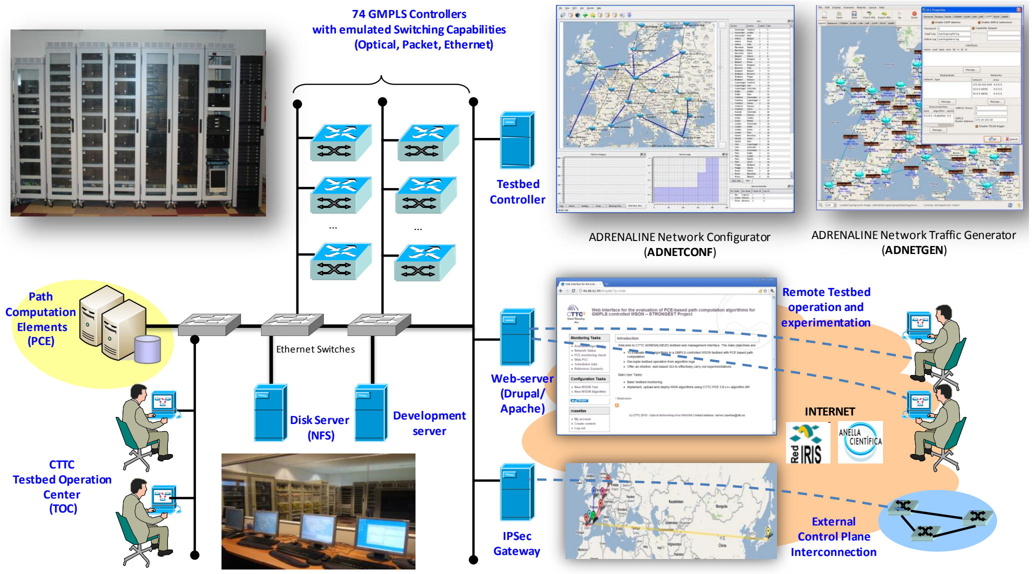


Figure 4.2: GMPLS/PCE control plane platform of the ADRENALINE Testbed

order to provide a flexible framework for topology reconfiguration, the IP Control Channels (IPCC) in the DCN are point-to-point IP interfaces over Ethernet interfaces. To do this, it uses virtual local area networks (IEEE 802.1q VLANs), configured both in the layer 2 Ethernet switches and in the GMPLS-enabled controllers within the testbed, with optional GRE or IP tunneling. This approach enables the deployment of arbitrary layer 2 interconnections between network nodes absolutely decoupled of the physical infrastructure.

The set of 74 GMPLS-enabled controllers are implemented in Linux-based routers with Intel Xeon 3.0GHz, Core 2 Duo E6550 2.33GHz and Intel Core 2 Duo E8200 2.66GHz processors. Typical values for Inter-Arrival Time (IAT) for connection requests towards the GMPLS controllers are around 3-5s. due to the necessary time to disseminate the different TE Link State Updates.

Chapter 5

PhD Thesis Objectives

In this Chapter, the objectives of this PhD Thesis are presented. This PhD Thesis focuses on experimental research, in the sense that each contribution proposed aims at being implemented in a real environment, such as the ADRENALINE Testbed. This provides important advantages such as agile prototyping, and provides architectures which can be deployed in a real word, as well as allowing to take into account specific issues or factors, that one would not consider them as long as either a simulator or emulator model is used, but result essential and determining within a real environment.

5.1 GMPLS-controlled Virtual Optical Networks

Objective 1 Design, implementation and experimental validation of an architecture for deploying GMPLS-controlled WSON as a service. Section 3.3 has exposed the need for a virtualization architecture including a control plane. To this end, an architecture for virtualizing both the data and control planes will be designed. Finally, this objective includes the development of a virtualizable GMPLS-controlled WSON platform in the ADRENALINE Testbed and its validation. This objective will be addressed in Chapter 6.

Objective 2 Design, implementation and experimental evaluation of a VON resource broker. Section 3.2 has demonstrated the need of a VON resource broker, which takes into account the virtualization of the data plane, as well as the virtualization of the control plane. Two architectures will be designed, implemented and evaluated:

- VON resource broker for VON connected subgraphs, where virtual optical links are directly mapped towards the underlying physical links.

- VON resource broker for VON Virtual Network Topologies (VNT), where the virtual optical links do not require to be directly mapped towards the underlying physical links.

Both architectures will be designed, developed and evaluated in the framework of the virtualizable GMPLS-controlled WSON platform in the ADRENALINE Testbed. This objective will be addressed in Chapter 7.

Objective 3 Resource allocation algorithms for VON connected subgraphs. The search for VON resource allocation has been described in Section 3.2.1. This objective includes a formal description of the VON RA problem. Later, an introduction of nomenclature describing the common resources and descriptors in optical networks will be detailed. VON connected subgraphs and VON VNT will be clarified and distinguished. VON RA for VON connected subgraphs will be designed, implemented in the VON resource broker and experimentally evaluated in different proposed scenarios. The proposed VON RA algorithms shall take into account the WCC. Algorithms introducing WC shall be designed, implemented and evaluated in order to ease the WCC. This objective will be addressed in Chapter 8.

Objective 4 Resource allocation algorithms for VON VNT. Different algorithms for VON VNT resource allocation shall be introduced. These algorithms will take advantage of Global Concurrent Optimization (GCO), which is performed at the PCE. The performance of the presented algorithms will be evaluated in terms of VON request blocking rate and VON setup delay. This objective will be addressed in Chapter 9.

Objective 5 Resource Allocation algorithms for VEONs. The required design modifications for the VON resource broker shall be discussed in order to provide EONs as a Service. A VON RA algorithm will be designed in order to allocate the requested transport resources in a VEON. An experimental performance evaluation shall be performed in terms of VEON request blocking rate.This objective will be addressed in Chapter 10.

Objective 6 Resource allocation algorithms for virtual GMPLS control plane. The design and validation of an architecture for the virtualization of the GMPLS control plane shall be provided. Resource allocation required algorithm for assigning the virtual control resources to the requested VON will be design, implemented and evaluated. Later, the impact of the virtualization of the GMPLS control plane will be analyzed, by providing performance results on the blocking probability and setup delays of the deployed VONs. This objective will be addressed in Chapter 11.

5.2 Virtual GMPLS-controlled MPLS-TP Networks over a shared WSON

Objective 9 Design, implementation and performance evaluation of a GMPLS-controlled MPLS-TP node. The architectural design of a MPLS-TP node with tunable 10Gbps transponders shall be described. Later, the implementation using commercial off-the-shelf hardware will be discussed. Finally, an evaluation of the node will be provided by means of analyzing the obtained throughput and CPU usage in different evaluation scenarios with different traffic grooming and traffic aggregation strategies. This objective will be addressed in Chapter 12.

Objective 10 Virtualization of a GMPLS-controlled MPLS-TP node. This objectives includes the design and implementation of the required MPLS-TP node virtualization, by partitioning its interfaces. This objective will be addressed in Chapter 12.

Objective 11 The design, implementation and evaluation of a resource broker for virtual MPLS-TP networks over a shared WSON. The proposed virtual MPLS-TP networks will overcome the limitations of the current Data Center interconnection infrastructure. This objective will be addressed in Chapter 13.

5.3 Optical Network Virtualization with Heterogeneous Control Domains

Objective 12 The design, implementation and evaluation of a virtualization orchestrator for the composition of multiple VONs with heterogeneous control domains (i.e., GMPLS and OpenFlow). Experimental assessment will be provided in terms of VON request blocking rate and setup delay. This objective will be addressed in Chapter 14.

Part II

GMPLS-controlled Virtual Optical Networks

Chapter 6

Deploying GMPLS-controlled WSON as a service

The objective of this chapter is to design, implement and evaluate a virtualizable GMPLS-controlled WSON platform in the ADRENALINE Testbed to deploy multiple virtual networks with full control of the internal routing policies or addressing schemes.

6.1 Architectures for virtualizing the GMPLS control plane

Optical node virtualization is the creation of a virtual representation of an optical network node (i.e., ROADM or OXC) with the same functionalities, based on an abstract model that is often achieved by partitioning and/or aggregation. A virtual optical link is defined as a connection between a port of a virtual optical network node to a port of another virtual optical network node. A physical optical link can be partitioned based on a TDMA, OFDMA or WDMA basis. Finally, a VON is the composition of isolated virtual optical resources (i.e., virtual nodes and links), which exist simultaneously over the same physical optical network infrastructure (Fig. 6.1).

A VON does not only consist of a transport plane, but also a control plane is need. Some authors have proposed the virtualization of optical network infrastructure by the creation of a middleware layer (e.g., an optical network service plane) to abstract and virtualize the network resources, but the optical network infrastructure (both at data and control plane level) remains unmodified [53].

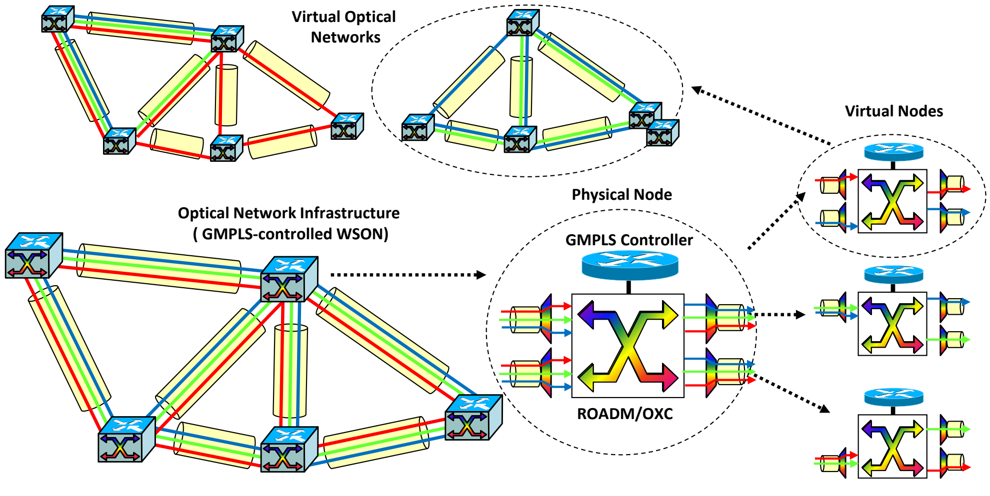


Figure 6.1: Virtual Optical Node and Link Partitioning and Aggregation

One of the main limitations of this approach is that the same instance of the GMPLS control plane is shared by all the virtual networks. Thus, there is no way to configure specific internal routing policies or network addressing schemes at each virtual GMPLS-controlled WSON network (Fig. 6.2, left).

The proposed network virtualization architecture is based on virtualizing both the GMPLS control and the WSON data planes. The virtualization of the GMPLS control plane allows creating virtual GMPLS connection controllers that can be used to deploy virtual networks with their own virtualized GMPLS control plane instance (Fig. 6.2, right). This approach allows service providers to have full control over the deployed virtual networks.

6.2 Virtualization of the GMPLS control plane

The virtualization of a distributed GMPLS-based control plane (which is responsible for handling, dynamically and in real-time, the resources of the optical node in order to manage the automatic provisioning and survivability of lightpaths [13]), relies on the usage of the GNU/Linux kernel KVM virtualization technology and the virtualization capabilities of the CPUs.

A non-virtualized GMPLS Controller (Fig.6.3) includes:

- RSVP-TE signaling protocol agent for wavelength reservation.
- OSPF-TE routing protocol agent for topology and optical resource dissemination.
- Management of optical resources (LRM).
- Hardware Abstraction Layers (HAL) managing the CCI interface.
- SNMP agents for SNMP based network management.

The Data Communication Network (DCN) is based on IP control channels (IPCC) carried at 1310 nm within the C-band at the optical fiber with a line rate of 100 Mb/s using point-to-point link with

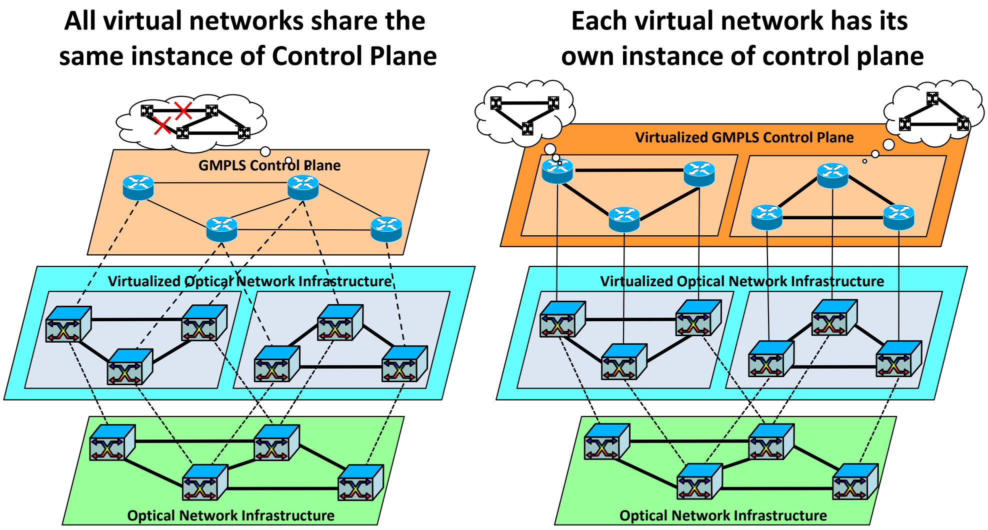


Figure 6.2: Architectures for Virtualizing the GMPLS Control Plane

GRE tunneling. Once virtualized, the same physical machine hosts a configurable number of virtual guests (Fig.6.4), each replicating a GMPLS controller.

A VON resource manager will be responsible for assigning the allocated resources for each VON to each of the Link Databases (LD) and a dedicated DCN will be deployed as an overlay on top of the shared physical infrastructure via KVM bridging modules, which allow connecting the physical Network Interface Card (NIC) that provides connectivity to a neighboring node to each of the virtualized network interfaces, resulting in isolated IPCCs.

6.3 Virtualization of the WSON transport plane

The ROADMS and OXCs of the ADRENALINE testbed are composed of three main optical systems (Fig. 6.5), namely, optical fiber switches, add-drop stages, and tunable optical transceivers. The optical fiber switches, acquired from external providers (Glimmerglass Networks and Polatis), already support virtualization, through the possibility of a control mechanism that allows to be called from multiple instances.

As for the add-drop stage of ROADM nodes, designed and developed by CTTC, the controller cards have been reprogrammed in order to support virtualization, being able to handle multiple TCP connections from virtual GMPLS controllers. Finally, the tunable optical transponders (Proteus Systems, acquired from W-OneSys), do not support virtualization because they only handle serial protocols (RS-422 and I2C). A new instance (optical transceiver system controller) has been introduced for each Proteus System, in order to work concurrently with several TCP connections from virtual GMPLS controllers. This new instance serializes the different isolated instructions received from the virtual GMPLS controllers, allowing the optical fiber switch to be called from multiple instances.

6.4 VON resource manager: ADNETCONF

ADNETCONF is the network configuration and deployment tool for the ADRENALINE Testbed [54]. ADNETCONF is used to design a particular network topology and configure the GMPLS-related

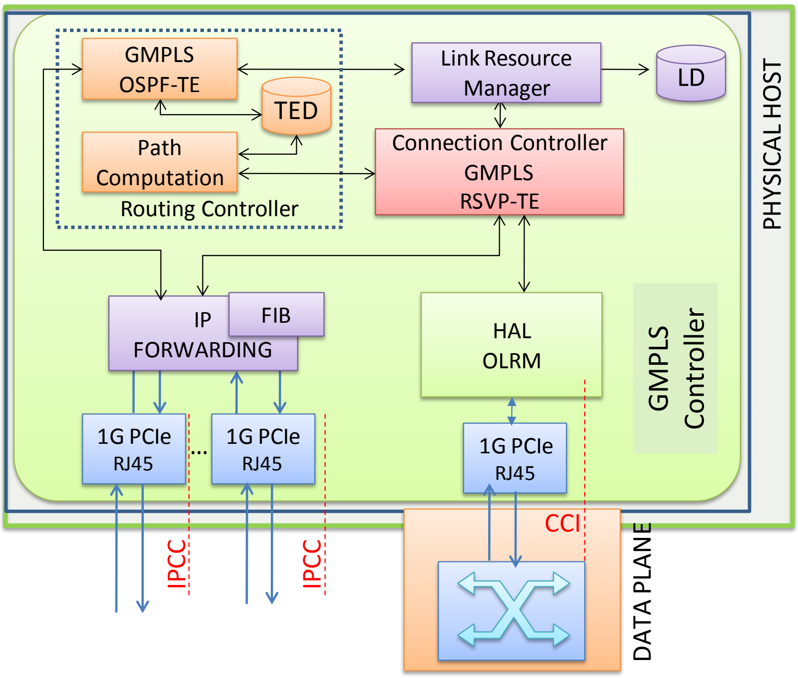


Figure 6.3: GMPLS Controller

processes (LRM, OSPF-TE, RSVP-TE, SNMP and OLRM), and the network scenario is deployed through the creation and usage of the corresponding XML configuration files. ADNETCONF has been extended to handle different virtual networks scenarios concurrently (Fig. 6.6). This required refactoring ADNETCONF from a Single Document Interface (SDI), where a single scenario could be defined, to a Multiple Document Interface(MDI).

ADNETCONF is able to manage the virtualized infrastructure resources (both at the control and data plane) with a fine-grained granularity covering optical nodes, full TE links and wavelength channels within a link. As a virtual network manager, ADNETCONF connects to the virtual GMPLS controllers and establishes GRE tunnels over the virtual Ethernet interfaces (that share the same physical connection through a KVM bridge), in order to isolate the different virtual scenarios. Moreover, ADNETCONF is able to cross-check the defined virtual scenarios with the physical resources in order to detect mismatches, or pre-reserved elements, effectively acting as a resource manager.

In next chapters, ADNETCONF will be used inside the VON resource broker in order to deploy the different instances of VONs (including each data and control planes).

6.5 Experimental performance evaluation

The proposed architecture is validated by means of the virtualizable GMPLS-controlled WSON platform of the ADRENALINE Testbed, which is a GMPLS-based Intelligent Optical Network composed of an all-optical WSON with 2 ROADMs and 2 OXCs providing reconfigurable (in space and in frequency) end-to-end lightpaths, deploying a total of 610 km of G.652 and G.655 optical fiber (Fig. 6.7).

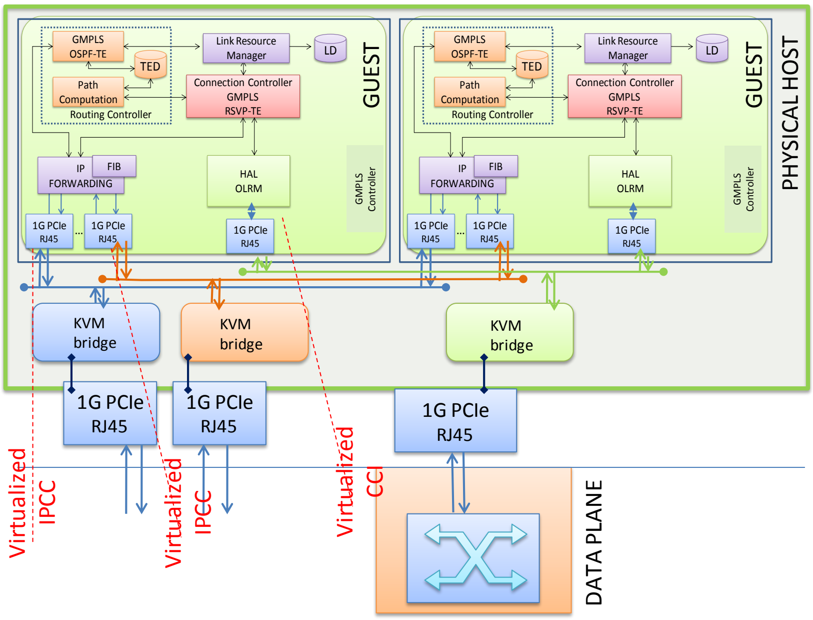


Figure 6.4: Virtual GMPLS Controller

Each optical node is equipped with a virtualization server running in a Linux-based router with an Intel Core 2 Duo E6550 2.33 GHz processor.

The experimental demonstration consists in, first, validating the virtualization of the real network infrastructure (Fig. 6.8), by means of the ADNETCONF tool in order to create two independent virtual network scenarios with the same number of nodes, links and topology connectivity as in the real network infrastructure but having half of the wavelength channels on each of their optical links.

Additionally, each physical machine needs to be configured to host two virtual GMPLS Controllers. Once the virtual network scenarios are set up, the performance is quantitatively evaluated, summarizing the results in Table 6.1. Performance values were obtained when dynamically provisioning and restoring connections in three network scenarios: real optical infrastructure and the two

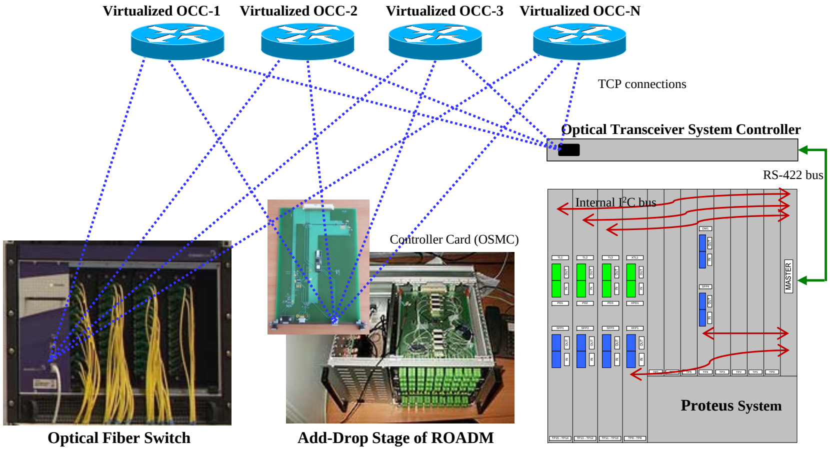


Figure 6.5: Virtualized optical systems

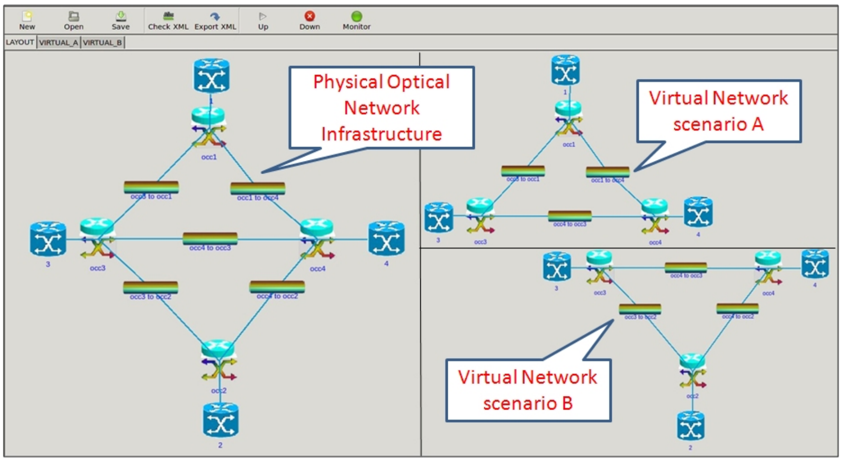


Figure 6.6: Virtual Network Resource Manager

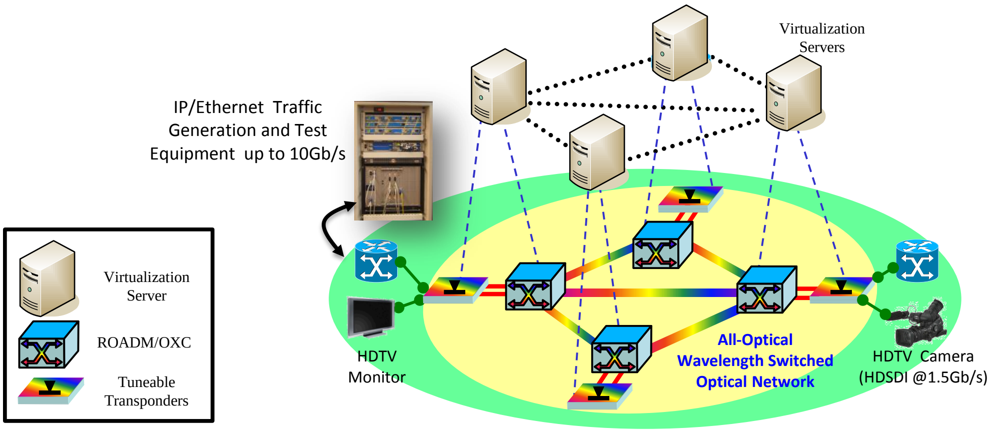


Figure 6.7: ADRENALINE Testbed Architecture

virtual networks (referred to as virtual A and B) using a stochastic model of lightpath requests and link failure events, in which the inter-arrival process is Poisson, and the Holding Time (HT) follows a negative exponential distribution, with requests (failures) events uniformly distributed among all distinct links.

For the LSP restoration, the link failure Inter-Arrival Time (IAT) was set to 30 s with a mean link failure HT set to 15 s. Each data point is obtained requesting 10^3 connections.

Both the average provisioning Setup Delay (SD) and the Restoration Time (RT) in virtualized scenarios increase with respect to the one attained in the real network infrastructure. This result is caused by the contention over the physical resources (CPU, hard disk, memory, and NIC) that is generated by the virtual GMPLS Controller.

Although in relative terms the values are noticeably increased, they still remain within acceptable values when operating a virtualized service network.

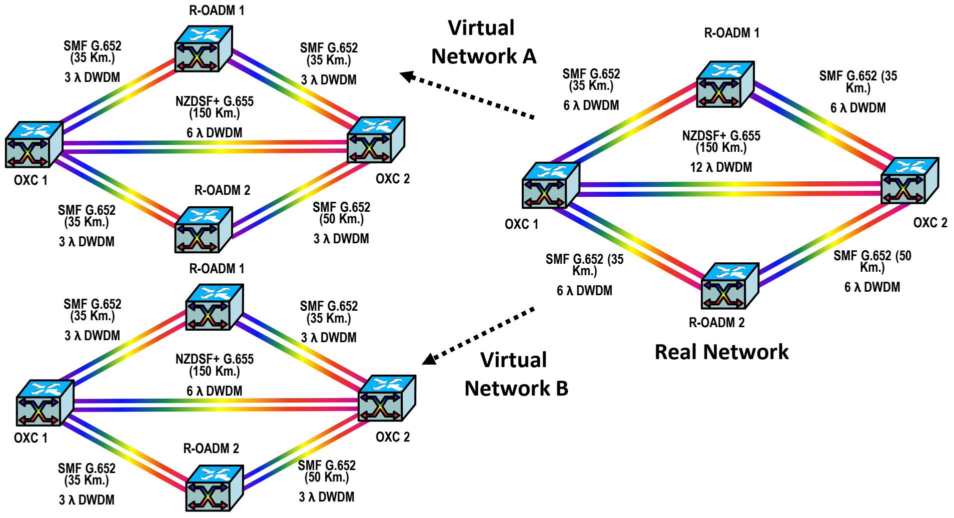


Figure 6.8: ADRENALINE Testbed Virtualization

Table 6.1: Numerical results provisioning/restoration

	Provisioning				Restoration				
Scenario	IAT (s.)	HT (s.)	Traffic Load (Er.)	SD (ms.)	IAT (s.)	HT (s.)	Traffic Load (Er.)	SD (ms.)	RT (ms.)
Real network	3	20	6.67	14.7	3	100	33.3	18.9	40.4
Virtual A	3	10	3.33	24.8	3	50	16.67	28.2	50.1
Virtual B	3	10	3.33	25.1	3	50	16.67	29.7	49.9

6.6 Conclusions

A network virtualization architecture has been presented for deploying GMPLS-controlled WSON networks as a service. Experimental evaluation of the deployed virtual networks have been carried out in the ADRENALINE testbed, assessing the feasibility of the proposed architecture with a low impact on the performance in terms of provisioning and recovery delay.

Next chapter will introduce a VON Resource Broker in order to dynamically provision virtual GMPLS-controlled WSONs as a service.

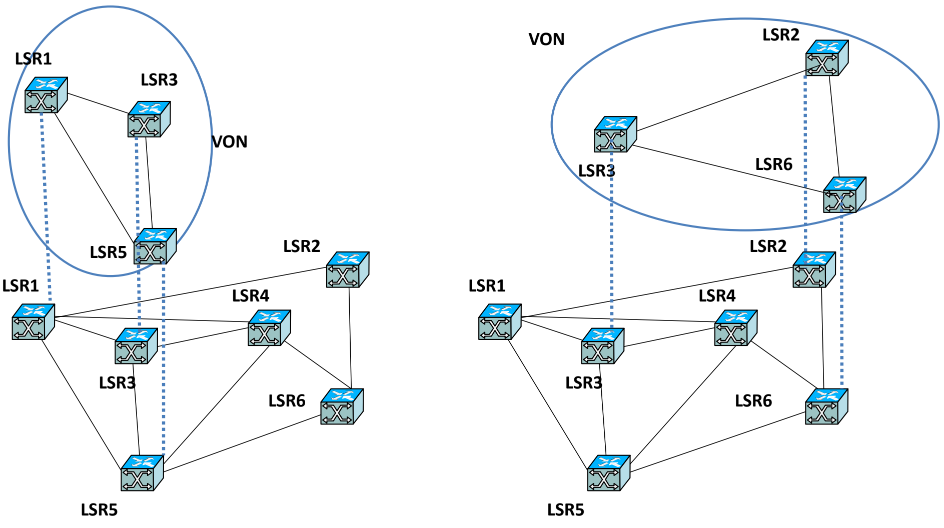
Chapter 7

Virtual optical network resource broker

Chapter 3 has explained how optical network virtualization supports the heterogeneous and stringent infrastructure network requirements of the emerging dynamic and bandwidth-hungry applications such as high-definition video streaming and cloud computing. Thus, service providers can dynamically request, on a per need basis, a dedicated VON for each application and have full control over it.

The objective of this chapter is to design, implement and evaluate a VON resource broker for deploying virtual GMPLS-controlled WSONs. The introduction of a VON resource broker will provide the required dynamicity for service providers. Two different VON network topologies will be deployed. Firstly, a VON connected subgraph is a VON which its virtual optical links are directly mapped towards the underlying physical links (Fig. 7.1.a). A VON Virtual Network Topology (VNT) does not require its virtual links to be included in the set of physical links (Fig. 7.1.b).

The work presented in [55] focuses on the design, implementation and validation of a virtualization architecture for deploying VON connected subgraphs. To overcome this limitation, in [56], a VON resource broker which includes a Virtual Network Topology Manager (VNTM) has been proposed. In this chapter both architectures will be described including its design, implementation, evaluation and comparison.



(a) VON connected subgraph (b) VON Virtual Network Topology

Figure 7.1: Comparison between VON connected subgraph and VON Virtual Network Topology (VNT)

7.1 VON Resource Broker for deploying VON connected subgraphs

In this section the design, implementation and assessment of a VON Resource Broker for deploying VON connected subgraph are presented.

7.1.1 System Architecture

Figure 7.2 shows the proposed system architecture, including the four main blocks of the proposed VON resource broker:

- Resource manager
- VON controller
- Resource allocator
- Resource configurator

7.1.1.1 Resource manager

The resource manager handles both the virtual control and transport resources. As for the virtual control resources, it manages the available IP subnetworks that shall be used to establish the virtual IPCC to later deploy dedicated DCN for each virtual transport plane. It is also responsible for managing the number and location of the available virtual GMPLS controllers on the virtualization servers in the ADRENALINE Testbed (Fig. 7.3), as well as their configuration information (including the management IP address, the amount of CPU power and available RAM). The virtualization of a GMPLS Controller has been previously described in chapter 6.

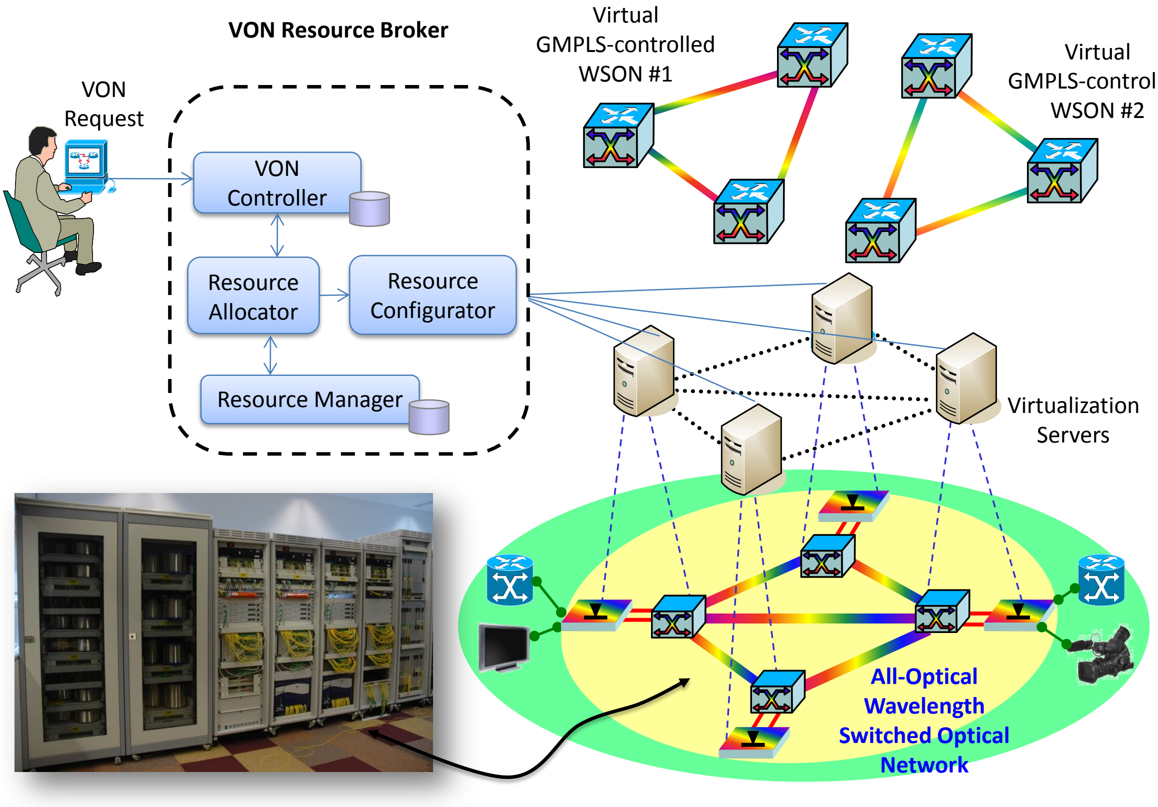


Figure 7.2: System Architecture for VON Resource Broker for deploying VON connected subgraphs

The resource manager also stores all the information required to configure the processes running in the virtual GMPLS controllers. For example, it stores the IP addresses and users (login and passwords) of the optical switches (each switch has a maximum number of users that equals to the maximum number of partitions supported), the set of available GMPLS router addresses and identifiers, the path computation algorithm identifiers, the set of wavelength assignment algorithm identifiers or the default protocol timers (e.g., hello, refresh).

Transport plane resources are described as optical links and switches. Each optical switch has a determined number of input and output ports and a maximum number of partitions is feasible. The specific partitioning of an optical switch is performed dynamically, based on the requirements of the requested VON. Each optical link stores the information about the edge optical switches the link is connected to, the number of supported wavelengths and their identifiers, as well as any link parameter (i.e. optical impairments). The partitioning of an optical link is on a wavelength granularity basis.

7.1.1.2 VON controller

The VON controller accepts incoming TCP sessions, which include VON requests, and handles these requests asynchronously and dynamically. Once the VON identifier, which is a unique identifier per each VON, is assigned, or found in the request database, the VON controller triggers the resource allocator in order to process the VON request, which consists on the allocation of the new VON, (i.e.,

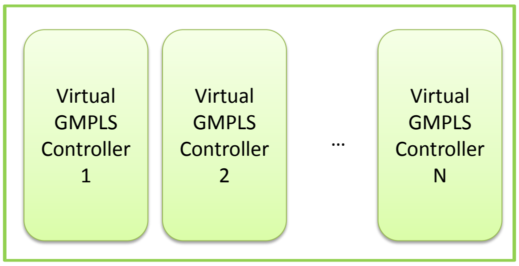


Figure 7.3: Virtualization Server for VON connected subgraphs

VON deployment) the modification of resources assigned to an existing VON or the releasing of the resources in case a VON is torn down.

A VON request is modeled as a graph that describes a set of virtual optical switches and links for the virtual transport plane, specifying for each one the number of requested input and output optical ports, and the number of wavelengths respectively. The VON request also includes some requirements for the virtual control plane, such as the needed capacities (i.e. CPU power, RAM) for the virtual GMPLS controllers, or the configuration values for the control processes of the virtual GMPLS controllers, which can be later modified by the service provider.

7.1.1.3 Resource allocator

The resource allocator assigns the virtual transport and control resources to the requested VON. For the virtual control plane, it allocates the virtual GMPLS controllers, based on the requested CPU power and RAM, and assigns the GMPLS router address. It also assigns IP addresses and GRE tunnels for the required IPCCs (it is assumed that an IPCCs are assigned following the same topology as the transport plane).

The configuration values for all GMPLS processes running in a virtual GMPLS controller are also assigned. For example, for the HAL process, it is specified the set of allocated wavelengths per link. The configuration for the LRM process is generated, including for each TE link the switching capabilities and the maximum and available bandwidth, depending on the allocated wavelengths. The OSPF-TE process is configured with the adjacent router addresses and the interfaces that shall be used for the VON topology and resource dissemination.

For the virtual transport plane, the target is to assign the requested number of available wavelengths (N) for each virtual optical link of the requested VON. Several wavelength assignment schemes can be applied, such as first-fit (select the first N wavelengths being available and common in all requested virtual optical links).

7.1.1.4 Resource configurator

The resource configurator generates the virtual transport and control plane configuration XML file, which describes a VON scenario model that can be set up, modified or torn down by means of ADRENALINE Network Configurator (ADNETCONF) [54], which is a software tool in charge of scenario model management in ADRENALINE Testbed. With ADNETCONF, the scenario model is then serialized to the formal representation of the scenario that the processing engine understands. Up to five different XML files are produced; one describing the logical DCN topology for the virtual control plane, and the others describing the configuration of the different GMPLS processes.

Table 7.1: Numerical results

VON request load (Er.)	Treq (us.)	Tup (s.)	stdev up (s.)	Tdown (s.)	stdev down (s.)
1	131	11.76	3.38	5.18	1.95
5	126	10.13	4.31	5.68	7.80
10	111	11.04	3.48	4.78	1.63

7.1.2 Experimental assessment

The experimental demonstration consists in validating the VON resource broker architecture by requesting different VONs to be deployed on the ADRENALINE Testbed. The ADRENALINE Testbed has been previously introduced in chapter 6.

The performance values are obtained when dynamically provisioning VONs, in which the inter-arrival process is Poisson, and the HT follows a negative exponential distribution. The average VON HT is set to 1000 s, and the mean IAT will be varied during all different experimentations, yielding an offered traffic load from 1 to 10 Erlangs. The topology for each VON request is randomly selected from the space of feasible topologies, which is determined by the physical optical network infrastructure. For the ADRENALINE Testbed, a set of 28 topologies has been considered. However, the VON request should be randomly generated, without being constrained by the physical network, that is why in next section a VON Resource Broker with VNT is introduced.

The number of wavelengths requested for each VON is set to one, and the wavelength assignment algorithm used is first-fit. In chapter 8, more insight on VON resource allocation algorithms for VON connected graphs is introduced. 100 requests have been preformed for each data point. The provided results are request processing time (Treq) at the VON resource broker, the VON setup and tear down average delays employed by ADNETCONF to configure both the control and data plane, along with their standard deviation (Tup and Tdown, respectively). The obtained results are shown at Table 7.1. The VON setup and tear down delay for the ADRENALINE Testbed is around 11 and 5 seconds respectively, whilst the request processing time is of the order of few ms per request.

7.2 VON resource broker for deploying VON Virtual Network Topologies

In this section the design, implementation and assessment of a VON resource broker for deploying VON VNT are presented. The proposed resource broker includes a VNTM to dynamically deploy virtual GMPLS-controlled WSON networks. This VNTM communicates with a GMPLS/PCE control plane which controls the physical WSON infrastructure in order to construct the requested virtual optical links as a set of LSPs and the VNTM will offer the established LSPs as virtual TE links for the allocated virtual GMPLS control plane.

This section extends the obtained results in [56], and introduces Global Concurrent Optimization (GCO), which allows the PCE to perform concurrent path computations, where a set of paths are computed concurrently in order to efficiently utilize the network resources. GCO allows the concurrent path computation of all the requested virtual optical links in order to perform an efficient VON RA. A GCO algorithm is used to evaluate the performance of a concurrent path computation for the requested virtual optical links. The VON Resource Broker with VNTM system architecture is experimentally assessed and the proposed system architecture on the ADRENALINE Testbed is evaluated, by providing performance results such as the VON setup delay.

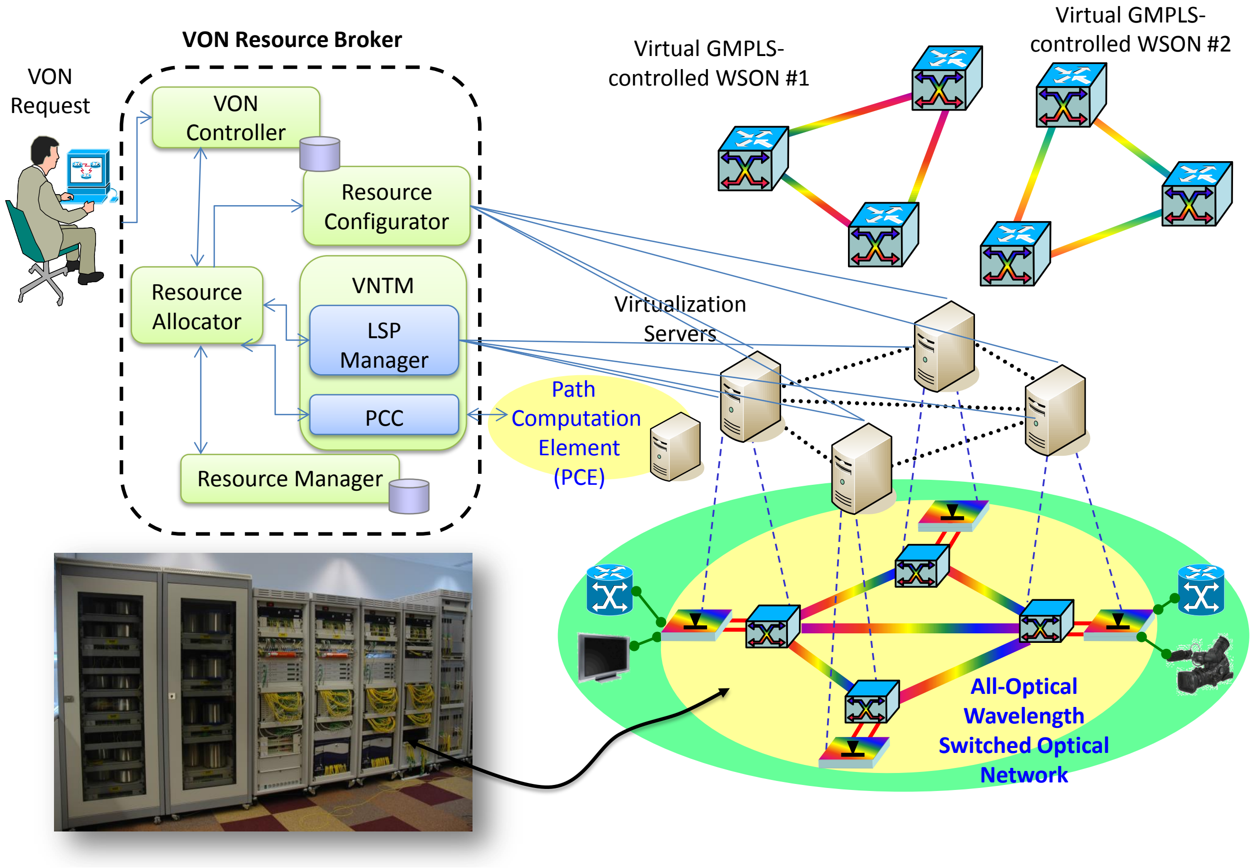


Figure 7.4: System architecture for VON resource broker for deploying VON Virtual Network Topologies

7.2.1 System Architecture

Figure 7.4 shows a virtualizable GMPLS/PCE-controlled WSON network managed by the proposed resource broker. The resource broker is the responsible for managing the incoming asynchronous and dynamic VON requests, which include a request for the allocation of the new VON, the modification of resources assigned to an existing VON or the releasing of the resources in case a VON is torn down.

A VON request is modeled as a graph that describes a set of virtual optical switches and links for the virtual transport plane, specifying for each one the number of requested input and output optical ports, and the number of wavelengths, respectively. The VON request also includes some requirements and constraints for the virtual control plane, such as the needed capacities for the virtual GMPLS controllers, or the selected values for configuring the parameters/attributes of the control processes running on the virtual GMPLS controllers, which can be later modified by the service provider. Internally, the resource broker is composed of 5 modules:

- VON Controller
- Resource Manager
- Resource Allocator

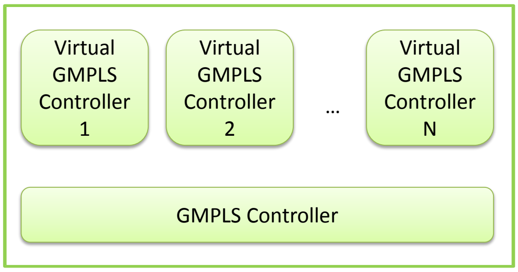


Figure 7.5: Virtualization Server for VON VNT

- VNT Manager
- Resource Configurator

Some of the modules share the same functionalities with the modules presented in the previous section, so in this section, only the different functionalities will be introduced.

7.2.1.1 Resource manager

The resource manager is a virtual resource database which handles the virtual partitioned resources. For the control plane, it is able to manage the same resources as in the previous resource broker. It also stores all the information required to contact the underlying GMPLS controllers of the physical infrastructure (Fig. 7.5). For the data plane, it maintains a copy of the physical network resources and topology stored in the TED, and performs sanity checks for the independent coexisting VONs. These are performed to ensure that the PCE has correctly allocated the data plane resources.

7.2.1.2 VON controller

The VON controller accepts incoming TCP sessions from a VON request client, used to reliably transport VON requests, and handles these requests asynchronously and dynamically. Once the VON identifier is negociated, the VON controller triggers the resource allocator in order to process the VON request.

A VON request consists of a XML file, describing the requested virtual nodes, the requested virtual links between these nodes and the required number of wavelength channels.

7.2.1.3 Resource allocator

The resource allocator assigns the control resources to the requested VON. For the virtual control plane, it allocates the virtual GMPLS controllers, and assigns the GMPLS router address. It also assigns IP addresses and GRE tunnels for the required IPCC.

7.2.1.4 Virtual Network Topology Manager

The VNTM is composed of a Path Computation Client (PCC) and a Label Switched Path Manager (LSP Manager). Using the GCO, all the requested virtual optical links are grouped into a Path Computation Request (*PC_Request*) for the PCE. Each requested virtual optical link includes the requested number of wavelengths. The PCE will reply with a Path Computation Reply (*PC_Reply*), which includes Explicit Route Objects (ERO). A virtual optical link can be defined in one or more EROs (one ERO per requested virtual link and wavelength).

Each ERO is used by the LSP Manager to request the necessary LSPs through the GMPLS controller (i.e., RSVP-TE protocol) that is running on the source node of each requested virtual link. Once each LSP has been established, the RSVP-TE protocol answers with a Record Route Object (RRO). Once all the required LSPs for a virtual optical link have been established, they are offered as a virtual optical TE link to the virtual GMPLS control plane, which is configured to map the virtual TE link to the physical port of the optical node (i.e., ROADM or OXC).

7.2.1.5 Resource configurator

The resource configurator generates the virtual transport and control plane configuration XML file, which describes a VON scenario model that can be set up, modified or torn down by means of ADRENALINE Network Configurator (ADNETCONF), which has been described in Section 6.4.

7.2.2 Experimental assessment

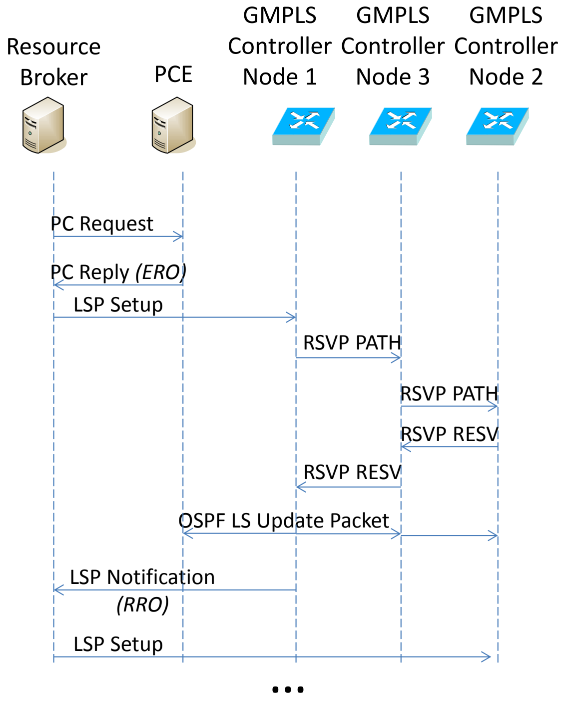


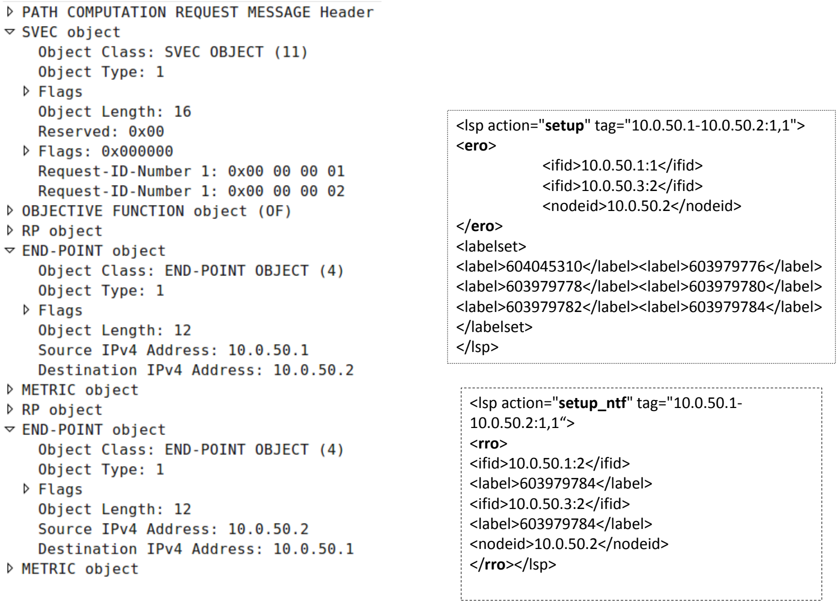
Figure 7.6: Resource broker with VNTM message exchange

In this section an experimental assessment and a performance evaluation are presented for the proposed system architecture on the virtualizable GMPLS-controlled WSON platform of the ADRENALINE Testbed [57], described in Section 6.5.

Figure 7.6 shows protocol details on the message exchange for VON request and provisioning, including the PCE request and reply for VON resource assignment, VNTM XML protocol for LSP request at the GMPLS-controlled nodes and LSP establishment (RSVP-TE) messages.

As an example, a VON is requested consisting of 2 virtual optical links between nodes 1 and 2 (i.e., 1-2 and 2-1) (see Fig. 7.6). The links are considered unidirectional, for simplicity. With the purpose of obtaining a path for the requested virtual optical links, the resource broker issues to the PCE a *PC_Request* message including a Synchronization VECtor (SVEC) object [58]. When a PCE computes sets of dependent path computation requests concurrently, use of the SVEC list is required for association among the sets of dependent path computation requests. The SVEC object is carried within the Path Computation Element Communication Protocol (PCEP) *PC_Request* message. In the SVEC object the different requested links are enumerated (Fig. 7.7.a). In this example, two END-POINT objects are requested, including source and destination.

As an example, the PCE uses the Shortest Path Unreserved Bandwidth First-Fit (SPUB-FF) algorithm, which consists on computing a constrained shortest path algorithm for each requested virtual optical link. SPUB-FF only takes into account the unreserved bandwidth. Once the spatial paths for the virtual optical links have been computed, the common set of available wavelengths of the computed paths are obtained, in order to overcome the WCC. The WCC requires that the same set of assigned wavelengths is allocated on all of the requested virtual optical links. The First-Fit (FF) wavelengths of the common set of available wavelengths of the computed paths are assigned. The PCE replies with a *PC_Reply* message which includes the EROs (spatial path) and the possible wavelengths (i.e., labelset), for each virtual optical link.



```
▷ PATH COMPUTATION REQUEST MESSAGE Header
▽ SVEC object
    Object Class: SVEC OBJECT (11)
    Object Type: 1
  ▷ Flags
    Object Length: 16
    Reserved: 0x00
  ▷ Flags: 0x000000
    Request-ID-Number 1: 0x00 00 00 01
    Request-ID-Number 1: 0x00 00 00 02
▷ OBJECTIVE FUNCTION object (OF)
▷ RP object
▽ END-POINT object
    Object Class: END-POINT OBJECT (4)
    Object Type: 1
  ▷ Flags
    Object Length: 12
    Source IPv4 Address: 10.0.50.1
    Destination IPv4 Address: 10.0.50.2
▷ METRIC object
▷ RP object
▽ END-POINT object
    Object Class: END-POINT OBJECT (4)
    Object Type: 1
  ▷ Flags
    Object Length: 12
    Source IPv4 Address: 10.0.50.2
    Destination IPv4 Address: 10.0.50.1
▷ METRIC object
```

```
<lsp action="setup" tag="10.0.50.1-10.0.50.2:1,1">
<ero>
        <ifid>10.0.50.1:1</ifid>
        <ifid>10.0.50.3:2</ifid>
        <nodeid>10.0.50.2</nodeid>
</ero>
<labelset>
<label>604045310</label><label>603979776</label>
<label>603979778</label><label>603979780</label>
<label>603979782</label><label>603979784</label>
</labelset>
</lsp>
```

```
<lsp action="setup_ntf" tag="10.0.50.1-
10.0.50.2:1,1“>
<rro>
<ifid>10.0.50.1:2</ifid>
<label>603979784</label>
<ifid>10.0.50.3:2</ifid>
<label>603979784</label>
<nodeid>10.0.50.2</nodeid>
</rro></lsp>
```

(a) Wireshark Capture of a Path Computation Request

(b) LSP Setup XML messages

Figure 7.7: Messages interchanged during VON deployment

Table 7.2: Pros and Cons of VON resource broker for VON connected subgraphs and VON resource broker for VON VNT

	VON connected subgraph	VON VNT
Pros	- Lower VON setup delays - Less complexity	- Higher VON flexibility - No need to directly configure optical equipment
Cons	- Only physical links can be requested	- Higher VON setup delays

The *PC_Reply* message is processed and, through the LSP manager, an LSP is requested through an XML proprietary interface to the RSVP process of the source node (Fig. 7.7.b). Once a setup LSP request message is received at node 1, a RSVP PATH message is issued to node 2, following RSVP standard procedure. Node 2 responds with RSVP RESV message which includes the RRO, and when the message is received at node 1, the optical resources for the LSP have been occupied. Finally, node 1 sends to the resource broker an acknowledgement for the setup of the LSP. The LSP establishment is performed for all the requested virtual optical links and the requested wavelengths on each optical link.

Once all the necessary LSPs have been established, the necessary virtual GMPLS control plane resources need to be allocated. To this end, the established LSPs are used as virtual TE links. Finally, once the VON configuration is generated, the VON configurator, by means of ADNETCONF, is the responsible to set up or tear down the requested VON. The virtual GMPLS-controlled WSON setup and tear down delay for the ADRENALINE Testbed are 17s and 7s, respectively.

7.3 Conclusions

Table 7.2 compares the advantages and disadvantages of using either a VON resource broker for VON connected subgraph or a VON resource broker for VON VNT.

A VON resource broker for VON connected subgraphs has lower setup delays and less complexity. On the other hand, virtual optical links which are included in the physical infrastructure can be requested. Instead, a VON resource broker with VNTM does not need to directly configure the optical equipment, as the underlying GMPLS control plane is the responsible for doing so. This architecture provides more flexibility on the VON requests, at the cost of providing higher VON setup delays.

A VON resource broker with two different architectures have been presented which act as interface between service providers and infrastructure providers to deploy virtual GMPLS-controlled WSON infrastructure services. Experimental assessment of both architectures has been carried out in the ADRENALINE testbed and has shown the feasibility of deploying independent instances of virtual GMPLS-controlled WSON, providing low delays for VON setup and tear down.

In the next chapters, the VON RA problem is presented and several VON RA algorithms are proposed and studied. The following chapter analyzes VON RA algorithms which consider a VON topology which is a connected subgraph of the Physical Infrastructure, and later next chapters focus on VON RA without this limitation.

Chapter 8

Virtual Optical Networks transport resource allocation: connected subgraphs

In this chapter, different resource allocation algorithms are proposed and analyzed. These algorithms which are run by the VON resource broker, as explained in the previous chapter, responsible for assigning the requested VON resources (i.e., virtual optical nodes and links), focusing on VON transport plane allocation and deployment.

One of the key elements for the VON RA for transport plane is the wavelength allocation algorithm, which allocates the requested wavelengths for each virtual optical link requested in a VON request. In this chapter, it is assumed that a VON request consists on a request for virtual optical nodes and links and the requested virtual links correspond to links of the physically partitioned infrastructure, also known as a VON connected subgraph. As expected, the selection of an heuristic algorithm will result on a big impact on the Blocking Rate of the requested VONs.

Two different types of algorithms are introduced:

- The former satisfies the WCC in transparent WSONs using FF or Random (RND) wavelength allocation for the different VON requests. The WCC requires that the same wavelength is allocated on all of the requested virtual optical links.
- The second type of algorithm allocates Wavelength Converters (WC) to bypass the WCC in case no transparent wavelengths can be assigned to the requested VON, relaxing the constraint.

This chapter focuses on the analysis of both types of wavelength allocation algorithms to be run by the VON resource broker and their impact on the blocking probability on different selected network topologies.

8.1 Virtual Optical Networks Resource Allocation Definitions

In this section, the different required definitions for describing the optical network resources are presented. To keep this PhD Thesis consistent, these definitions are shared between the different chapters.

The physical optical infrastructure is modeled as a directed graph and denoted by $G^P = (V^P, E^P)$, where V^P is the set of physical optical nodes and E^P is the set of physical optical links.

Each optical link $e^P(i, j) \in E^P$ between two physical optical nodes e_i and e_j is associated with a certain number of available wavelengths $w(e^P)$.

A VON request consists on a directed graph denoted by $G^V = (V^V, E^V)$ and a unsigned integer value of requested wavelengths $\mathcal{W}$. The requested set of virtual links is a subset of the available physical links, meaning G^V is a connected subgraph of G^P ($E^V \subseteq E^P$). It can be observed that in a VNT $E^V \nsubseteq E^P$.

The following functions are defined with the purpose of helping to obtain certain parameters of the optical infrastructure:

- $w(e)$ $\forall e \in E^P$ corresponds to the available wavelengths for edge e.
- FF($w(e)$, $\mathcal{W}$) is defined as the function that returns the first $\mathcal{W}$ wavelengths of $w(e)$ and triggers an error in case not enough wavelength are available.
- RND($w(e)$, $\mathcal{W}$) selects $\mathcal{W}$ random wavelengths from the input vector $w(e)$ with a uniformly distributed probability.
- $a(e)$ $\forall e \in E^V$ assigned wavelengths for edge e.
- $visit(e)$ $\forall e \in E^V$ checks if a wavelength has been assigned for edge e.
- $out(e)$ $\forall e \in E^V$ includes all requested virtual links that have as source node e_j.

8.2 VON RA algorithms that satisfy the Wavelength Continuity Constraint

Two different algorithms are used to assign the requested number of wavelengths for the requested virtual optical networks by the proposed VON resource broker. Both algorithms share the common procedure by which they obtain the common available wavelengths in all the requested virtual optical links. The algorithm 1, named TWA-FF, selects the first common available wavelengths for all the requested virtual optical links, while algorithm 2, named TWA-RND, selects the required number of wavelengths randomly from the common available wavelengths on all the requested virtual optical links.

For all requested virtual links the available wavelengths on the physical optical link are obtained. The intersection of all the available wavelengths in the requested links is defined as *labelset*. Once the available wavelengths have been obtained for the requested virtual optical network, the process FF or RND depending on the algorithm is applied. The TWA-FF selects the First-Fit wavelengths and TWA-RND selects randomly the requested wavelengths.

Selecting the First-Fit wavelengths implies to select the number of requested wavelengths starting from above the wavelength grid, while selecting randomly implies the random selection of the number of requested wavelengths from the per-request available wavelengths vector.

Algorithm 1 Transparent Wavelength Allocation First-Fit algorithm (TWA-FF)

```
procedure TWA-FF($G^P$, $G^V$,$\mathcal{W}$)
    $labelset = \bigcap_{\forall e \in G^V} w(e)$
    $\lambda = FF(labelset, \mathcal{W})$
    $a(e) = a(e) + \lambda \forall e \in G^V$
    $w(e) = w(e) - \lambda$
end procedure
```

Algorithm 2 Transparent Wavelength Allocation Random algorithm (TWA-RND)

```
procedure TWA-RND($G^P$, $G^V$,$\mathcal{W}$)
    $labelset = \bigcap_{\forall e \in G^V} w(e)$
    $\lambda = RND(labelset, \mathcal{W})$
    $a(e) = a(e) + \lambda \forall e \in G^V$
    $w(e) = w(e) - \lambda$
end procedure
```

8.3 VON RA algorithms using Wavelength Converters

In this section, the WCs are introduced to bypass the previously explained WCC. These WCs are located on each physical node, an a limited number of WCs has been considered. With the purpose to improve the efficiency on VON Resource Allocation (VON RA), the proposed algorithms (WAWC-FF and WAWC-RND) take advantage of wavelength converters availability.

Algorithm 3 details the necessary procedures for the wavelength allocation with wavelength converters for WAWC-FF algorithm. The proposed algorithm firstly tries to assign the requested wavelengths using the common available transparent wavelengths, by previously running TWA-FF(G^P, G^V, $\mathcal{W}$).

If not enough wavelengths are allocated, for each visited requested virtual optical link, the assigned wavelengths are checked against the available wavelengths in the outgoing edges of the requested link. If the assigned wavelengths in the visited requested virtual optical link are available in its outgoing edges, they are allocated, meaning that no WC is needed. In case they do not match, a WC will be allocated and the first available wavelength will be allocated in the outgoing link. For each not visited requested virtual optical link, the available wavelengths are obtained and compared against the available wavelengths of the outgoing links. A WC will be assigned in case no common wavelengths can be allocated. The destination node of the requested link will be marked as visited.

In Figure 8.1 an example of WAWC-FF algorithm is provided. A VON is requested with the virtual links (0-1,1-2,2-3,1-4,2-4,3-4) and 3 wavelengths per virtual link. In link table the available wavelengths are annotated (e.g., link 0-1 has available wavelength ids: 0, 1, 2).

The algorithm 3 allocates in link 0-1 the wavelength ids 1 and 2. A WC is required for converting available wavelength id 0 to wavelength id 3 at outgoing link 1-2. In node 1 three WC are required for converting wavelength ids 0, 1 and 2 to wavelength ids 3, 4 and 5, respectively.

Algorithm 4 follows a similar procedure than Algorithm 3, but wavelengths are assigned randomly when selecting from a set of available wavelengths.

8.4 Experimental Results

With the purpose to evaluate the performance of the different presented algorithms (i.e., TWA-FF, TWA-RND, WAWC-FF and WAWC-RND), three different test network scenarios have been set up in the GMPLS/PCE control plane platform of the ADRENALINE Testbed.

Algorithm 3 Wavelength Allocation with Wavelength Converters First-Fit algorithm (WAWC-FF)

```
procedure WAWC-FF($G^P$, $G^V$,$\mathcal{W}$)
    TWA-FF($G^P$, $G^V$,$\mathcal{W}$)
    for all $e \in E^V$ do
        if VISIT($e$) then
            $\lambda = a(e)$
        else
            $\lambda = \mathrm{FF}(w(e), \mathcal{W})$
            $a(e) = a(e) + \lambda$
            $w(e) = w(e) - \lambda$
        end if
        for all $outEdge \in out(e)$ do
            for $i = 0 \rightarrow \mathcal{W} - 1$ do
                if $\lambda \in w(outEdge)$ then
                    $a(outEdge) = a(outEdge) + \lambda$
                    $w(outEdge) = w(outEdge) - \lambda$
                else
                    $wc(e_j) = wc(e_j) - 1$
                    $\lambda' = \mathrm{FF}(w(outEdge), \mathcal{W})$
                    $a(outEdge) = a(outEdge) + \lambda'$
                    $w(outEdge) = w(outEdge) - \lambda'$
                end if
            end for
        end for
    end for
end procedure
```

Algorithm 4 Wavelength Allocation with Wavelength Converters Random algorithm (WAWC-RND)

```
procedure WAWC-RND($G^P$, $G^V$,$\mathcal{W}$)
    TWA-RND($G^P$, $G^V$,$\mathcal{W}$)
    for all $e \in E^V$ do
        if VISIT($e$) then
            $\lambda = a(e)$
        else
            $\lambda = \text{RND}(w(e), \mathcal{W})$
            $a(e) = a(e) + \lambda$
            $w(e) = w(e) - \lambda$
        end if
        for all $outEdge \in out(e)$ do
            for $i = 0 \rightarrow \mathcal{W} - 1$ do
                if $\lambda \in w(outEdge)$ then
                    $a(outEdge) = a(outEdge) + \lambda$
                    $w(outEdge) = w(outEdge) - \lambda$
                else
                    $wc(e_j) = wc(e_j) - 1$
                    $\lambda' = \text{RND}(w(outEdge), \mathcal{W})$
                    $a(outEdge) = a(outEdge) + \lambda'$
                    $w(outEdge) = w(outEdge) - \lambda'$
                end if
            end for
        end for
    end for
end procedure
```

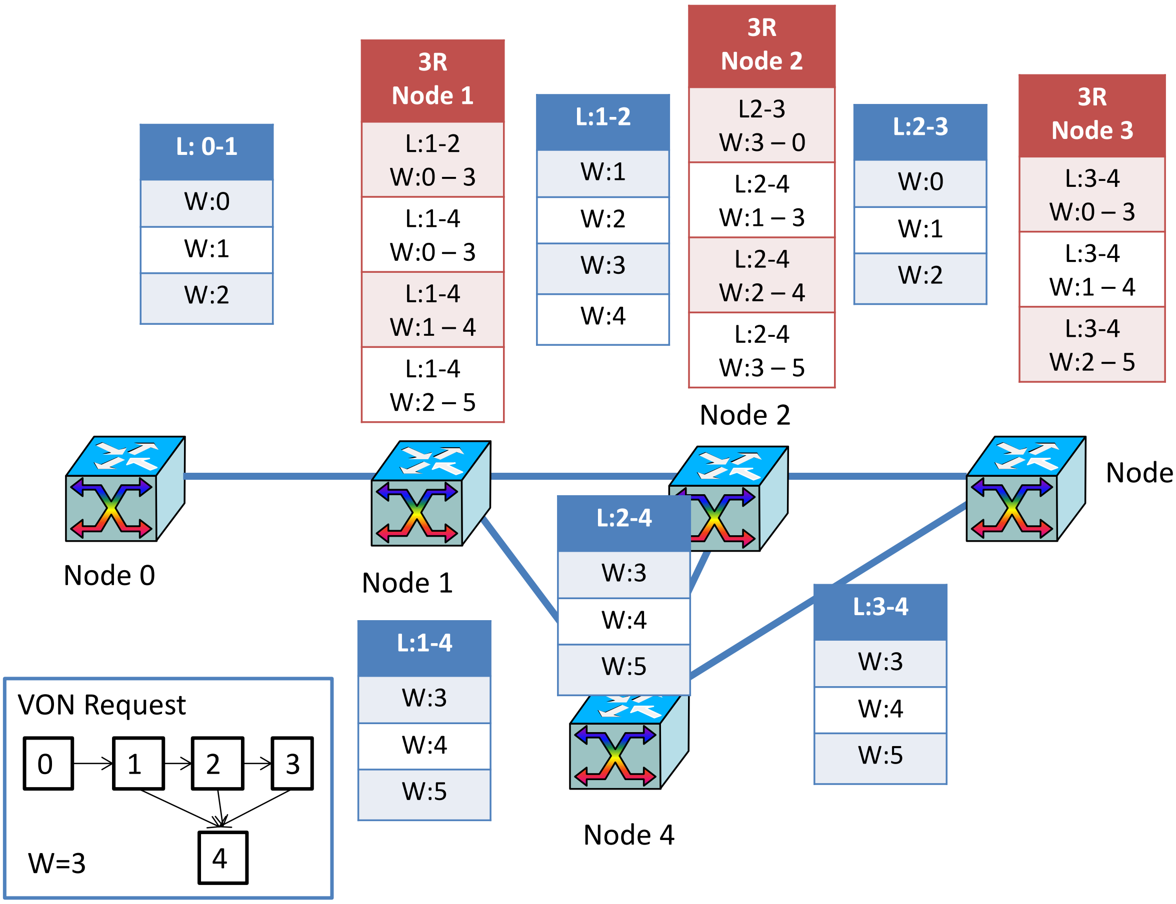


Figure 8.1: Wavelength Converter First-Fit algorithm example

The test network scenarios consist of 14 physical optical nodes connected using an European Optical Network topology (Figure 8.3), a National Science Foundation (NSF) topology (Figure 8.2) and a full-meshed topology (Figure 8.4). Each of the considered optical nodes included in the proposed scenarios contains 32 virtual GMPLS controllers and 8 WCs. For each physical optical link of the considered scenarios, 32 different wavelengths are available. The EON topology consists of 24 bidirectional links; the NSF topology consists of 21 bidirectional links and the full-meshed topology consists of 91 bidirectional links.

In order to evaluate the performance of VON setup, the performance values are obtained when dynamically provisioning VONs, in which the inter-arrival process is Poisson, and the HT follows a negative exponential distribution. The mean IAT is set to 3 seconds and the average VON-holding time is varied during all different experimentations, yielding an offered traffic load from 10 to 50 Erlangs. The topology for each VON request is randomly selected from the space of feasible topologies, which is determined by the selected evaluation scenario. The number of wavelengths requested for each VON is selected randomly between 1 and 4 wavelengths. 10^4 VONs have been requested for measuring the blocking rate, which has been computed in the 3 presented scenarios for the 4 proposed algorithms.

Figure 8.5 shows the blocking rate of VON requests in the European topology scenario. The different blocking rate results have been obtained for the four proposed algorithms (TWA-FF, TWA-RND, WAWC-FF and WAWC-RND). For a given VON request load of 20 Erlangs, the obtained VON blocking rate is 0.44%, 0.93%, 0.43% and 0.64% for TWA-FF, TWA-RND, WAWC-FF and WAWC-RND, respectively. The WAWC-FF algorithm is the most suitable algorithm in the presence of multiple

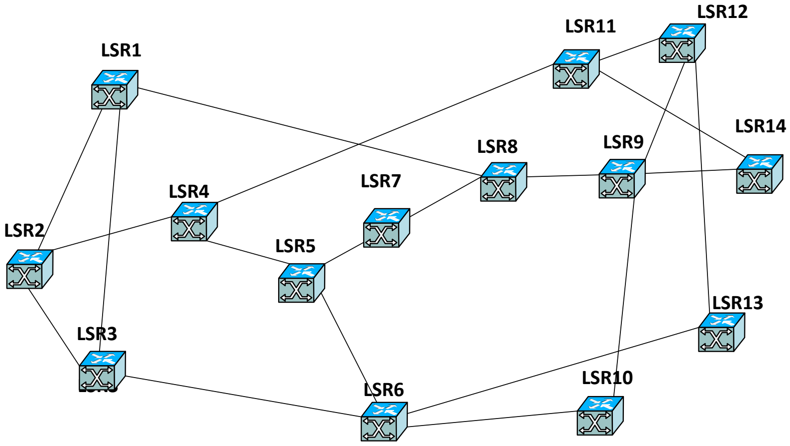


Figure 8.2: 14-node NSF topology scenario

wavelength converters on each optical node. In case of absence of wavelength converters, a better performance is given for the TWA-FF algorithm. The better performance of TWA-FF algorithms towards TWA-RND algorithm can be explained because of the bad alignment of available wavelengths on each link because of random allocation. Bigger differences in blocking rate of VON Requests can be observed for a given request load of 30 Erlangs, in which the obtained values are 3.3% and 0.43% for the TWA-FF and WAWC-FF, respectively.

The results of the proposed algorithms in the NSF network topology scenario are analyzed in Figure 8.6. It can be observed how for a VON Request load of 30 Erlangs the obtained VON blocking rate is of 8.73%, 10.08%, 2.72% and 2.95% for TWA-FF, TWA-RND, WAWC-FF and WAWC-RND algorithms, respectively. The differences between the performance of the proposed algorithms tends to be less in case of lousy coupled network topologies, so that in higher node degree network topologies, such as EON topology in comparison to NSF network topology.

Finally, in Figure 8.7 the blocking rate of VON requests is shown for the full-meshed topology scenario (Fig. 8.4). For a VON Request load of 30 Erlangs the obtained blocking rate of VON Requests is of 3.49%, 6%, 0.76% and 4.76% for the TWA-FF, TWA-RND, WAWC-FF and WAWC-RND algorithms. For a VON Request load of 50 Erlangs the results are between 12% and 18%.

For all the considered scenarios the WAWC-FF algorithm is the most suitable algorithm in case of wavelength converter availability in the physical optical nodes. In case of absence of wavelength converters the TWA-FF algorithm is preferred. These algorithms make better usage of the physical optical resources and are able to provide less blocking rate of VON requests.

8.5 Conclusions

Four different algorithms for VON Resource Allocation have been presented (i.e., TWA-FF, TWA-RND, WAWC-FF and WAWC-RND). These algorithms have been analyzed by obtaining the blocking rate of VON requests on three proposed scenarios (i.e., European, NSF, full-meshed network topologies).

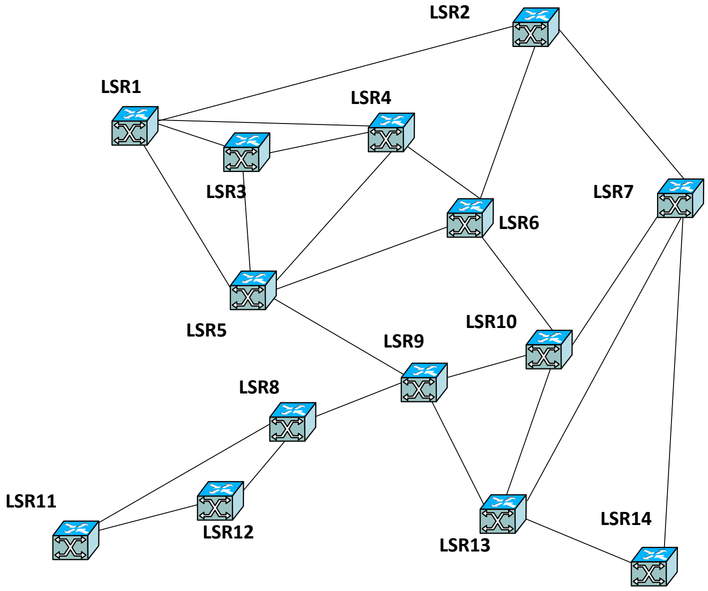


Figure 8.3: 14-node European Optical Network topology scenario

The obtained results show that in the presence of wavelength converters in the physical optical nodes, and the usage of an algorithm that takes them into account, really improves the blocking rate of VON requests.

The proposed algorithms are able to allocate virtual links for VON connected subgraphs. In next chapter a new set of algorithms is presented, allowing VON VNT.

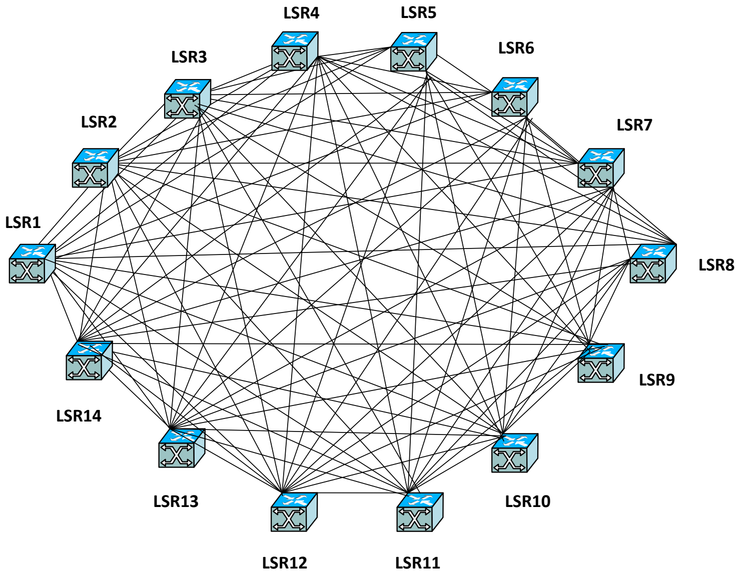


Figure 8.4: 14-node full-meshed topology scenario

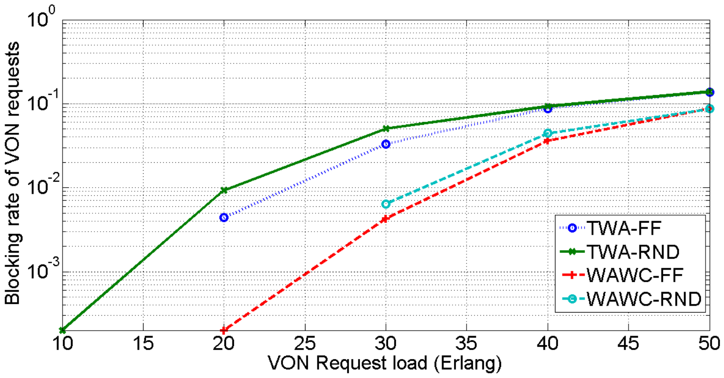


Figure 8.5: Blocking rate of VON requests in a EON topology scenario

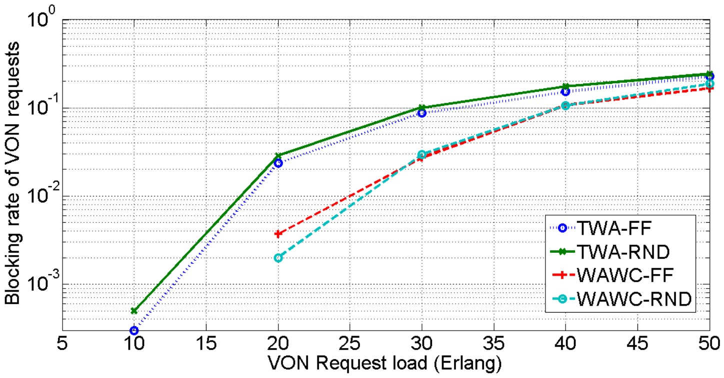


Figure 8.6: Blocking rate of VON requests in a NSF topology scenario

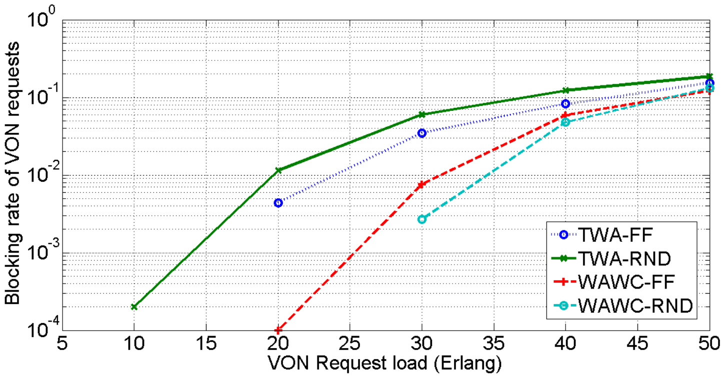


Figure 8.7: Blocking rate of VON requests in a Full-meshed topology scenario

Chapter 9

Virtual Optical Networks transport resource allocation: Virtual Network Topologies

In chapter 7, a VON resource broker architecture with a VNTM for serving the requested VONs and deploying them as virtual GMPLS-controlled WSONs instances over the same physical infrastructure [56] has been introduced.

The Global Concurrent Optimization (GCO) [59] functionality allows the PCE to perform concurrent path computations, where a set of paths are computed concurrently in order to efficiently utilize the network resources. GCO allows the concurrent path computation of all the requested VON virtual links in order to perform an efficient VON RA. The dynamic VON RA problem consists on the on-demand resource allocation for the requested VONs. The resources to be allocated are the requested virtual optical nodes and links over the physical optical infrastructure. In comparison with the previous chapter, the proposed algorithms are able to allocate VON VNT, meaning that there is no restriction on the requested VON topologies.

This chapter first introduces a formulation for VON RA and later, novel dynamic VON RA algorithms are provided for VON VNT. These algorithms take the advantage of the GCO for efficiently solving VON RA. Then, the proposed algorithms are experimentally evaluated. Finally, conclusions are provided.

9.1 VON Resource Allocation algorithms

Although the optical network resources definitions have been presented in the previous chapter, some of them are repeated in this section, because of their usage in this chapter. The physical optical infrastructure is modeled as a directed graph and denoted by $G^P = (V^P, E^P)$, where V^P is the set of physical optical nodes and E^P is the set of physical optical links.

Each optical link $e^P(i,j) \in E^P$ between two physical optical nodes i and j is associated with a certain number of available wavelengths $w(e^P)$.

The following functions are defined with the purpose of helping to obtain certain parameters of the optical infrastructure:

- $w(e)$ $\forall e \in E^P$ corresponds to the available wavelengths for edge e. w_T is the vector containing the different wavelengths supported by the physical optical infrastructure.
- FF($w(e)$, $\mathcal{W}$) is defined as the function that returns the first $\mathcal{W}$ wavelengths of $w(e)$ and triggers an error in case no wavelength is available.
- RND($w(e)$, $\mathcal{W}$) selects $\mathcal{W}$ random wavelengths from the input vector $w(e)$ with a uniformly distributed probability.

A VON request consists on a directed graph denoted by $G^V = (V^V, E^V)$ and a unsigned integer value of requested wavelengths $\mathcal{W}$. The requested set of virtual nodes is a subset of the available physical nodes, but G^V is not a subgraph of G^P ($E^V \nsubseteq E^P$). $p(e)$ is defined for each $e \in E^V$ (i.e., virtual optical link) as the ordered set of links $e^P \in E^P$ that constitute the underlying physical optical path.

In the following subsections, the proposed VON RA algorithms are introduced.

9.1.1 Shortest Path Unreserved Bandwidth First-Fit (SPUB-FF) algorithm

For each requested virtual optical link a shortest path is computed, only taking into account the unreserved bandwidth. Once the spatial virtual optical links have been computed, the common set of available wavelengths of the computed paths are obtained. The FF wavelengths of this common set are assigned.

9.1.2 Shortest Path Common Wavelength Allocation First-Fit (SPCWA-FF) algorithm

Algorithm 5 details the necessary procedures for the wavelength allocation using Shortest Path Common Wavelength Allocation First-Fit algorithm (SPCWA-FF). This algorithm makes use of a shortest path with WCC which was firstly introduced in [60] (shortest_path_wcc(e)). This shortest path algorithm uses a dijkstra algorithm that takes into account the WCC.

SPCWA-FF computes the presented shortest path with WCC for each requested virtual optical link. w_C is defined as the list of the common available wavelengths (i.e., $\textit{labelset}(p(e))$) for the computed shortest paths with WCC. SPCWA-FF assigns the first wavelengths of the obtained w_C.

Algorithm 5 SPCWA-FF algorithms

```
procedure SPCWA-FF($G^P$, $G^V$,$\mathcal{W}$)
    $p(e) = shortest\_path\_wcc(e)\ \forall e \in G^V$
    $w_C = \bigcap_{\forall e \in G^V} labelset(p(e))$
    $\lambda = FF(w_C, \mathcal{W})$
    $w(p) = w(p) - \lambda\ \forall p \in p(e)\ \forall e \in G^V$
end procedure
```

Algorithm 6 SPCWA RND algorithms

```
procedure SPCWA-RND($G^P$, $G^V$,$\mathcal{W}$)
    $p(e) = shortest\_path\_wcc(e)\ \forall e \in G^V$
    $w_C = \bigcap_{\forall e \in G^V} labelset(p(e))$
    $\lambda = RND(w_C, \mathcal{W})$
    $w(p) = w(p) - \lambda\ \forall p \in p(e)\ \forall e \in G^V$
end procedure
```

9.1.3 Shortest Path Common Wavelength Allocation Random (SPCWA-RND) algorithm

The SPCWA-RND algorithm (Alg. 6) is a variation of the previously proposed SPCWA-FF, by selecting $\mathcal{W}$ random wavelengths of w_C. These are the wavelengths that will be assigned for all the requested virtual optical links.

9.1.4 Wavelength Tree Minimum Hop Shortest Path (WTMHSP) algorithm

A 1-VON is defined as a VON in which each link has only one wavelength allocated.

Let $\mathcal{S}$ be the set of already computed 1-VONs obtained by iterating the scenario wavelengths (w_T). To compute each 1-VON, G^P is filtered with a selected wavelength $w_i \in w_T$, removing the edges which do not have the wavelength w_i available, obtaining G'. G' will be used to compute the shortest path for each requested virtual link. The obtained 1-VON is added to $\mathcal{S}$, which will include at maximum $|w_T|$1-VONs.

Algorithm 7 WTMHSP algorithm

```
procedure ASSIGNW($G^P$, $G^V$,$\mathcal{W}$)
    for all $w_i \in w_T$ do
        $G' = filter(G^P, w_i)$
        $p'(e) = shortest\_path\_wcc(e) \forall e \in G^V$
        $G'^V = G'^V \cup p'(e)$
        $\mathcal{S} = \mathcal{S} \leftarrow G'^V$
    end for
    for all $i \leftarrow 0 : (\mathcal{W} - 1)$ do
        $A = argmin(h(s)) \forall s \in \mathcal{S}$
        $G^V = G^V \cup A$
        $\mathcal{S} = \mathcal{S} - A$
    end for
end procedure
```

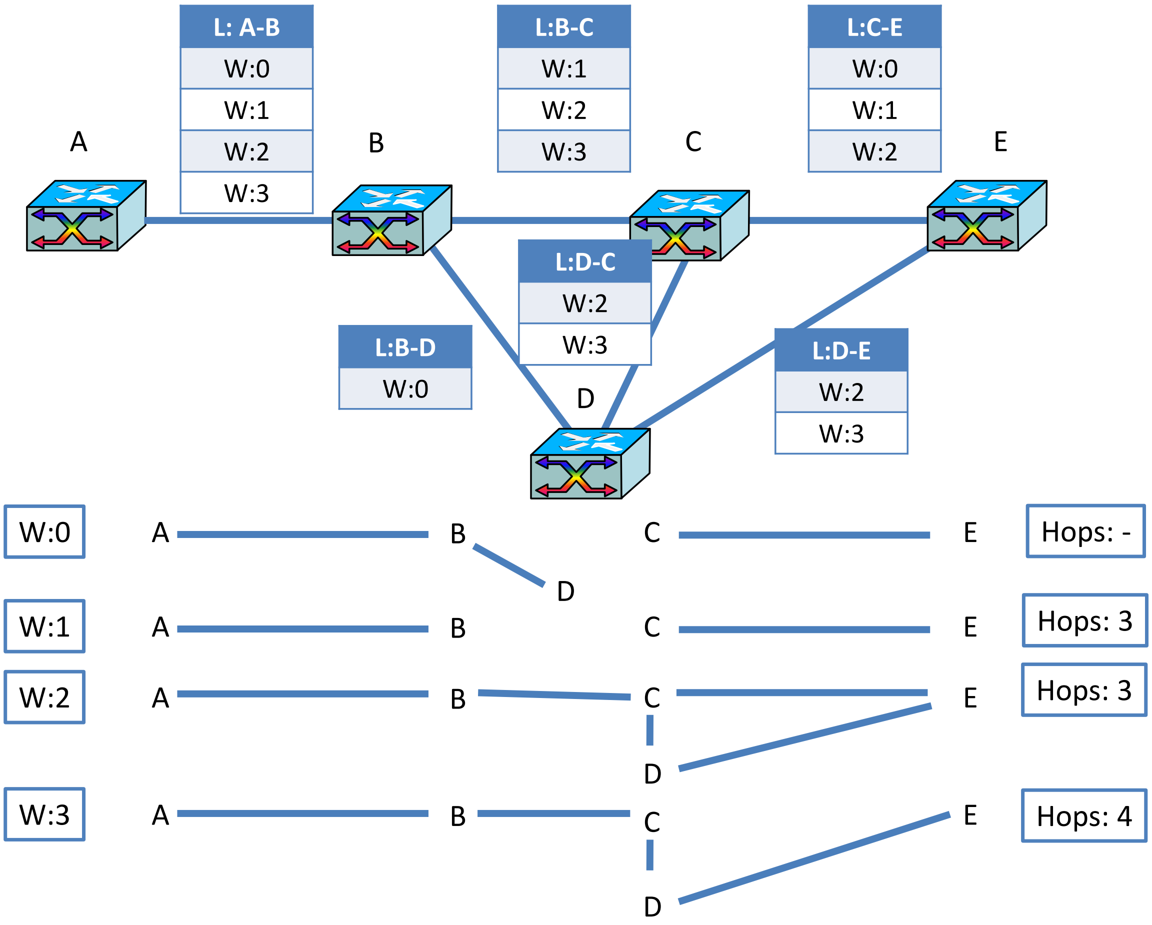


Figure 9.1: WTMHSP algorithm for virtual link request between nodes A and E

$$h(s) = \sum_{\forall e \in E^V} hops(p(e)) \tag{9.1}$$

$$G^V = G^V \cup argmin(h(s)) \forall s \in \mathcal{S} \tag{9.2}$$

Once all possible 1-VONs for $\mathcal{S}$ have been computed, they are evaluated by means of defining an heuristic function (Eq. 9.1) that evaluates the total number of hops.

Once all possible solutions are evaluated, the first $\mathcal{W}$ 1-VONs that minimize $h(s)$ (Eq. 9.2) are selected. These 1-VONs will be the ones allocated. WTMHSP performs a minimum hop count taking into account all the requested virtual optical links (i.e., GCO).

Figure 9.1 describes a simple network topology consisting of 5 nodes and 6 links. The available wavelengths at each link are noted. For example, link A-B has 3 available wavelengths: 0, 1, and 2. Next, a virtual link request is received between nodes A and E. Figure 9.1 details how the WTMHSP algorithm allocates the resources for that virtual link request. For all the possible wavelengths, the algorithm draws the possible paths. Taking into account the WCC, a shortest path is computed for each possible path. Later, the number of hops for each path is computed and the first solution with the minimum number of hops ($h(s)$) is selected. In the proposed example, no path is obtained for wavelength 0. For wavelengths 1 and 2, paths of 3 hops are obtained, while for wavelength 3 a path of 4 hops is obtained. The WTMHSP algorithm selects the first solution with the minimum number of hops, that is the path obtained with wavelength 1.

9.2 Experimental Performance Evaluation

9.2.1 System Architecture on the ADRENALINE Testbed

The proposed VON RA algorithms (i.e., SPUB-FF, SPCWA-FF, SPCWA-RND, WTMHSP) have been implemented and evaluated in the 14-node NSF network topology (Fig. 8.2) using the GMPLS/PCE control plane platform of the ADRENALINE testbed. Each optical link supports 8 different wavelengths.

The inter-arrival process is Poisson, and the HT follows a negative exponential distribution. The average IAT is set to 5s. and the average VON HT is varied for an offered traffic load ranging from 1 to 40 Er. Each VON requests 1 wavelength. 10^3 VONs have been requested for each data point. The VON request topology is generated randomly selecting the number of nodes ($V \in [2-4]$) and the number of links ($E \in [1-6]$).

To ensure that a connected graph is generated the next procedure is followed, starting with one node. Then, it iterates, creating a new node and a new link. The link is to connect the new node with a random node from the previous node set. As only bidirectional VONs make sense, also a link between the random node to the new node is added. After all nodes are created, random links are created until the VON request is fulfilled.

Figure 9.2 depicts the blocking rate of VON requests in the 14-node NSFNet topology scenario, for the different proposed VON RA algorithms, SPUB-FF, SPCWA-FF, SPCWA-RND and WTMHSP. For a given VON request load of 10 Er., the obtained blocking rate of VON requests is 19.7%, 18.3%, 34.9% and 16.8% for SPUB-FF, SPCWA-FF, SPCWA-RND and WTMHSP, respectively. Results for SPUB-FF and SPCWA-FF are similar due to the fact, that are both based on shortest path algorithm, being the first one less suitable because resource allocation algorithm does not take into account WCC, blocking the request if it is not satisfied. The WTMHSP algorithm is the most suitable. Bigger differences in the blocking rate of VON requests can be observed for a given request load of 50 Er., in which the obtained values are 41%, 40%, 52% and 31.5%, respectively. These results can be explained because the WTMHSP algorithm computes a minimal hop solution for the requested VONs, while other algorithms only take into account the shortest paths for the requested virtual optical links. In other words, computing the minimal hop solution leads to obtain a more efficient use of the network resources, reducing the blocking rate of the subsequent VON requests.

The VON setup delay is defined as the time required to allocate the resources of a virtual WSON. The VON setup delay does not include the time required for provisioning the virtual optical links. Figure 9.3 shows that the WTMHSP algorithm requires 69 ms to allocate a virtual WSON, while SPUB-FF algorithm requires 63 ms. This can be explained by the number of operations performed by each of the algorithms. As WTMHSP performs more operation in order to find optimal VON RA it is understandable that the VON setup delay is higher, although the improvement in performance justifies the higher delays.

9.3 Conclusions

Several VON RA algorithms (i.e., SPUB-FF, SPCWA-FF, SPCWA-RND and WTMHSP) have been proposed and compared in terms of VON blocking rate and VON setup delay. These algorithms have

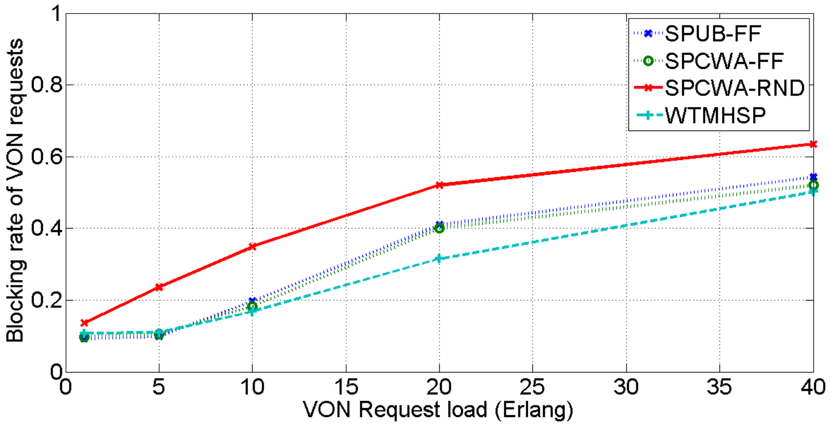


Figure 9.2: VON request Blocking rate

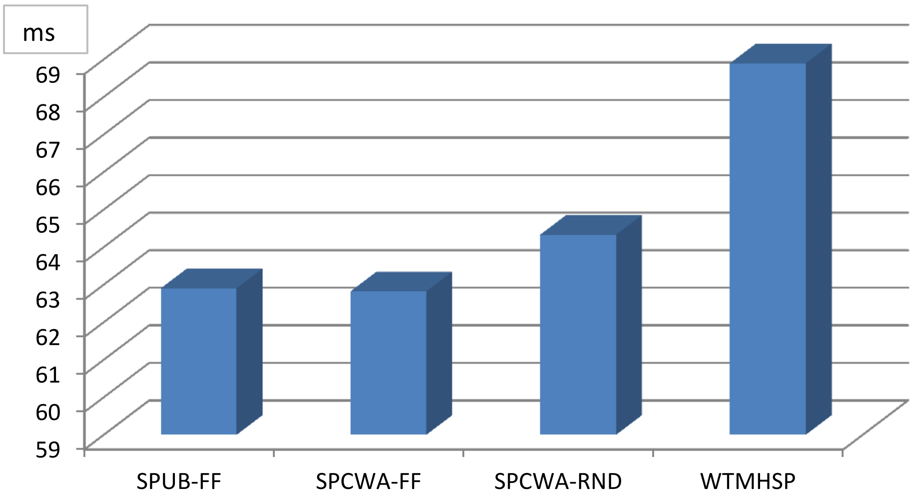


Figure 9.3: VON Setup Delay

been experimentally evaluated in a system architecture including the proposed VON resource broker with VNTM.

The obtained results show how the proposed WTMHSP algorithm does lower the blocking rate of VON requests in the analyzed 14-node NSFNet scenario, while not increasing the complexity of the VON resource broker, minimizing the total number of hops for the requested virtual optical links. The SPUB-FF algorithm has demonstrated to be a good alternative in case that the VON setup delay is a critical value.

Chapter 10

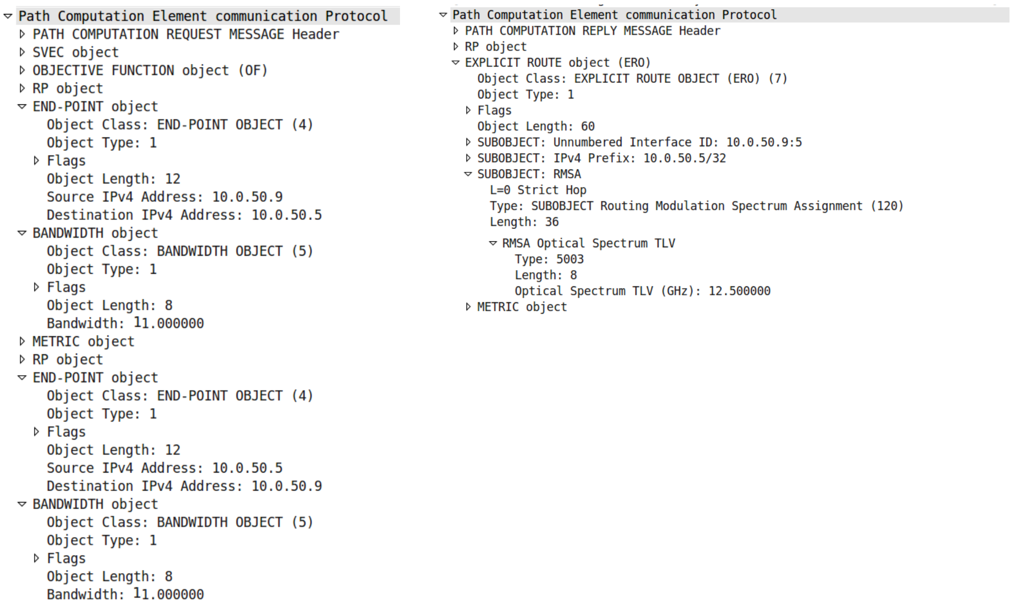
Virtual Optical Networks transport resource allocation: Elastic Optical Networks

In section 1.4, EON have been introduced. A GMPLS/PCE control plane for EONs has been discussed in [61], where required extensions and adaptations of enhancements to GMPLS and protocols and functions are proposed, in order to deploy a control plane for elastic and flexi-grid DWDM optical networks using the CO-OFDM transmission technology. Such control plane relies on Frequency Slot (FS) allocation by means of a PCE.

The objective of this chapter is to propose extensions to the VON resource broker presented in chapter 7, present a Virtual Elastic Optical Network (VEON) resource allocation algorithm and evaluate its performance in the GMPLS/PCE control plane platform of the ADRENALINE Testbed. The proposed extensions allow the resource broker to dynamically deploy GMPLS-controlled VEON whose virtual optical links are derived from established elastic optical connections.

10.1 System Architecture

A virtualizable GMPLS/PCE-controlled elastic optical network is managed by the VON resource broker with VNTM, described in chapter 7. This section describes the main extentions that have been designed and implemented in the VON resource broker, in order to dynamically deploy virtual GMPLS-controlled EONs as a service.



```
▽ Path Computation Element communication Protocol
  ▷ PATH COMPUTATION REQUEST MESSAGE Header
  ▷ SVEC object
  ▷ OBJECTIVE FUNCTION object (OF)
  ▷ RP object
  ▽ END-POINT object
      Object Class: END-POINT OBJECT (4)
      Object Type: 1
    ▷ Flags
      Object Length: 12
      Source IPv4 Address: 10.0.50.9
      Destination IPv4 Address: 10.0.50.5
  ▽ BANDWIDTH object
      Object Class: BANDWIDTH OBJECT (5)
      Object Type: 1
    ▷ Flags
      Object Length: 8
      Bandwidth: 11.000000
  ▷ METRIC object
  ▷ RP object
  ▽ END-POINT object
      Object Class: END-POINT OBJECT (4)
      Object Type: 1
    ▷ Flags
      Object Length: 12
      Source IPv4 Address: 10.0.50.5
      Destination IPv4 Address: 10.0.50.9
  ▽ BANDWIDTH object
      Object Class: BANDWIDTH OBJECT (5)
      Object Type: 1
    ▷ Flags
      Object Length: 8
      Bandwidth: 11.000000
```

```
▽ Path Computation Element communication Protocol
  ▷ PATH COMPUTATION REPLY MESSAGE Header
  ▷ RP object
  ▽ EXPLICIT ROUTE object (ERO)
      Object Class: EXPLICIT ROUTE OBJECT (ERO) (7)
      Object Type: 1
    ▷ Flags
      Object Length: 60
    ▷ SUBOBJECT: Unnumbered Interface ID: 10.0.50.9:5
    ▷ SUBOBJECT: IPv4 Prefix: 10.0.50.5/32
    ▽ SUBOBJECT: RMSA
        L=0 Strict Hop
        Type: SUBOBJECT Routing Modulation Spectrum Assignment (120)
        Length: 36

      ▽ RMSA Optical Spectrum TLV
          Type: 5003
          Length: 8
          Optical Spectrum TLV (GHz): 12.500000
    ▷ METRIC object
```

Figure 10.1: *PC_Request* message and *PC_Reply* messages

The VON resource broker is the responsible for serving the incoming, asynchronous and dynamic VEON requests, encompassing the provisioning of a new VEON, the allocation and/or release of the resources (both at transport and control plane level) assigned to an existing VEON.

A VEON request is modelled as a graph that describes a set of virtual optical switches (BV-OXC and BV-ROADM) and elastic links for the virtual transport plane specifying for each one the number of requested input and output optical ports, and the requested slot width (in Hz), including the number of 6.25 GHz slots for each virtual link of the requested VEON. The VON resource broker has been modified to distinguish between a request of different wavelengths or a request for a number of slots.

The PCC has been adapted to include a bandwidth object (Fig. 10.1), which includes the requested slot width in Hz. This feature allows to request the virtual elastic optical links to the PCE.

Finally, the XML message protocol between LSP Manager and RSVP-TE running at GMPLS controllers has been extended to include the requested Frequency Slot (FS), with the usage of a pre-standard label, which includes the first slot and the number of allocated slots.

10.2 Virtual Elastic Optical Network resource allocation

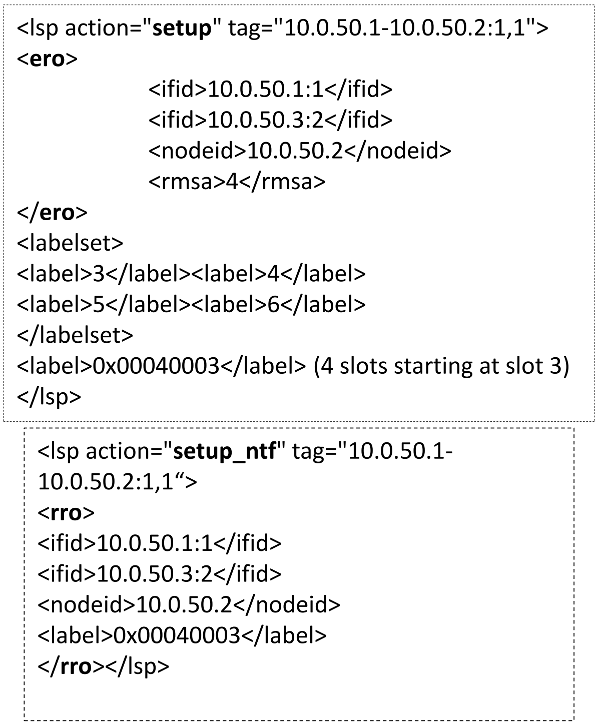
This section details a proposed VEON resource allocation algorithm (Alg. 8). The Shortest Path Spectrum Allocation First-Fit algorithm (SPSA-FF) has been designed and implemented for the VON resource broker with VNTM and runs in the PCE.

A VEON request consists on a directed graph denoted by $G^V = (V^V, E^V)$ and a unsigned integer value of requested slots $\mathcal{SW}$, meaning the spectrum bandwidth requested divided per the slot grid (i.e., 6.25 GHz).

A disjoint shortest path is computed for each requested virtual link, meaning that once a path has been computed for a virtual link, it cannot be used for another requested virtual optical link. By this means, the Spectrum Continuity Constraint (SCC) is satisfied. Finally, the first common available set of bandwidth slots are assigned.

Algorithm 8 Shortest Path Spectrum Allocation First-Fit algorithm (SPSA-FF)

```
procedure SPSA-FF($G^P$, $G^V$, $\mathcal{SW}$)
    $p(e) = disjoint\_shortest\_path(e)\ \forall e \in G^V$
    $\lambda = \bigcap_{\forall p \in p(e) \forall e \in G^V} w(p(e))$
    $\lambda' = FF(\lambda, \mathcal{SW})$
    $w(p(e)) = w(p(e)) - \lambda'\ \forall p \in p(e)\ \forall e \in G^V$
    $a(p(e)) = a(p(e)) + \lambda'\ \forall p \in p(e)\ \forall e \in G^V$
end procedure
```



```
<lsp action="setup" tag="10.0.50.1-10.0.50.2:1,1">
<ero>
                <ifid>10.0.50.1:1</ifid>
                <ifid>10.0.50.3:2</ifid>
                <nodeid>10.0.50.2</nodeid>
                <rmsa>4</rmsa>
</ero>
<labelset>
<label>3</label><label>4</label>
<label>5</label><label>6</label>
</labelset>
<label>0x00040003</label> (4 slots starting at slot 3)
</lsp>
```

```
<lsp action="setup_ntf" tag="10.0.50.1-
10.0.50.2:1,1“>
<rro>
<ifid>10.0.50.1:1</ifid>
<ifid>10.0.50.3:2</ifid>
<nodeid>10.0.50.2</nodeid>
<label>0x00040003</label>
</rro></lsp>
```

Figure 10.2: RSVP XML interface

10.3 Experimental Assessment and Performance Evaluation

10.3.1 Experimental assessment

The message exchange for VON request and provisioning has been previously described in chapter 7. In this section, protocol details are described for the PCE request and reply for VEON resource assignment, the protocol using an XML encoding between the VNTM and the ingress node and, finally, LSP establishment (RSVP-TE PATH/RESV messages exchange). As an example a VEON request is created consisting of 2 virtual optical links between nodes 1 and 2 (i.e., 1-2 and 2-1) . With the purpose of obtaining a path for the requested virtual optical links, the resource broker issues to the PCE a *PC_Request* message including an SVEC object where the different requested links are enumerated (Fig. 10.1).

In this example, two paths are requested including respective source and destination nodes. As an example, the PCE uses the SPSA-FF algorithm (Alg. 8), where the FF necessary slots of the common

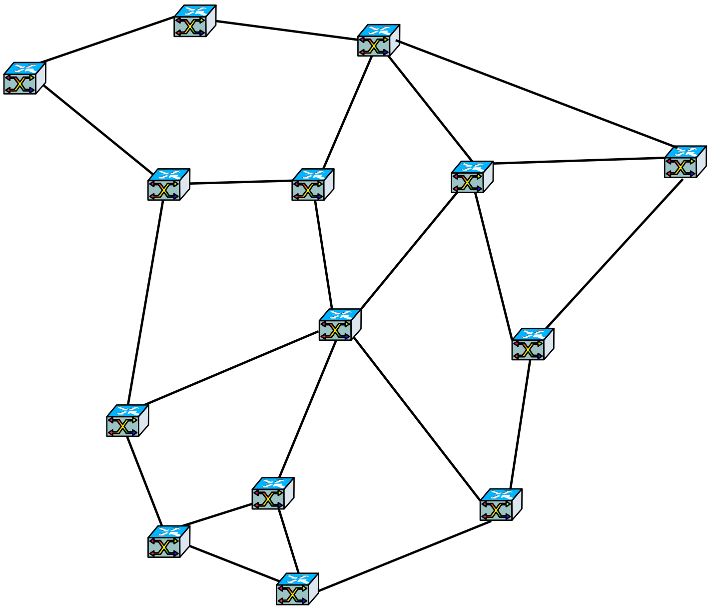


Figure 10.3: Spanish EON topology)

set of available slots of the computed paths are assigned. The PCE replies with a *PC_Reply* message which includes the EROs (spatial path) and the generalized label for each virtual optical link, which encodes the frequency slot to use.

The *PC_Reply* message is processed and, through the LSP manager, an LSP is requested through an XML proprietary interface to the RSVP-TE process of the source node (Fig. 10.2). Once a setup LSP request message is received at source node 1, a RSVP-TE PATH message is issued to node 2, following RSVP-TE standard procedure. Node 2 responds with RSVP-TE RESV message which includes the RRO, and when the message is received at node 1, the optical resources for the LSP have been occupied. Finally, node 1 sends to the RB an acknowledgement for the setup of the LSP. This is performed for all the requested virtual optical links.

Once all the necessary LSPs have been established, the necessary virtual GMPLS control plane resources need to be allocated. To this end, the established LSPs are used as virtual TE links. Finally, once the VEON configuration is generated, the VEON configurator, by means of ADNETCONF, is the responsible to set up or tear down the requested VEON.

10.3.2 Experimental performance evaluation

The experimental assessment consists on validating the VON resource broker architecture by requesting different VEONs to be deployed on the 14-node Spanish topology (Fig. 10.3) in the GMPLS/PCE control plane platform of the ADRENALINE Testbed. Each link has 128 frequency slots of 6.25 GHz. Frequency slots are allocated assigning one or more slots. The arrival process is Poisson, and the HT follows a negative exponential distribution. The average IAT is set to 10s and the average VEON HT is varied for an offered traffic load ranging from 1 to 50 Er.

Each VEON request ranges from 1 to 16 slots (i.e., 6.25 Ghz to 100 GHz) being uniformly distributed. 10^3 VEONs have been requested for each data point. The VEON request topology is generated randomly selecting the number of nodes (V $\in$ [2-4]) and the number of links (E $\in$ [1-6]).

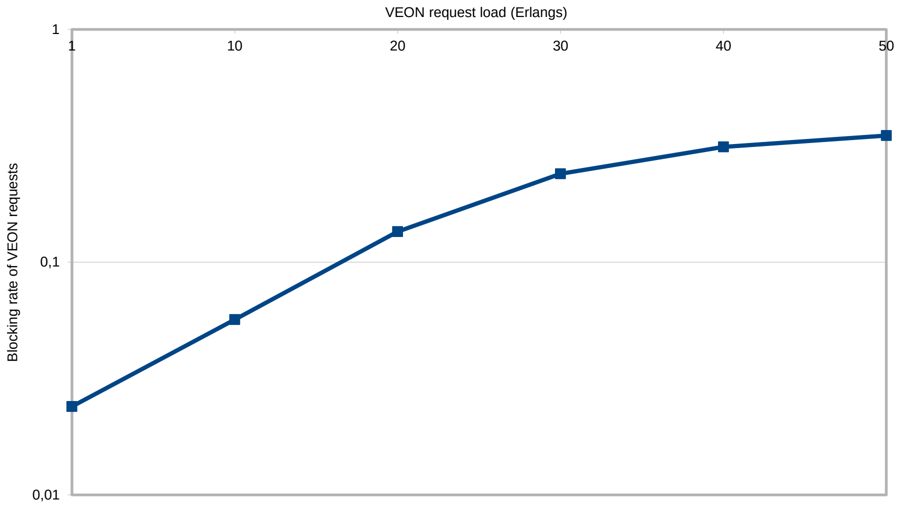


Figure 10.4: Blocking rate of VEON requests

Figure 10.4 depicts the blocking rate of VEON requests in the Spanish topology scenario, for the proposed algorithm. For a given VEON request load of 10 Er., the obtained blocking rate of VEON requests is 5.7%.

10.4 Conclusions

The previously proposed resource broker has been expanded to deploy virtual GMPLS-controlled EON topologies by establishing elastic optical connections over a GMPLS-controlled physical optical network infrastructure and offering them as virtual TE links. The SPSA-FF VEON resource allocation algorithm has been presented and experimentally evaluated in the GMPLS/PCE control plane platform of the ADRENALINE Testbed. The modified VON resource broker has shown the feasibility of deploying independent instances of virtual GMPLS-controlled EON.

Chapter 11

Virtual Optical Networks control resource allocation

In chapter 6 is remarked that a VON must be composed of not only a virtual transport plane but also of a virtual control plane, with the purpose of providing the required independent and full control functionalities (i.e., optical connection provisioning, traffic engineering, protection/restoration, etc.).

This chapter presents and evaluates from a GMPLS control plane perspective the VON resource broker for dynamic GMPLS-controlled WSON infrastructure services presented in chapter 7 and evaluates the effects of the introduction of the virtual control resources to the deployed VONs.

11.1 Virtualization of the GMPLS control plane for Virtual Optical Networks

A virtual GMPLS control plane is a distributed entity composed of virtual GMPLS controllers (one per virtual optical switch) executing several collaborative processes and a Data Communication Network (DCN) based on virtual IP Control Channels (IPCC) to allow the exchange of control messages between the virtual GMPLS controllers (see Fig. 11.1). The processes running in a virtual GMPLS controller are RSVP-TE signalling agent for connection provisioning, OSPF-TE routing agent for topology and resource dissemination, the LRM agent for local resource management and HAL agent for managing the Connection Controller Interface (CCI).

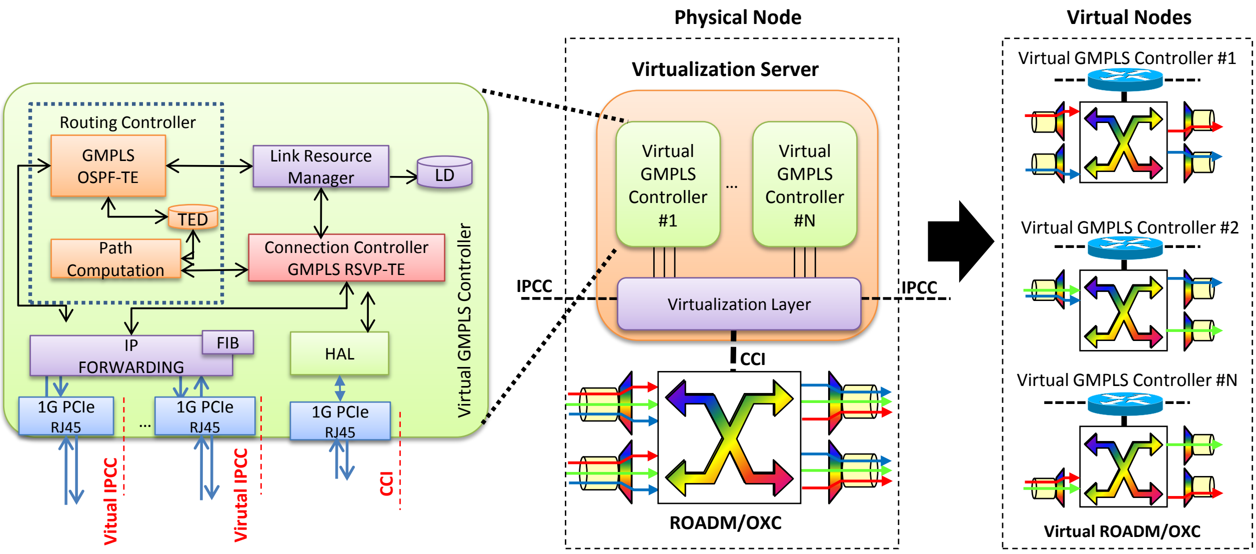


Figure 11.1: Partitioning of physical transport and control resources

The process of virtualizing a GMPLS controller has been explained in detail in chapter 6. Each virtualization server hosts a configurable number of virtual guests, each replicating a virtual GMPLS controller.

The VON resource broker is also responsible for assigning the allocated resources and deploying a dedicated DCN as an overlay on top of the shared physical infrastructure via KVM bridging modules, which allow connecting the physical NIC that provides connectivity to a neighboring node to each of the virtualized network interfaces, resulting in isolated IPCCs.

The presented virtual control plane resource allocation algorithm 9 describes the followed procedure to obtain the necessary configuration parameters once transport plane resources have been allocated for a virtual optical network request.

The inputs to this algorithm are a graph (G^V) composed of the transport plane allocated nodes and edges (links) and a virtual optical network counter, which is increased for each allocated virtual optical network request.

Firstly, the algorithm configures each requested virtual link, including configuration for the DCN, as well as the required data plane information, such as the allocated wavelength (Alg. 10). The node information is updated to configure the different GMPLS processes.Secondly, the virtual node configuration is created (Alg. 11). Finally, ADNETCONF [54] general properties are included.

11.2 Experimental performance evaluation

The experimental evaluation consists on the evaluation of the impact of the allocation of virtual control resources (i.e., virtual GMPLS controllers) when the VON resource broker is requested to deploy different VONs on the ADRENALINE Testbed.

The experimental evaluation has been performed on the virtualizable GMPLS-controlled WSON platform of the ADRENALINE Testbed, detailed in Section 6.5.

The virtual machines have the same amount of RAM (which is the RAM of the virtualization server divided by the number of active virtual machines) and share the CPU power of the server. The

Algorithm 9 Virtual optical network control plane configuration algorithm

```
procedure GenerateConfig($G^V$, $counter$)
    string conf
    for all $e \in E^V$ do
        conf += edgeConfig($e$,$counter$,$V^V$)
    end for
    for all $v \in V^V$ do
        conf += nodeConfig($v$,$counter$)
    end for
    conf += adnetconfProperties
end procedure
```

Algorithm 10 Edge configuration algorithm

```
procedure edgeConfig($e$,$counter$,$V^V$)
    string conf
    $conf + = e \rightarrow Peer0$
    $conf + = e \rightarrow Peer1$
    $conf + = e \rightarrow EthDevPeer0$
    $conf + = e \rightarrow EthDevPeer1$
    $conf + = IpPeer0(counter, e \rightarrow index, e_i)$
    $conf + = IpPeer1(counter, e \rightarrow index, e_j)$
    $conf + = IpGrePeer0(counter, e \rightarrow index,$
    $e_i)$
    $conf + = IpGrePeer1(counter, e \rightarrow index, e_j)$
    $conf + = UserDefinedEdgeConfig$
    $llid = EthDevPeer0$
    $remote\_lid = EthDevPeer1$
    for all $v \in V^V$ do
        if $v = e_i$ then
            for all $lambda \in e \rightarrow assigned\_lambdas$ do
                $v \rightarrow add\_lambda(lambda, llid,$
                $Peer0, Peer1)$
            end for
            $v \rightarrow add\_network\_entry(llid, remote\_lid,$
            $Peer0, Peer1, EthDevPeer0, EthDevPeer1,$
            $IpGrePeer0, IpGrePeer1)$
        end if
        if $v = e_j$ then
            for all $lambda \in e \rightarrow assigned\_lambdas$ do
                $v \rightarrow add\_lambda(lambda, remote\_lid,$
                $Peer1, Peer0)$
            end for
            $v \rightarrow add\_network\_entry(remote\_lid,$
            $llid, Peer1, Peer0, EthDevPeer1,$
            $EthDevPeer0, IpGrePeer1, IpGrePeer0)$
        end if
    end for
end procedure
```

Algorithm 11 Node configuration algorithm

```
procedure NODECONFIG($v$,$counter$)
    string config
    $config+ = nodeBasics(v, counter)$
    $config+ = nodeBins(v \rightarrow nodeName)$
    $config+ = nodeOlrm(v \rightarrow nodeName)$
    $config+ = nodeLrm(v)$
    $config+ = nodeOspf(v, counter)$
    $config+ = nodeRsvp(v \rightarrow nodeName)$
    $config+ = nodeSnmp(v \rightarrow nodeName)$
end procedure
```

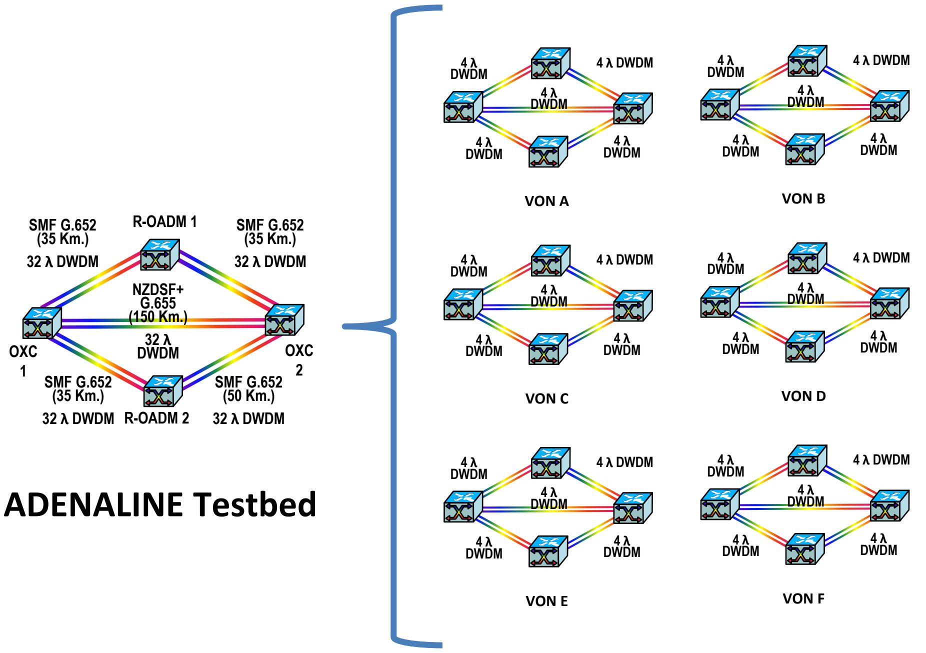


Figure 11.2: ADRENALINE Testbed WSON scenario and multiple VONs

virtualization of the GMPLS control plane of the ADRENALINE Testbed to partition the physical control plane into multiple virtual control plane instances has been previously addressed in Section 6.2.

The performed evaluation has only taken into account the performance measurements for the control plane, so that there has been no deployment of the configurations for the transport plane. 32 wavelengths per each optical link have been considered.

11.2.1 Impact of the available virtual control resources on the VON deployment

In order to evaluate the performance of VON provisioning, the performance values are obtained when dynamically provisioning VONs, in which the inter-arrival process is Poisson, and the HT follows a negative exponential distribution. The mean IAT is set to 3 seconds and the average VON-holding time is varied during all different experimentations, yielding an offered VON Request load from 0.1 to 50

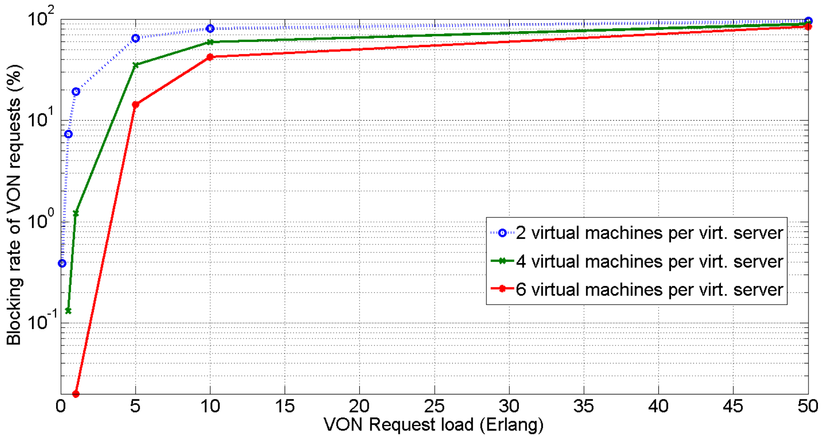


Figure 11.3: Blocking rate of VON requests

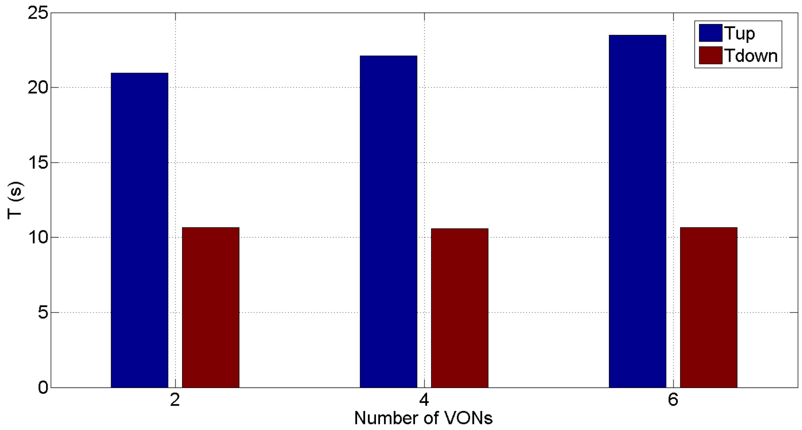


Figure 11.4: Setup and tear down mean times for VON in the ADRENALINE Testbed

Erlangs. The topology for each VON request is randomly selected from the space of feasible topologies, which is determined by the physical optical network infrastructure. For the ADRENALINE Testbed, a set of 28 topologies is considered, due to the fact that the VON Resource Broker for connected subgraphs has been used.

Three different scenarios are proposed for the evaluation of the VON resource broker performance when deploying VONs. In the first scenario, each virtualization server has 2 virtual machines running on top of it, whilst in the second scenario 4 virtual machines are running on top of the virtualization server and in the third scenario, 6 virtual machines.

The number of wavelengths requested for each VON is set to 4, and Algorithm 1 TWA-FF is used. 10^4 VONs have been requested for measuring the blocking rate, which has been computed in the 3 presented scenarios.

In Figure 11.3, the blocking rate of VON requests is lower when more virtual machines per virtualization server are used. For a given VON request load of 5 Erlangs, the obtained blocking rate of VON requests is 14.3% in the case of using 6 virtual machines and 64.9% in the case of using only 2 virtual machines. These results are coherent with the expectation of lower blocking probability in case of providing more virtual machines, which act as GMPLS controllers.

The impact of the virtual control resources is mesured on the setup and tear down delay when deploying the VON requests, only measuring the setup and tear down delay of the control plane, depending on the number of virtual machines in the virtualization servers. 100 requests have been performed for each data point. Figure 11.4 shows how the necessary time for setting up a VON is increased linearly when the number of virtual machines in the virtualization servers increases (e.g. setup time is 20.97 seconds with 2 VMs, whilst it is 22.10 seconds for 4 VMs). Instead, the tear down time does not increase with regard the number of virtual machines. These results can be explained because for setting up a VON and its control plane a higher amount of CPU and RAM resources need to be used, whilst for tearing down the VON this amount is not such significant. The time for setting up a VON increases with the number of VMs in use, due to the fact that other virtual GMPLS controllers are competing for the same CPU and RAM resources.

11.2.2 Impact of allocated virtual control resources on the VON performance

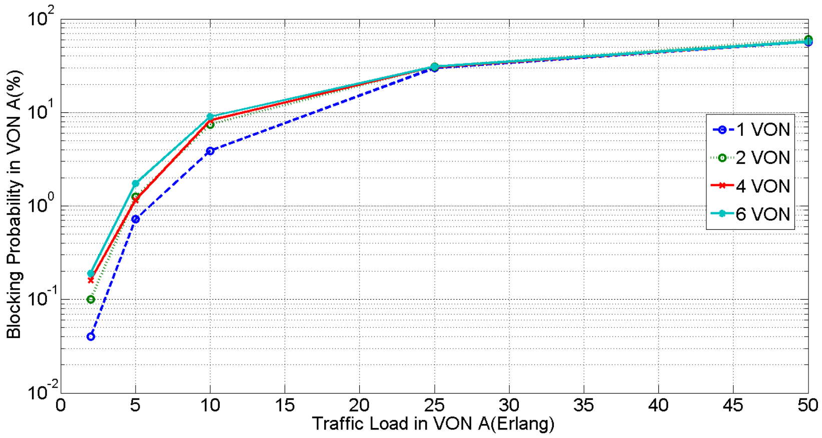


Figure 11.5: Blocking probability in VON A

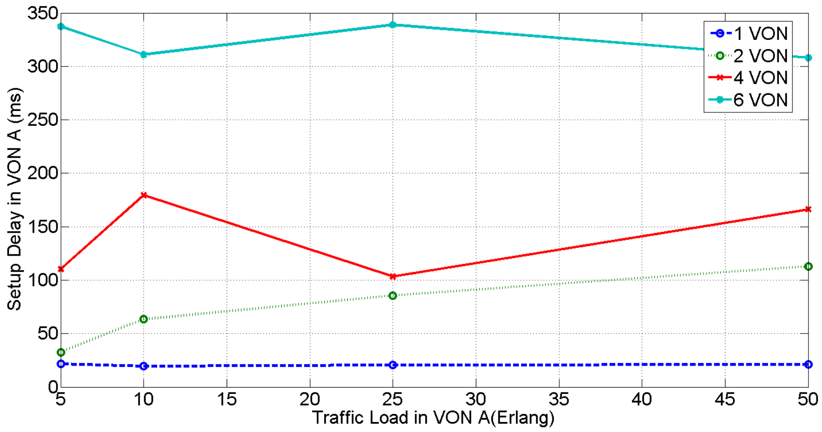


Figure 11.6: Setup Delay for connection requests in VON A

In order to evaluate the impact of the allocated virtual resources on virtual control plane performance four scenarios have been considered, depending on the number of VON deployed on the ADRENALINE Testbed. Figure 11.2 depicts the scenario where 6 VONs have been deployed on the

ADRENALINE Testbed. Scenarios for 4, 2 and 1 VON deployed have also been considered. Each VON has been assigned with 4 lightpaths for each of its virtual optical links.

The performance values are obtained when dynamically provisioning connections inside each VON, the lightpath request inter-arrival process is Poisson, and the HT follows a negative exponential distribution. The mean IAT is set to 3 seconds and the average holding time is varied during all different experimentations, yielding an offered traffic load from 2 to 50 Erlangs inside each VON. 10^4 connections have been requested inside each VON for the four scenarios.

In Figure 11.5, the connection blocking probability inside the VON A (for the other active VONs the results are equivalent) is significantly affected depending on the number of deployed VON virtual control planes, which depends on the number of VM in the virtualization servers upon each optical node. For example the Blocking Probability (BP) in VON A is 0.72% when the traffic load is of 5 Erlangs and only VON A is active, whilst the BP in VON A for the same traffic load is 1.74% when 6 VONs are deployed on the ADRENALINE Testbed. The obtained results can be explained because of the increase of the Setup Delay (SD), implies a longer lightpath establishment time, so there are more concurrent requests being processed at the time.

Figure 11.6 shows the mean setup delay in VON A and how it is affected depending the number of deployed VON virtual control planes. The setup delay in VON A is one order of magnitude higher (337 ms) in the case of 6 VON control planes than in the case of 2 VON control planes (32.61 ms) for a traffic load in both VONs of 5 Erlangs. This poor performance is explained by the virtualization of the GMPLS controllers and the amount of GMPLS controllers running on top of a virtualization server. The obtained results of SD in the case of 6 VON virtual control planes (337 ms / 5 Erlangs) might be suitable for lightpath provisioning, but they may not be tolerable in case of connection restoration.

11.3 Conclusions

In this chapter, it has been analyzed the impact of deploying several virtual VON GMPLS control planes on the performance of a single allocated virtual control plane (i.e., setup delay and blocking probability) and it has resulted in higher setup delay and higher blocking probability in lightpath provisioning. The obtained results show the trade off between the diminution of blocking rate of VON requests, through the provision of more virtual resources for the virtual control planes, and the deterioration of the virtual control plane performance.

In case that a maximum setup delay of 150 ms and a blocking probability lower than 2% are desired when requesting lightpaths to a VON with 5 Erlangs traffic load, the maximum number of virtual machines to be deployed in the presented virtualization server are 4 VMs, which will result in a blocking rate of 1.21% for a VON request rate of 1 Erlang in the ADRENALINE Testbed.

Part III

Virtual GMPLS-controlled MPLS-TP Networks over a shared WSON

Chapter 12

Design and development of a MPLS-TP node

This chapter has two objectives. The first objective is to present the architectural design, implementation and the packet forwarding performance evaluation of a MPLS-TP/PWE3 node with software-based forwarding and integrated 10Gbps tunable DWDM transponders to extend the GMPLS-enabled WSON transport network of the ADRENALINE Testbed, as shown in Fig 12.1. The second objective is the design and implementation of an architecture for the virtualization of the MPLS-TP/PWE3 node.

This work is organized as follows: section 12.1 describes the architecture of the proposed GMPLS-controlled MPLS-TP/PWE3 node. Section 12.2 provides the software implementation details of the MPLS-TP/PWE3 label forwarding engine. Section 12.3 discusses the implementation of a 10Gbps tunable DWDM transponder. Later, the experimental performance evaluation of the implemented forwarding engine and the used 10Gbps tunable DWDM transponder is carried out in section 12.4. Finally, the virtualization of the MPLS-TP/PWE3 node is provided in section 12.5.

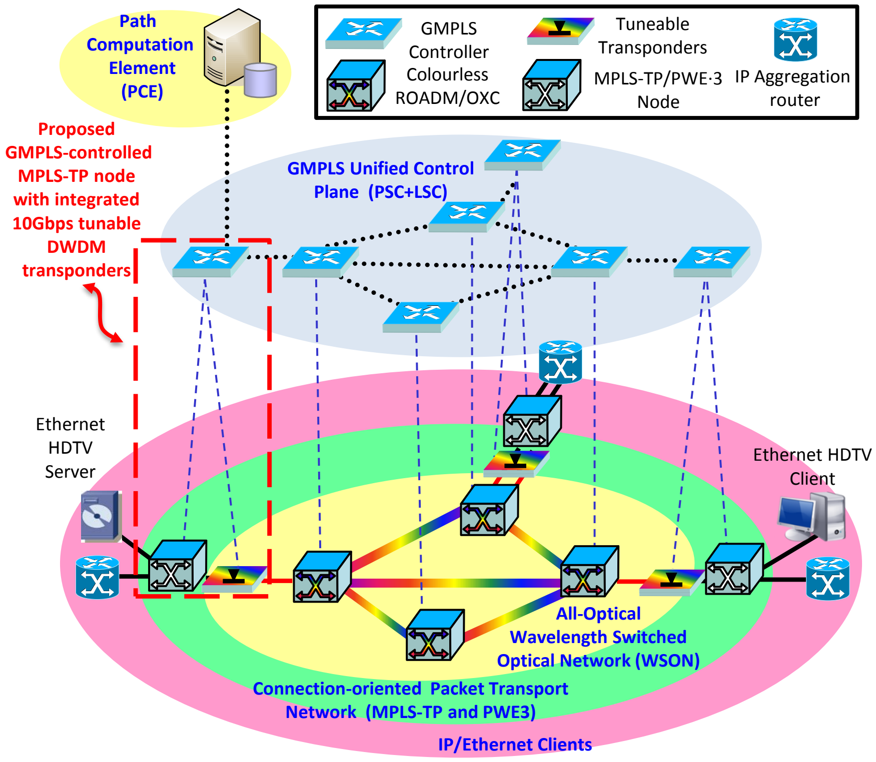


Figure 12.1: Architecture of enhanced single-domain dual-region (MPLS-TP and WSON) ADRENALINE Testbed

12.1 GMPLS-controlled MPLS-TP/PWE3 node with integrated 10Gbps tuneable DWDM transponders

Figure 12.2 shows the logical architecture of the proposed GMPLS-controlled MPLS-TP/PWE3 node with integrated 10Gbps tuneable DWDM transponders. In general, the node architecture is divided into two main elements, the control and the data planes.

]

12.1.1 Functionalities

The control plane is responsible for handling the establishment, maintenance, deletion, and protection of MPLS-TP connections in a dynamic and on-the-fly fashion, by using the RSVP-TE connection controller. Meanwhile, it should also take care the discovering and disseminating the topology and available resources in MPLS-TP and WSON networks, and store the results in the TED, by using the OSPF-TE routing controller. In the peer model a unified control plane allows the node's TED to contain the information relative to both WSON and MPLS-TP layers (i.e., lambda switched capable - LSC TE links and packet switched capable - PSC TE links, respectively) constituting the dual-layer network. Consequently, a path across multiple regions can be computed. To this end, GMPLS defines the concept and notion of Forwarding Adjacency (FA), which favors the use of efficient multi-layer routing algorithms exploiting the traffic aggregation and grooming strategies [62]. Indeed, the purpose of the traffic grooming aims at merging/grouping low-speed and flexible packet connections with small bandwidth requirements into high-speed optical tunnels with coarse bandwidth, reusing already established lightpaths (i.e., optical 10GE interfaces). That is, assuming a local control policy, a GMPLS node may advertise an LSC connection (LSP in the context of the GMPLS) as a PSC TE link. Such a link is referred to as FA. The corresponding LSP is thus referred to as a FA LSP.

The TE attributes (e.g., end-point LSRs, bandwidth, etc.) of a particular FA are disseminated by the routing protocol instance(i.e., OSPF). This allows other MPLS-TP nodes to use the created FA as a regular PSC TE link for path computation purposes. Figure 2.3 has shown an example of two LSC FA LSPs that are used as PSC TE links, and a PSC LSP that make use of the created PSC TE links. Finally, if an Ethernet service is requested, it is necessary to set up the PW tunnel through the just created PSC LSP, according to the PWE3 architecture. To do this, the source and destination MPLS-TP nodes exchange the PW labels using the T-LDP protocol. It is worth mentioning that the control plane is also responsible for creating and managing the forwarding tables of the data plane.

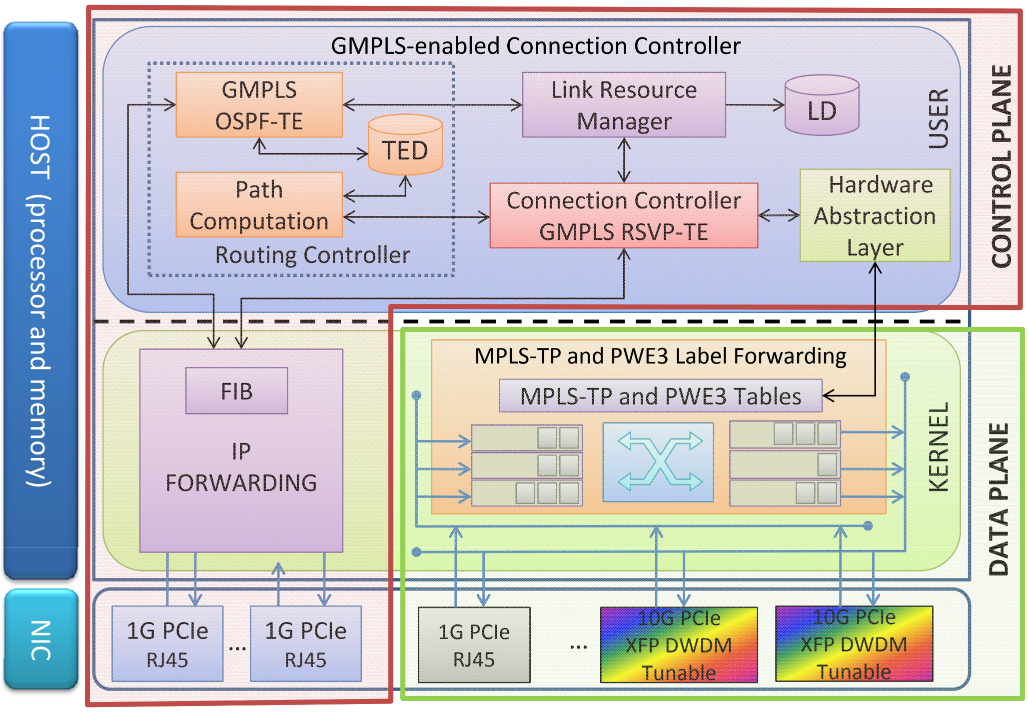


Figure 12.2: Architecture of GMPLS-controlled MPLS-TP node with integrated 10Gbps tunable DWDM transponders

The data plane enables a connection-oriented packet transport capability primarily for point-to-point connections for a number of link layer networks (e.g., Ethernet). It is based on the MPLS-TE and PseudoWire (PW) architectures. MPLS-TP uses the standard MPLS mechanisms for traffic forwarding: PUSH/POP/SWAP procedures for LSP labels, and standard LSPs and PWs encoding. In Figure 2.3, is shown an emulated (Ethernet) service for a point-to-point connection between two customer edges (i.e., CE1 and CE2). Provider Edges 1 and 2 (i.e., PE1 and PE2) are the ingress and egress routers, respectively, and are responsible for creating the PWE3 service. An Attachment Circuit (AC) is the physical or virtual circuit attaching the CE to a PE. For the Ethernet service, the AC is an Ethernet port. Indeed, the function of the PE1 is to receive the (Ethernet) PDU coming from CE1 and to encapsulate and transport it to PE2, where it is decapsulated and transmitted out on the attachment circuit (AC) of the PE2. The PW starts at a logical port within the PE (PE1). This port provides an Ethernet MAC service that will deliver each Ethernet frame to the other end (PE2) of the PW. An important function of the PE is the NSP that includes the frame processing that is required for the Ethernet frames that are forwarded to the PW termination point. The ingress NSP strips the preamble and the FCS from the Ethernet frame and transports the frame in its entirety across the PW. Then the control word , if necessary, is prepended to the resulting frame. The function of the forwarder is to select the PW based on the incoming AC, the contents of the payload or some statically and/or dynamically configured forwarding information. Next, the proper PW demultiplexer (PW label) is prepended to the resulting packet. Since the PW is transported over an MPLS-TP PSN, an inner MPLS-TP label is used to provide the PW demultiplexing function. The proper tunnel encapsulation (MPLS-TP label) is prepended to the packet, and the packet is transmitted. When the packet arrives over a PW, the tunnel encapsulation and PW demultiplexer are stripped off. If the control word is present, it is processed and stripped off. The egress NSP function receives the Ethernet frame from the PW and regenerates the preamble or FCS before forwarding the frame to the attachment circuit.

12.1.2 Implementation

The GMPLS-controlled MPLS-TP/PWE3 node has been implemented using a standard PC architecture, running Linux as operating system and equiped with several NICs.

One of the main advantages of using a general-purpose Linux software platform is that it is a cost-effective solution (open-source software), but an even more important advantage is that it provides maximum flexibility in terms of development of new functionalities both at the user-level (applications) and at the kernel-level. Reusing and extending the GMPLS protocol stack already available for the WSON network of the ADRENALINE Testbed, is of particular interest because it reduces the development time in comparison with a new implementation from scratch. Moreover, a wide range of open-source software is available and ready to be used. For example, the MPLS-TP and PWE3 label forwarding engine has been implemented with open source software easily (using the Click modular router [63]), reducing the development time in comparison with a hardware implementation based on Field Programable Gate Arrays (FPGAs).

The main disadvantage for using this proposed architecture is the potential performance bottleneck due to the communication between the host (processor and memory) and the NICs, and the software/hardware interruptions. Thus, it is not possible to reach line rate card packet processing for high data bit rates (e.g., 10 Gb/s or above). The impact of this bottleneck in the control plane is almost negligible, since the bandwidth consumed by the control channels are a few tens of Kb/s. However, it may be critical for the data plane implementation, one of the objectives of this chapter being the performance evaluation of the software MPLS-TP/PWE3 forwarding engine implementation and the integrated 10 Gbps DWDM transponder implementation.

12.2 MPLS-TP and PWE3 label forwarding engine implementation

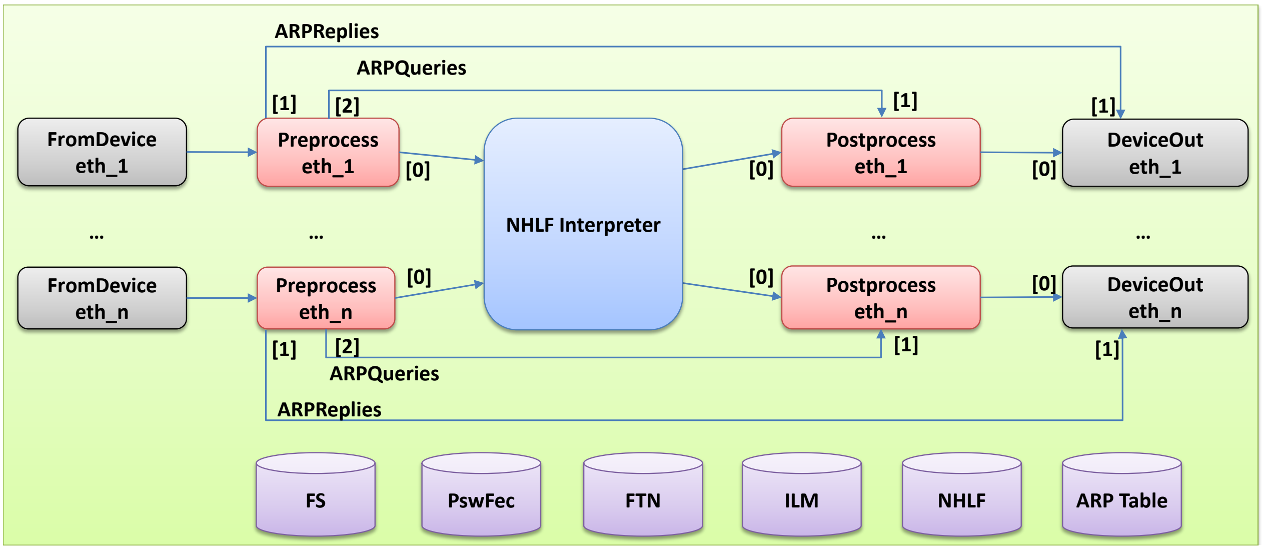


Figure 12.3: MPLS-TP/PWE3 forwarding engine click implementation

Figure 12.3 shows the architecture of the proposed MPLS-TP and PWE3 label forwarding engine, consisting of a first stage (Fig. 12.4) for the PWE3 Encapsulation of IP packets and/or Ethernet frames and the obtention of the required Next Hop Label Forwarding Entry (NHLFE), plus a second stage composed of a Next Hop Label Forwarding (NHLF) interpreter responsible for MPLS label stack processing and forwarding. Finally, a last stage (Fig. 12.5) reassembles Ethernet frames or IP packets and delivers them to the next hop.

The implementation of the forwarding engine is based on the Click modular router [63], which is an open source software architecture for building flexible and configurable routers. A Click router is assembled in a configuration file from several packet processing modules called elements. The proposed implementation includes the necessary elements to manage MPLS-TP datagrams and PWE3 ”frames”. The Click modular router can be run in kernel mode (as a kernel module) or in user mode. Kernel mode provides better performance, so it is the selected implemented solution.

The Click kernel module uses Linux /proc filesystem to communicate with user processes. To bring the MPLS-TP/PWE3 node with 10GE interfaces data plane up and running, the user creates a configuration description file in the Click language, which is loaded with the kernel module.

The /proc/click directory is used to call the necessary commands (handlers) that will fill the explained databases with the information corresponding to the switching operations of the desired LSPs. This is the main interface for GMPLS control of hardware.

12.2.1 Ethernet or IP transport services

The MPLS-TP forwarding engine architecture considers two different services: transport of either Ethernet frames or IP packets. In the proposed architecture, all Ethernet frames from a defined set of NIC used for the data plane are sniffed (using promiscuous mode). Each incoming Ethernet frame is mapped to a Forwarding Equivalent Class (FEC). The service provided to the FEC (Ethernet or IP transport) is determined in the FecToService (FS) Table.

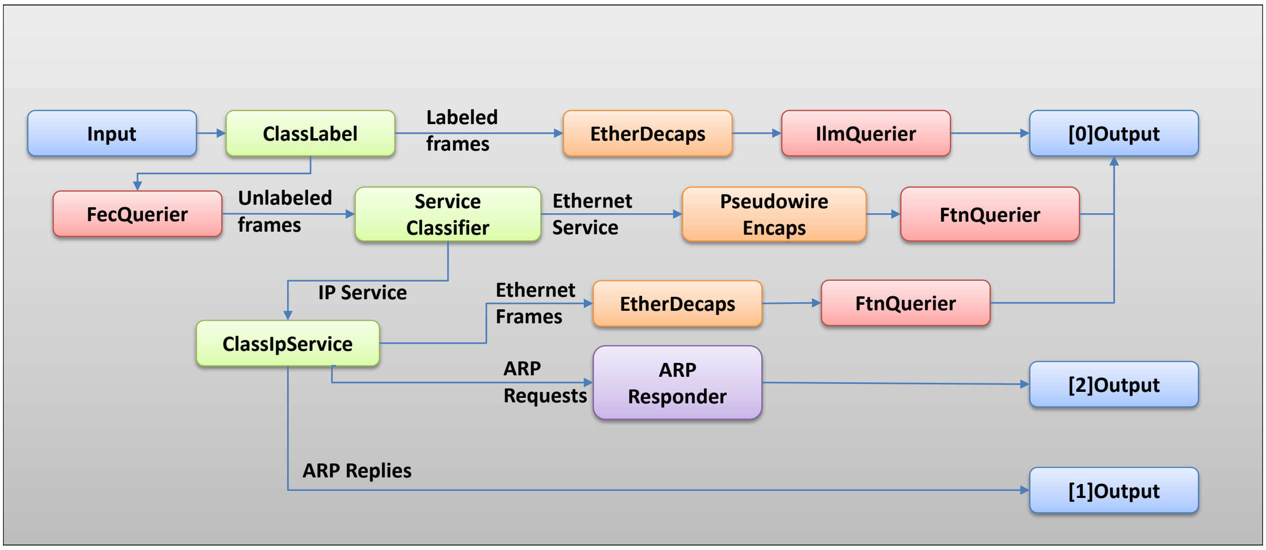


Figure 12.4: Preprocess module of the MPLS-TP/PWE3 forwarding engine click implementation

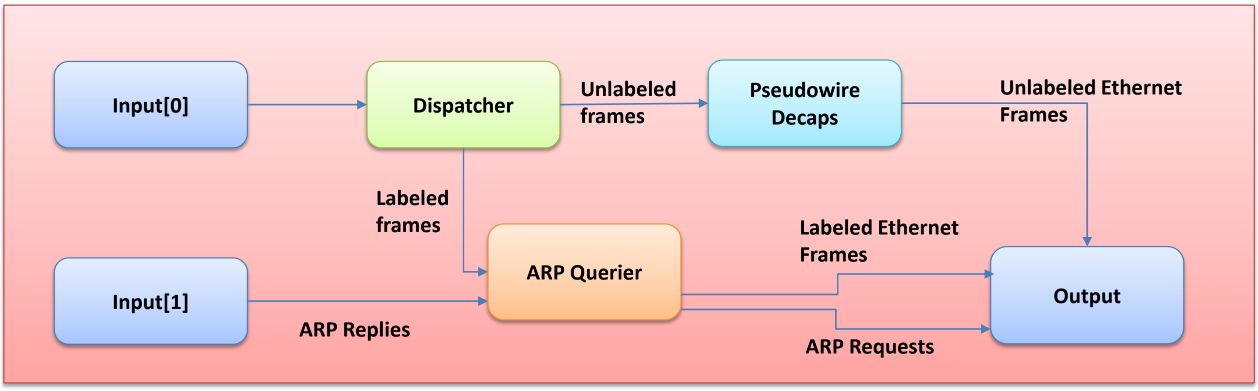


Figure 12.5: Postprocess module of the MPLS-TP/PWE3 forwarding engine click implementation

12.2.2 IP packet or Ethernet frame PWE3 Encapsulation and identification of the NHLFE

The sniffed incoming frames are classified into two types of frames: MPLS-TP labeled frames and unlabeled frames.

In case of the MPLS-TP labeled frames, the Incoming Label Map (ILM) maps each incoming labeled frame to an entry of the Next Hop Label Forwarding database (NHLFE).

For the unlabeled frames, the FEC to Service (FTS) determines whether the required service of the frame received from outside the MPLS-TP domain at the ingress node is Ethernet or IP transport. For Ethernet services, the frame is encapsulated inside a PW; the ingress Native Service Processing (NSP) function strips the preamble and FCS from the Ethernet frame. After that, the control word is prepended to the resulting frame, and optionally, fragmentation can be used since the maximum frame size that can be supported is limited (typically 1500 bytes) . PW fragmentation takes place immediately prior to PW encapsulation. MPLS-based PWE3 uses the Control Word (CW), which includes 2 bits for fragmentation information (BE)[64]. These bits are used to indicate first, intermediate or last fragment. In case that no fragmentation is applied these bits remain zero. Path Maximum Transmission Unit (MTU) can be configured dynamically to determine the maximum size for the fragments. Sequencing is provided by usage of the last 16 bits of the CW, which include the Sequence Number. The Sequence Number is incremental and it skips zero and wraps from 65535 to 1.

Finally, a FEC to NHLFE (FTN) maps each FEC to a NHLFE. As for IP service, the Ethernet header of the received frames is stripped of (i.e., Ethernet decapsulation) and the FEC is mapped into a NHLFE. At this point, both labeled and unlabeled frames have their NHLF entry, which will include the list of operations to be applied onto the frame and information about the next hop for the frame.

12.2.3 Next Hop Label Forwarding (NHLF) interpreter

The NHLF interpreter is the core element of the design. It is responsible for the operations on the MPLS label stack of each datagram and the forwarding of the datagram to the corresponding port.

Each NHLFE contains a list of operations including the next hop for the datagram and the operations to be performed on the label stack of the datagram [65]. Three operations to the label stack are defined: swap, pop and push. Swap operation replaces the label at the top of the label stack with a specified new label. Pop operation strips off the outermost label. Finally, the push operation appends a specified new MPLS label onto the label stack.

12.2.4 Forwarding of the MPLS-TP frame and delivery of Ethernet frames or IP packets

After performing the operations described in the NHLFE at the NHLF interpreter, the labeled frames and unlabeled frames without control word (i.e., IP service) are encapsulated with the proper Ethernet header at the ARPQuerier and forwarded to the next hop through the corresponding NIC.

As for the unlabeled frames including a control word (i.e., Ethernet Service), the PW Decapsulator element strips off the control word, and optionally frame reassembly is performed. If frames have been fragmented (include BE bits), they are queued following the sequence number. Out-of-order frames (and segments) are discarded. Complete frames are reassembled and dispatched directly to the corresponding NIC. If the forwarding bit-rate exceeds the NIC maximum bit-rate, the frames are discarded so the output bit-rate is bounded to the maximum NIC bit-rate.

12.3 Integrated 10Gbps tunable DWDM transponder implementation

The implemented 10Gbps tunable DWDM transponder integrated in the MPLS-TP/PWE3 node is composed of a 10 GE PCI Express NIC with a DWDM tunable OTN XFP Transceiver. This transceiver has a XFP Multi-Source Agreement (MSA) compliant footprint and allows Inter-Integrated Circuit (I2C) management. The DWDM tunable OTN XFP is 10GbE and SONET/SDH compatible, includes an Integrated G.709 compliant Digital Wrapper and is able to tune full C-Band (196.10THz - 191.60THz) with channel spacing of 50GHz.

Management of the XFP module is achieved via the XFP MSA Rev 4.5 I2C interface. Specific registers are located in vendor specific tables not used for normal XFP MSA management to avoid conflict. OTN XFP MSA I2C interface electrical properties meet specifications of the MSA Rev 4.5 Chapter 4.

The registers to access in order to support the data plane implementation, which includes Channel Tune and BER estimation, are detailled:

- The OTN tunable XFP complies with SFF 8477 laser wavelength setting and laser channel setting commands. The laser wavelength setting is controlled by writing to register 72 and 73. The laser channel setting is controlled by registers 112 and 113.
- Register-Page 5, Register 133 gives a 25 second averaged BER of the signal Pre-FEC correction. It is a rolling average over 25 seconds, and if there is a fault or no errors to correct it will always read 1e-12. When the same vendor OTN tunable XFP is connected at both ends, both Pre-FEC BER rates from each XFP can be accessed. The Far End BER information is passed through the ODU overhead.

12.4 Experimental Performance Evaluation

12.4.1 Node implementation

The GMPLS-controlled MPLS-TP/PWE3 node is deployed in a platform based on a dual processor PC architecture (2 x Xeon Intel Processor E5520 2.26Ghz 8 Threads) on a motherboard, as shown in Figure 12.6. The motherboard has one PCI-E x16 electrical and physical slot and two PCI-E x8 physical and x4 electrical. One node configuration is equipped with two quad port Cooper Gigabit Ethernet (GE) PCI Express NIC (4x1GE) from Intel (Intel PRO/1000 GT Quad Port Server Adapter), and an optical 10 GE PCI Express NIC with a XFP transceiver port. The other node configuration is equipped with an additional 10GE PCI Express NIC with XFP. The 10 GE NIC shall be connected to a PCI slot x8 electrical to obtain the best throughput performance of the NIC. Tests have been performed with a Full-C Band

The MPLS-TP/PWE3 nodes are tested under three scenarios with different traffic grooming strategies. In the first scenario (Scenario A in Fig. 12.7.a), the MPLS-TP/PWE3 node is configured to aggregate Ethernet traffic from 8 Attachment Circuits (AC, 8x1GE Cooper interfaces) and forward it to a high-speed optical tunnel (1x10GE optical interfaces). The operations involved in the MPLS-TP/PWE3 label forwarding engine are the PW encapsulation of the Ethernet frames received from the 8x1GE ports, and the push of the MPLS-TP tunnel and PW demultiplexor labels. In Scenario B (Fig. 12.7.b), the MPLS-TP/PWE3 node aggregates traffic received from 4 AC (4x1GE Cooper interfaces) and a high-speed optical tunnel (1x10GE optical interface) and forwards it to another high-speed optical tunnel (1x10GE optical interface). In this scenario, the MPLS-TP/PWE3 label forwarding engine performs the PW encapsulation, PW demultiplexor push, and MPLS-TP label push for the 4x1GE ports, and MPLS-TP label swapping for the 1x10GE port. In Scenario C (Fig. 12.7.c), the MPLS-TP/PWE3 node only considers the forwarding of incoming high-speed optical tunnel (1x10GE optical interface) to another high-speed optical tunnel (1x10GE optical interface). In this case, the MPLS-TP/PWE3 node performs MPLS-TP label swapping.

The followed methodology for the benchmark of the studied nodes is based on [66], which provides a benchmarking methodology for Network Interconnect Devices.

12.4.2 Forward engine evaluation

In this subsection, the performance of the system is evaluated, discussing the obtained key indicators such as CPU usage and throughput for frames of different sizes and depending on the number of entries in the NHLF.

Two scenarios have been configured, depending on the Symmetric MultiProcessing (SMP) affinity, since one of the key goals is to obtain more insight on the potential benefits of using multi-core architectures with high speed rates when compared to single core ones. For this:

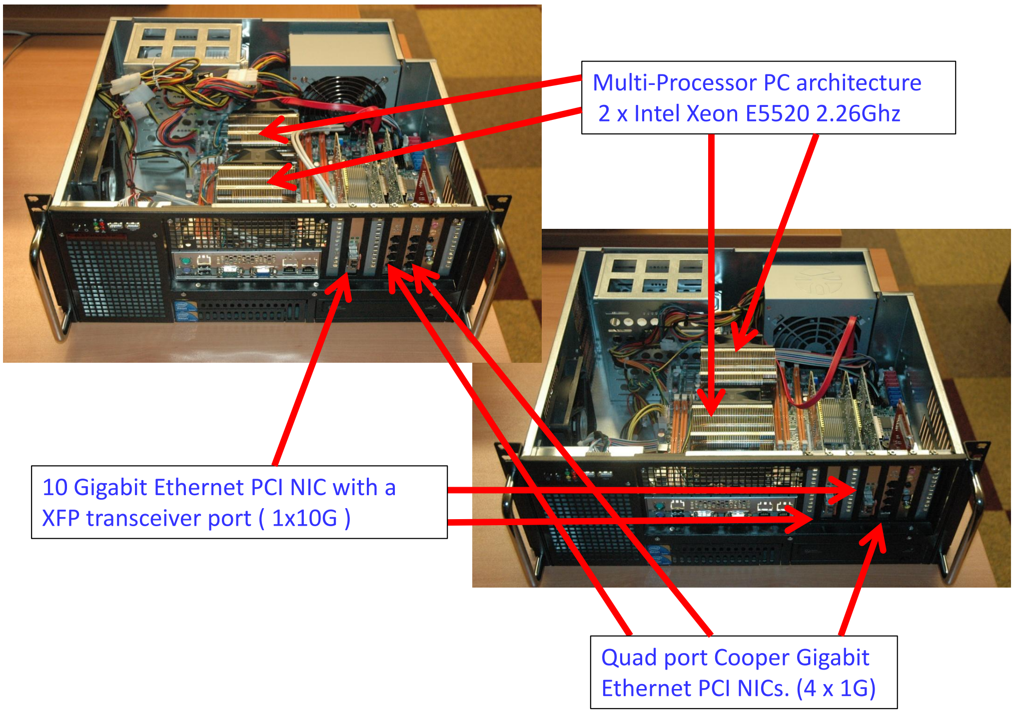
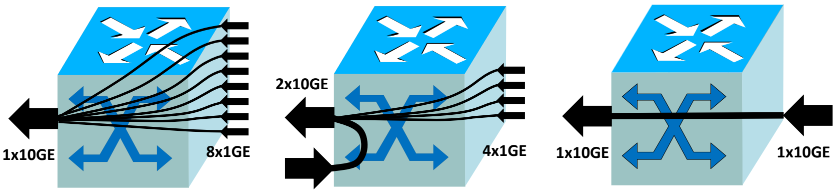


Figure 12.6: View of the MPLS-TP host node implementation

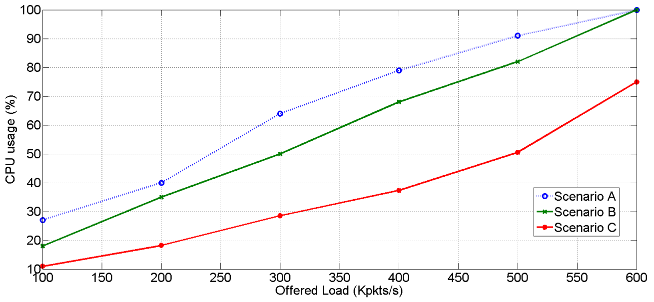


(a) Scenario A (b) Scenario B (c) Scenario C

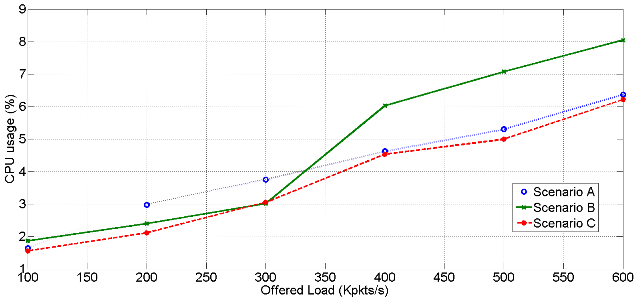
Figure 12.7: Performance scenarios

- In the first case, named *no affinity* the default processing is that the first core of the first CPU socket (CPU0) gets all interrupts coming from the network interfaces and processes the incoming frames.
- In the second case, named *full affinity* the CPU affinity is modified in order to map a given network interface IRQ to a given CPU core, configuring the SMP affinity via the /proc/irq interface.

Additionally, regarding to the actual frame forwarding, the Click SMP [67] can be used to take advantage of the multi-processor multi-core architecture, e.g. by means of a balanced thread scheduler which manages threads by minimizing variance in CPU load. However, in the proposed configuration, it does not have a strong impact in the results, because the CPU usage caused by processing of frames within the Click elements (i.e., NHLFE operations) is two orders of magnitude lower than the actual processing of interrupts from interfaces including the corresponding soft IRQ/bottom halves/takslets (i.e., the Click SMP does not improve the performance of transmitting and receiving frames to/from an interface).



(a) CPU usage - No Affinity



(b) CPU usage average - Full Affinity

Figure 12.8: CPU usage

Let us note that affinity is configurable by means of *affinity masks*, where bits can be masked to restrict an IRQ to a given core, and the mapping used in the second case is relatively straightforward. More advanced IRQ balancing could be used, since a simple mapping as the one used can still give issues due to the way that cores share the cache and the chipset (called FSB or Front Side Bus). More advanced or dynamic mappings are left for further study (for example, mapping interface IRQs to cores from different sockets, or using irqbalance [68]).

Figure 12.8.a shows the usage of the core of a CPU in the *no affinity* case when varying the offered load, the frame size is fixed to 1024 bytes and establishing only a single pseudowire tunnel in a packet switched circuit tunnel. From the figure, it can be observed that scenario A, with the highest number of interfaces, presents a higher CPU usage than scenario B. In general, the higher the number of input interfaces, the higher the number of IRQs to be processed [69]. Consequently, the CPU usage in scenario B, with less input interfaces is lower. For example, for an offered traffic of 500 Kpkts/sec, scenario B has a CPU usage of 82%, whilst scenario A consumes 91%. The same effect can be observed for Scenario C, which uses only two NICs, and consumes 50.5% of CPU for the same offered traffic load.

In the *full affinity* case, the CPU usage is computed as an average of the different CPUs loads (*mpstat* has been used for measuring this average). Figure 12.8.b shows that for a load of 500 Kpkts/sec Scenario A uses 5.2% average CPU load, Scenario B uses 7% average CPU load and Scenario C consumes 4.98% of the average CPU load. Taking a look at the segregated CPU rate provided per interface, it is noted that the CPU handling the 10 Gbps interface is almost working at 100%, acting as a bottleneck. Using the full affinity helps us improve improving the performance, but does not completely eliminate the bottleneck effect, so that the obtained results are not as good as expected.

Figure 12.9 shows the throughput of the nodes in function of the frame size in both *full affinity* case and *no affinity* case. Focusing on the *no affinity* case, it can be observed that scenario B also performs

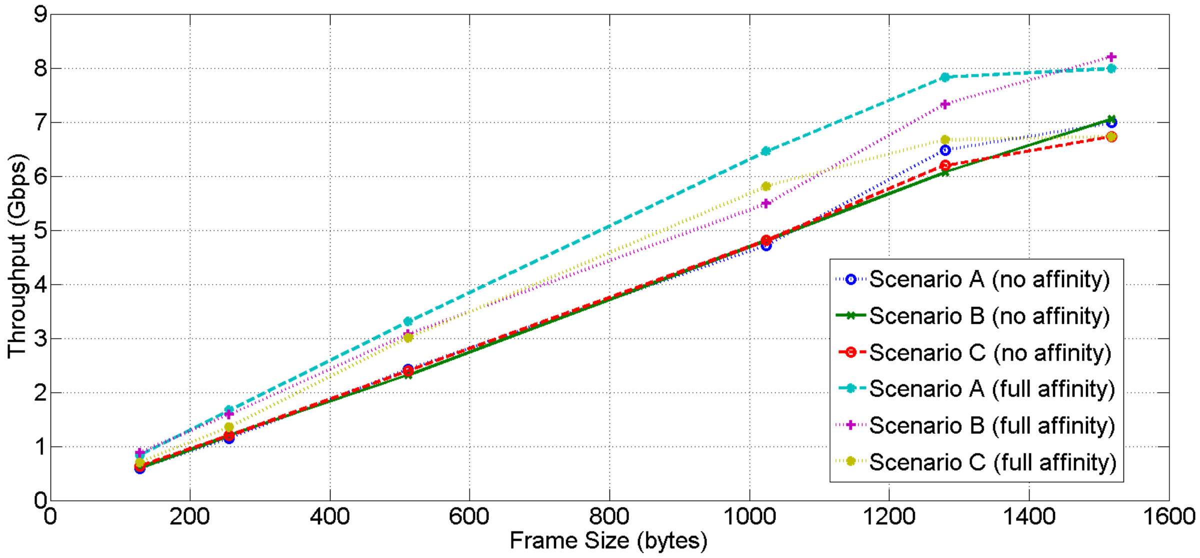


Figure 12.9: Throughput measured in gigabits per second

slightly better than scenario A. The maximum throughput in scenario B is 7.05 Gbps when the frame size is 1518 bytes, whilst the maximum throughput obtained for scenario A is 6.99 Gbps, also with the same frame size. As stated above, the scenario with the lower number of input interfaces can process a higher number of IRQ. Therefore, a higher number of packets per second can be processed, and by extension, the throughput is higher. Scenario C performs slightly worse than scenarios A and B, obtaining a maximum throughput of 6.67 Gbps for a frame size of 1518 bytes. This poor performance is due to hardware limitations. In this scenario where two XFP NICs are used, one is connected to a PCI-E x4 electrical slot. This limits the maximum bitrate of incoming traffic load, so that resulting in a worse throughput behavior. If a server motherboard had been used with two PCI x8 electrical slots instead of the proposed one the performance figure would have been much better for Scenario C.

Focusing on the obtained throughput in the *full affinity* case (Fig. 12.9), scenario A reaches the maximum expected throughput (7.99 Gbps), while scenario B reaches a maximum of 8.20 Gbps, for a theoretical maximum expected of 10 Gbps. It can be observed that the *full affinity* case always performs better than the *no affinity* case. The obtained throughput in scenario A in the *full affinity* case is 1 Gbps higher than in the *no affinity* case. In scenario B the obtained throughput is 1.15 Gbps higher for the *full affinity* case. For scenario C, the maximum obtained throuput remains the same (6.67 Gbps) due to the previously explained hardware limitations.

If the obtained throughput results are combined in the *no affinity* and the *full affinity* case, the maximum throughput is obtained with Scenario B (8.20Gbps), even though scenario A performs better on average.

It is also important to analyze the obtained throughput in units other than gigabits per second. Figure 12.10 shows the measured throughput in packets per second. It can be observed that the full affinity case is able to process more packets per second given a fixed frame size in all the proposed scenarios. For example, in scenario A, in the full affinity case the node processes 788,574 packets per second, while in the no affinity case only 576,171 packets per second are processed.

Finally, the reduction of the throughput is analyzed depending on the number of entries in the NHLF Table, which is directly related with the number of established pseudowire tunnels (see Section 12.1). In order to analyze this effect scenario A and scenario C have been used, because they make

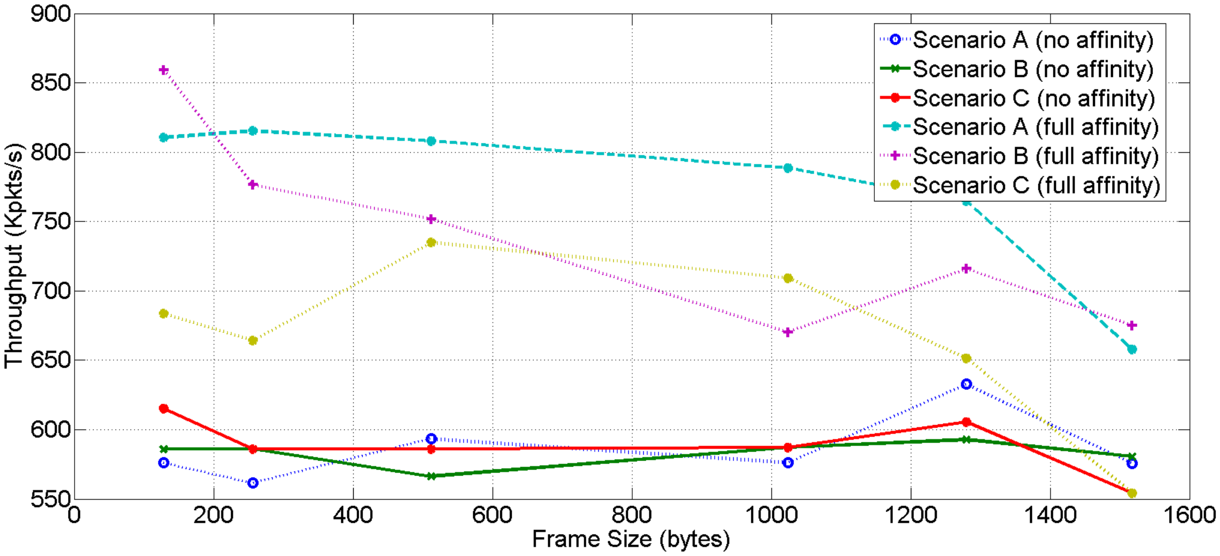


Figure 12.10: Throughput measured in packets per second

use of FTN and ILM, respectively. scenario B has been discarded because it uses both Maps, and results could be misleading. Lower throughput performance is expected as the number of entries in the NHLF table is increased.

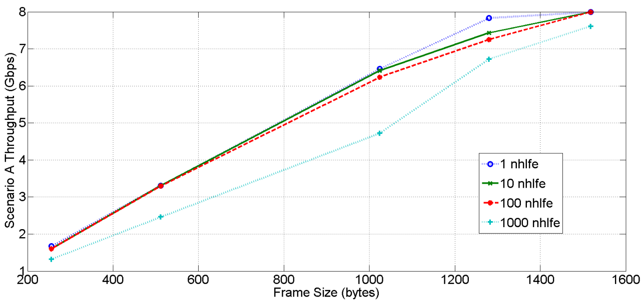
In scenario A, the throughput obtained with 1000 entries in the NHLF table for a frame size of 1024 bytes is 27% less (see Fig. 12.11.a) than the obtained throughput with a single entry (4.72Gbps and 6.46 Gbps, respectively).

Figure 12.11.b displays the obtained throughput in scenario C. The throughput variance depending on the number of entries in the NHLF table is of less significance and only observable in frame sizes below 512 bytes. For a frame size of 1024 bytes, the obtained throughput with 1000 NHLFEs is only 0.3% less than the obtained throughput with a single NHLFE (5.81 Gbps).

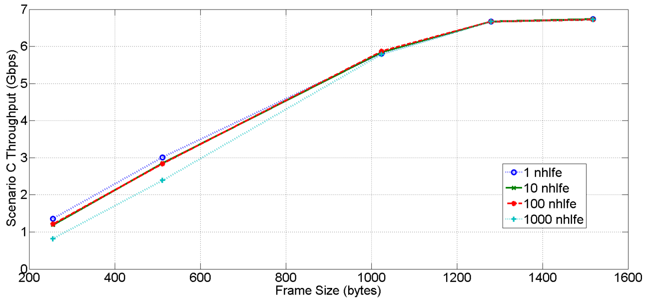
Scenario A is more sensitive to the number of entries in the NHLF Table. This can be explained because of the number of searches which are necessary to process an incoming frame. Whereas in scenario C, only one search is performed in ILM and another in NHLF, in scenario A, there are 4 related searches for each incoming frame, which devaluates significantly the obtained throughput. These 4 searches are: FEC determination (depending on the incoming port), FEC to Service classification, FEC to NHLFE and finally obtaining the action list contained in the NHLFE.

12.4.3 Evaluation of the integrated 10Gbps tunable DWDM transponder

Figure 12.12 shows the experimental set-up developed in the laboratory that is based on two MPLS-TP/PWE3 nodes, each of them including a 10Gbps DWDM tunable XFP transceiver. The transmitted signal is attenuated by 5 dB (i.e., L1, as depicted in the figure) and is received by a 50/50 optical coupler. This coupler also attaches, from the other input, the C-Band noise generated by an optical broadband source (EXFO FLS-5800) with different values of attenuation (i.e., L_2). Then, the output of this first optical coupler is connected to the input of another 50/50 optical coupler through a Single Mode (SM) fiber optic patch cord. Finally, one of the outputs of this second optical coupler is permanently connected to the receiver node, and the other output is connected to an Optical Spectrum



(a) Scenario A



(b) Scenario C

Figure 12.11: Throughput variation depending on NHLFEs

Analyzer (ANDO AQ6317C) or to a Wide-Bandwidth Oscilloscope (Agilent Infiniium DCA-J 86100C). The optical attenuators (i.e., L_1 and L_2) are necessary in order to regulate the contribution of the C-Band noise and to avoid saturation at the receiver.

The first 16 tunable 50GHz channels (196.10-195.35 THz) are shown in Figure 12.13, which represents the different traces obtained by the Optical Spectrum Analyzer, in absence of C- Band noise. The tune delay is of 30.32 milliseconds on average for all the possible 93 channels.

As a second experiment, the transmitter node is tuned to ITU-T channel H34 (equals 193.45 THz or 1549.72 nm) and the optical broadband source is switched on. Firstly, the generated C-Band noise is attenuated by 17 dB (i.e., L_2). Figure 12.14 shows the obtained Eye diagram for the received signal by the oscilloscope. In this case, the Optical Signal-to-Noise Ratio (OSNR) obtained by the optical spectrum analyzer has a value of 37.57dB. Using the previously described register, a BER of 1e-12 is obtained, which is the default value when no errors have been detected. Secondly, the C-Band noise is attenuated by 15 dB, thus increasing its impact on the data signal (see Eye diagram in Fig. 12.15). The obtained OSNR has a value of 36.16 dB in this case, and the obtained BER is 1e-10, which is worse than the previously obtained BER.

12.5 Virtualization of the MPLS-TP node

The MPLS-TP node virtualization consists on the partitioning of an MPLS-TP network node into several virtual MPLS-TP nodes with the same functionalities. A virtual MPLS-TP node has an assigned subset of the physical port interfaces.

Figure 12.16 displays the system architecture of a virtualizable GMPLS-controlled MPLS-TP node. The MPLS-TP forwarding engine has been explained in this Chapter. The virtualization server

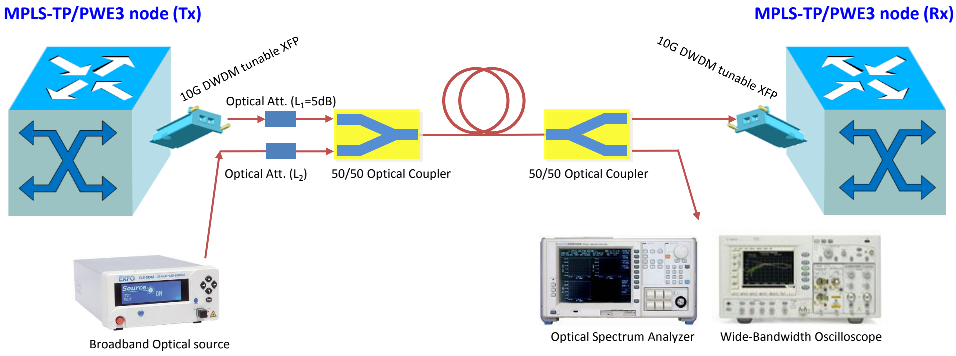


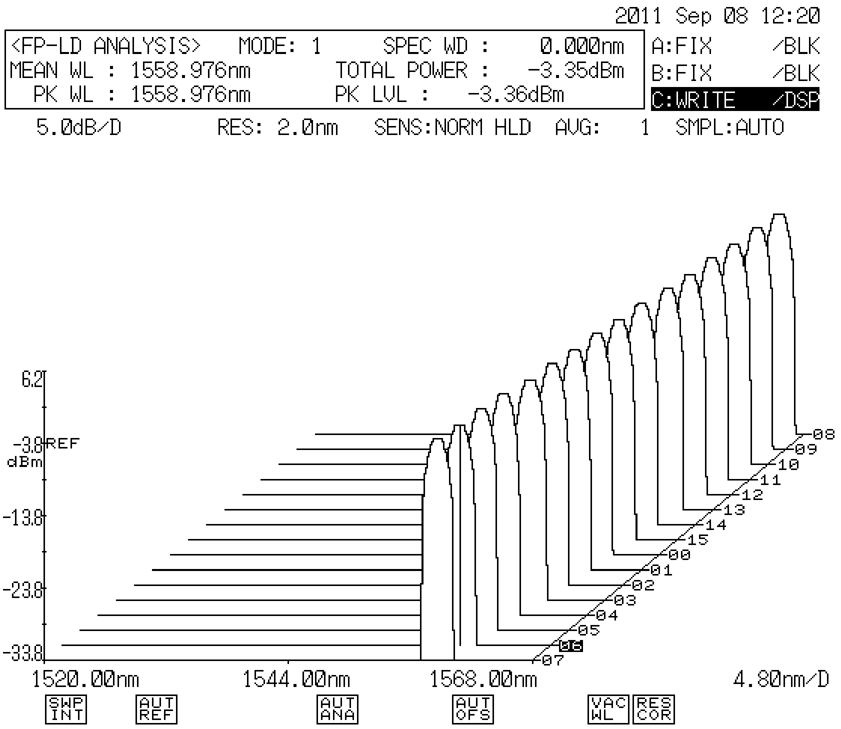
Figure 12.12: Evaluation Scenario Laboratory Set-Up

running on top of the MPLS-TP forwarding engine hosts different virtual GMPLS controllers. Each MPLS-TP port interface is managed from a single virtual GMPLS Controller. A dedicated DCN is deployed as an overlay on top of the shared physical infrastructure via KVM bridging modules, which allow connecting the physical NIC that provides connectivity to a neighboring node to each of the virtualized network interfaces, resulting in isolated IPCCs.

12.6 Conclusions

Software routers based on off-the-shelf hardware and open source software can provide a reliable, low-cost, open solution to develop forwarding nodes. In this chapter the architecture of a GMPLS-enabled MPLS-TP/PWE3 node with 10Gbps tunable DWDM transceivers has been presented and the performance of the implemented software MPLS-TP/PWE3 label forwarding engine has been evaluated in two representative multi-layer scenarios with different traffic grooming strategies. A maximum throughput around 8.20 Gbps can be achieved with a frame size of 1518 bytes, demonstrating the feasibility of the proposed solution.

A design for a virtualization architecture of the proposed MPLS-TP node has been discussed. In the next chapter, a virtual GMPLS-controlled MPLS-TP network, consisting on the presented nodes, over a shared WSON infrastructure is presented.



```
                                                        2011 Sep 08 12:20
<FP-LD ANALYSIS>   MODE: 1      SPEC WD :     0.000nm   A:FIX      /BLK
MEAN WL : 1558.976nm       TOTAL POWER :   -3.35dBm     B:FIX      /BLK
  PK WL : 1558.976nm       PK LVL :    -3.36dBm         C:WRITE    /DSP
   5.0dB/D         RES: 2.0nm    SENS:NORM HLD   AVG:    1   SMPL:AUTO
```

Figure 12.13: 10GE Tunable XFP channels 1-16

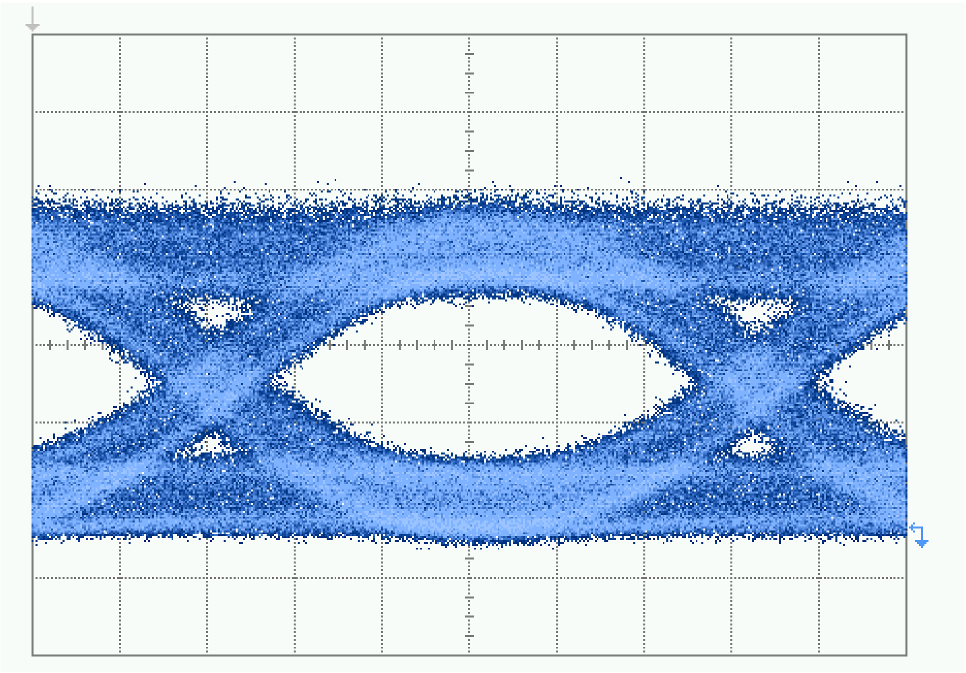


Figure 12.14: Reception eye diagram, coupling a C-band noise attenuated by 17dB

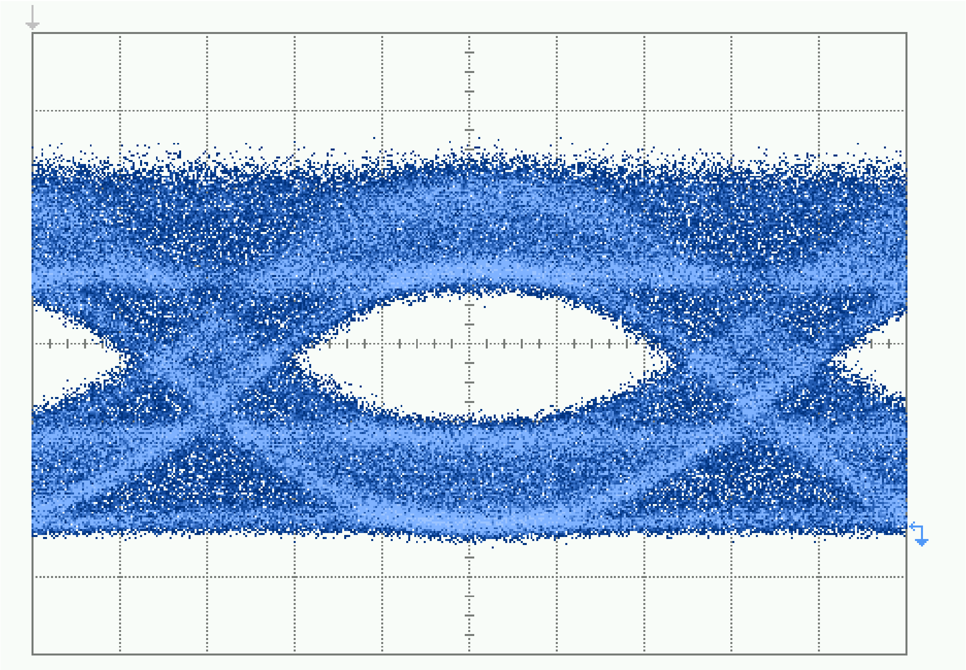


Figure 12.15: Reception eye diagram, coupling a C-band noise attenuated by 15dB

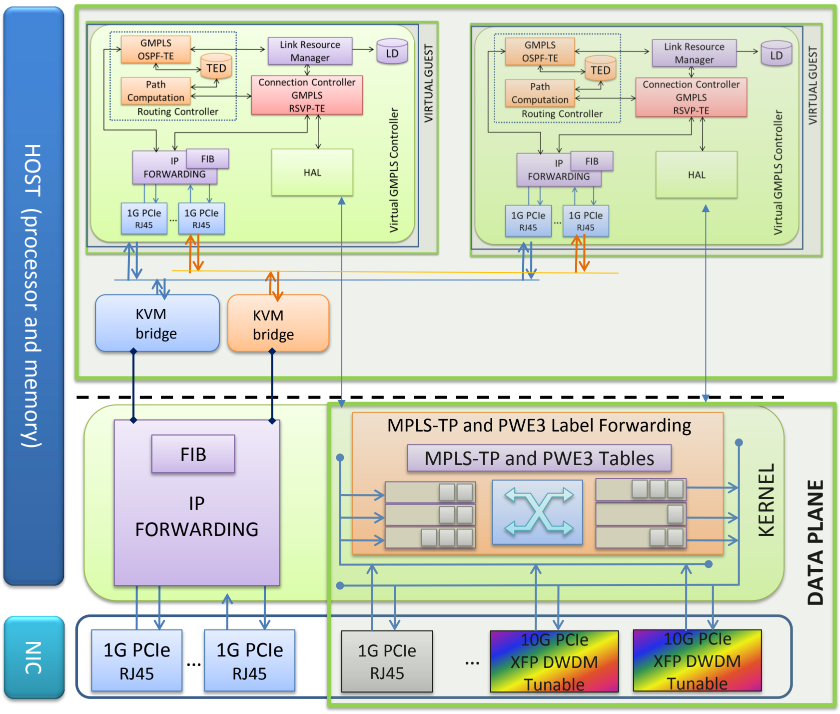


Figure 12.16: Virtualization of the MPLS-TP node

Chapter 13

Dynamic deployment of virtual GMPLS-controlled MPLS-TP networks over a shared WSON through a resource broker

In chapter 7, the author has presented an resource broker for the dynamic deployment of virtual GMPLS-controlled WSON networks over a physical GMPLS-controlled WSON infrastructure. In chapter 12, a GMPLS-controlled MPLS-TP node has been designed, implemented and evaluated. Also, a virtualization architecture for the MPLS-TP node has been proposed.

This chapter extends the previous work in the resource broker with new functionalities to dynamically deploy virtual GMPLS-controlled MPLS-TP networks over a GMPLS/PCE WSON infrastructure. The system architecture has been experimentally validated on the ADRENALINE Testbed.

13.1 Data Center Interconnection

Emerging cloud applications (e.g., real-time data backup, remote desktop) require more traffic being delivered between Data Centers (DC). Moreover, multiple, independent SLAs will be required to satisfy the different dynamic QoS requirements (e.g., low latency) of cloud applications.

Current DC interconnections are based on Ethernet transport services, which are provided through the classical IP/MPLS stack [70]. There is a need for the development of efficient and dynamic technologies for Ethernet transport services. In this context, MPLS-TP technology is seen as an appealing packet transport solution of Ethernet services. On the other hand, the accelerating growth of data rates has driven network operators to seek for more efficient alternatives enabling traffic offloading

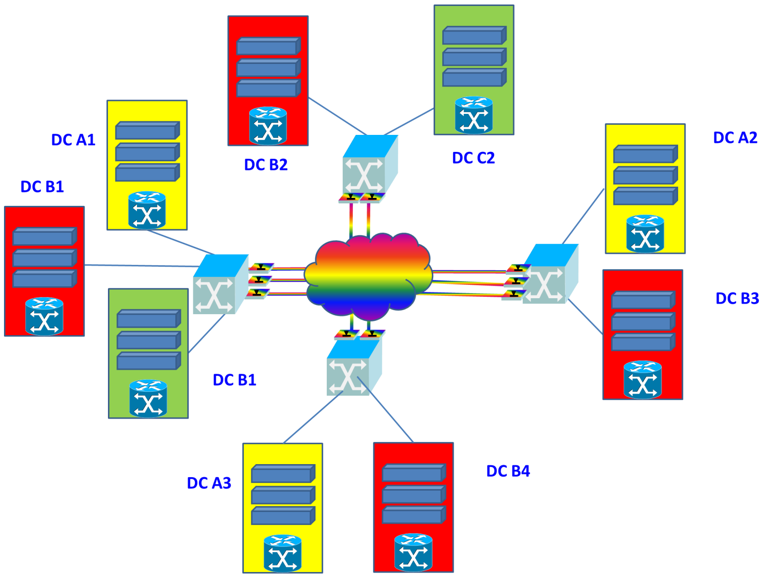


Figure 13.1: Data center interconnection

to the optical layer. WSON provide the required coarse transport capacity, while packet optical switches, located at the edge of each data center (see Fig.13.1), provide the benefits of bandwidth flexibility and connection-oriented packet-based services.

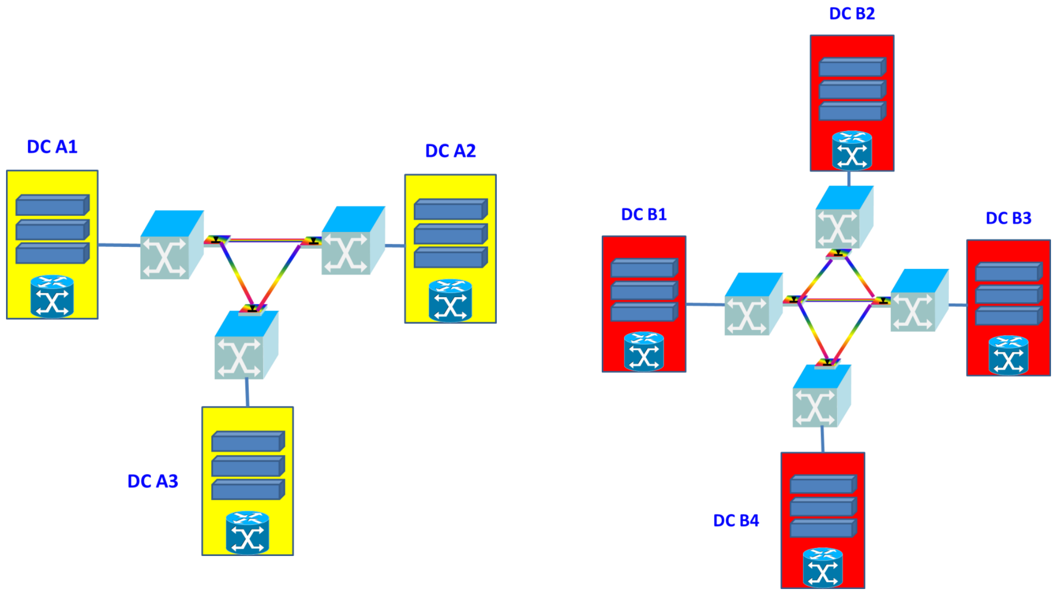
An overlay solution is proposed where the packet optical switches are virtualized, allowing the deployment of different virtual packet (MPLS-TP) infrastructures over optical (WSON) network topologies for connecting different DCs. A virtual GMPLS-controlled MPLS-TP network [8] provides the required flexibility for the proposed cloud applications.

Packet Switch Capable (PSC) Label Switched Paths (LSP) can be established upon the virtual MPLS-TP network for each cloud application (Fig. 13.2.a and Fig. 13.2.b).

The proposed virtual MPLS-TP nodes have several ports with tunable transponders in order to connect as clients of the WSON (i.e., acting as alien wavelengths, a ”coloured” optical signal which has not been originated under the direct control of the network operator). A Resource Broker is introduced to interact with a shared GMPLS-controlled WSON, while controlling the virtual MPLS-TP nodes. The established optical LSPs over the GMPLS-controlled WSON are offered as packet TE links to the virtual MPLS-TP nodes. In order to provide the virtual MPLS-TP networks with automated connection provisioning, traffic engineering and dynamic protection/restoration, a virtual GMPLS control plane manages the constructed virtual MPLS-TP network. On top of each virtual MPLS-TP node a virtual GMPLS controller is deployed.

13.2 Virtual GMPLS-controlled MPLS-TP Network Resource Broker System Architecture

As explained in section 12.5, each virtualizable MPLS-TP node is equipped with a virtualization server, which provides different virtual GMPLS controllers. Each virtual MPLS-TP node will be configured with independently partitioned tunable ports. For example, a physical MPLS-TP node with tunable DWDM ports a, b and c, connected as clients to an optical ROADM, can be virtualized



(a) Virtual GMPLS-controlled MPLS-TP Network 1 (b) Virtual GMPLS-controlled MPLS-TP Network 2

Figure 13.2: Virtual GMPLS-controlled MPLS-TP networks

into two independent virtual MPLS-TP nodes, with port a, and ports b-c, respectively (Fig. 13.3). The virtualization of a distributed GMPLS-based control plane responsible for handling dynamically and in real-time the resources of the virtual MPLS-TP nodes for the automatic provisioning and survivability of LSPs, has also been studied in section 12.5.

Figure 13.3 shows the proposed system architecture including a resource broker for dynamic deployment of virtual GMPLS-controlled MPLS-TP networks. The resource broker has been fully refactored in order to manage the incoming asynchronous and dynamic virtual MPLS-TP network requests. A virtual MPLS-TP network request is modeled as a graph that describes a set of virtual MPLS-TP nodes and links. The request, which is handled by the request controller, also includes the requirements for the control plane, such as the needed capacities for the virtual GMPLS controllers, or the selected values for configuring the parameters of the control processes running on the virtual GMPLS controllers. A MPLS-TP node must be linked to at least one optical switch.

The request controller accepts incoming TCP sessions, used to reliably transport virtual GMPLS-controlled MPLS-TP network requests, and handles these requests asynchronously and dynamically. The request controller triggers the resource allocator in order to process the request for setup, modify or tear down the virtual GMPLS-controlled MPLS-TP network. The proposed resource broker is able to interact with a GMPLS/PCE WSON by means of a VNTM. The VNTM is composed of a PCC and a LSP Manager. Using the GCO [59], all the requested links are grouped into a SVEC object within a single *PC_Request* for the PCE [58].

The PCE will reply with a *PC_Reply* specifying the Explicit Route Object (ERO) for each LSP. Each ERO is used by the LSP Manager to set up the optical LSP using the RSVP-TE signaling that is running on the source optical node connected to each requested virtual MPLS-TP node. A soft permanent connection including the ingress port of the source optical node and the egress port of the destination optical node is requested. Once each optical LSP has been established, the ingress GMPLS controller replies with the final path detailed into the Record Route Object (RRO). For the

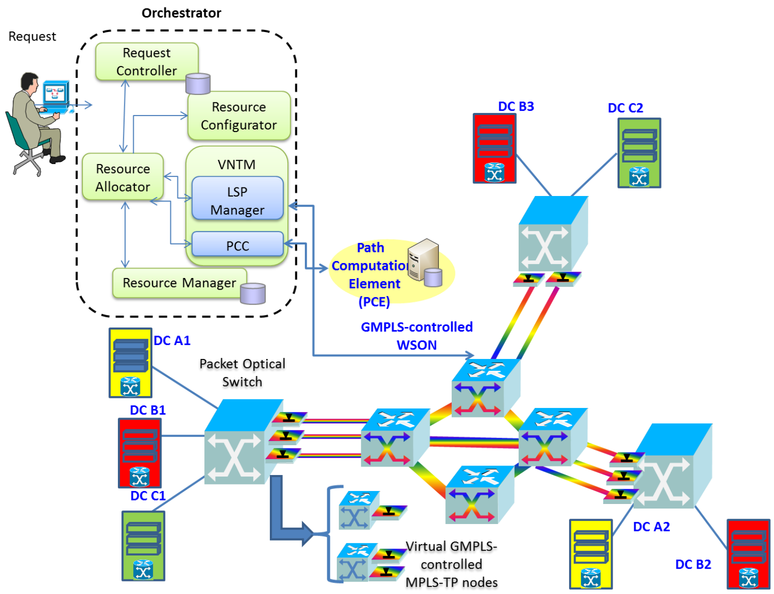


Figure 13.3: Virtual GMPLS-controlled MPLS-TP Network Resource Broker System Architecture

established optical LSP, the Resource Broker assigns a packet TE link identifier, and configures this LSP as a packet TE link in the configuration file of the virtual GMPLS control plane processes (e.g., OSPF-TE and local link resource manager) of the virtual GMPLS controller of the virtual MPLS-TP node. The TE link identifier is mapped to the physical port of the MPLS-TP node at the hardware abstraction layer process.

The resource manager, resource allocator and resource configurator share the same functionalities which have been described in chapter 7.

Once the virtual GMPLS-controlled MPLS-TP network are deployed, different PSC LSPs can be established on top of the network to cope with the dynamic requirements of incoming cloud applications.

13.3 Experimental evaluation

Figure 13.3 describes the resource broker system architecture on the GMPLS/PCE WSON platform of the ADRENALINE Testbed. It is composed of an all-optical PCE/GMPLS-controlled WSON infrastructure with 2 ROADMs and 2 OXCs providing reconfigurable (in space and in frequency) end-to-end lightpaths. The previously presented 3 MPLS-TP nodes with 10 Gbps tunable transceivers have been included. Each MPLS-TP node is equipped with a virtualization server running in a Linux-based router.

As an example (Fig. 13.4), a virtual MPLS-TP network is requested consisting of 2 virtual links between nodes 6 and 7 (i.e., 6-7 and 7-6). The links are considered unidirectional, for simplicity. With the purpose of obtaining a path for the requested virtual links, the Resource Broker issues to the PCE a *PC_Request* message including an SVEC object, where the different requested links are introduced. As an example, the PCE uses a shortest path algorithm. Once the spatial paths for the links have been computed, the PCE replies with a *PC_Reply* message (Fig. 13.5) which includes the EROs (spatial path) and the possible labelsets.

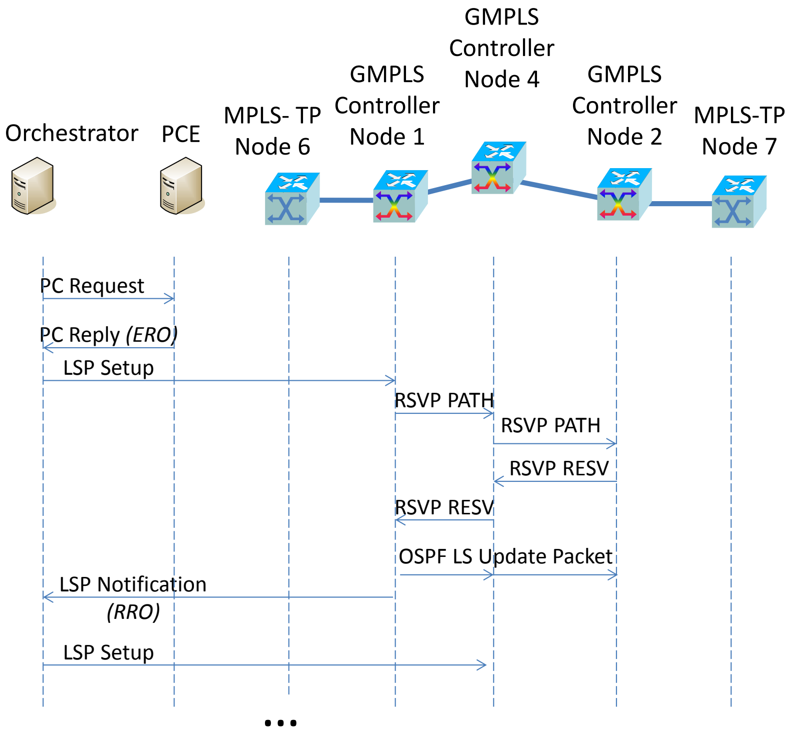


Figure 13.4: Resource broker message exchange

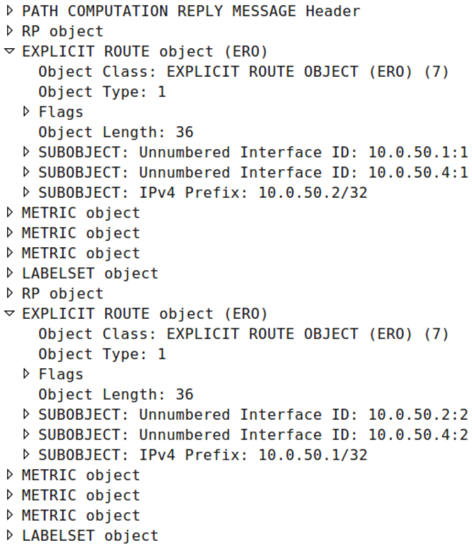
The *PC_Reply* message is processed and, through the LSP Manager, a soft permanent connection is requested via an XML interface to the RSVP process of the optical source node. Once the LSP setup request message is received at node 1, the LSP is established following the RSVP standard procedure. Finally, node 1 acknowledges to the Resource Broker the setup of the LSP. The LSP establishment is performed for all the requested virtual links.

Once all the necessary LSPs have been established, the virtual GMPLS control plane resources need to be allocated. To this end, the established LSPs are used as virtual packet TE links. Last, once the configuration is generated, the resource configurator, by means of ADNETCONF, is responsible to set up or tear down the requested virtual GMPLS-controlled MPLS-TP network.

Figure 13.6.a and Figure 13.6.b show the time for PCE to compute the requested paths (Tpce) for each proposed network (10 and 15ms), respectively. Tcc refers to the time for the LSP Manager to set-up the required LSPs (50 and 104ms). The virtual MPLS-TP network service setup delay is defined as the time required to allocate the resources (i.e., the delay to compute the paths and provision all the required LSPs for a virtual MPLS-TP network) for a virtual MPLS-TP network. Setup delay results in 62 ms and 121 ms, respectively.

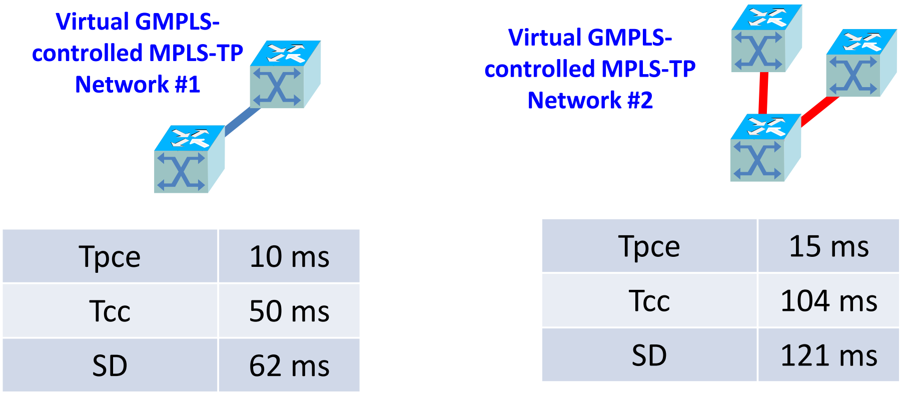
13.4 Conclusions

Experimental assessment carried out in the ADRENALINE testbed has shown the feasibility of deploying independent instances of GMPLS control plane for each deployed virtual MPLS-TP network under dynamic virtual network requests, providing low setup delays.



```
▷ PATH COMPUTATION REPLY MESSAGE Header
▷ RP object
▽ EXPLICIT ROUTE object (ERO)
    Object Class: EXPLICIT ROUTE OBJECT (ERO) (7)
    Object Type: 1
  ▷ Flags
    Object Length: 36
  ▷ SUBOBJECT: Unnumbered Interface ID: 10.0.50.1:1
  ▷ SUBOBJECT: Unnumbered Interface ID: 10.0.50.4:1
  ▷ SUBOBJECT: IPv4 Prefix: 10.0.50.2/32
▷ METRIC object
▷ METRIC object
▷ METRIC object
▷ LABELSET object
▷ RP object
▽ EXPLICIT ROUTE object (ERO)
    Object Class: EXPLICIT ROUTE OBJECT (ERO) (7)
    Object Type: 1
  ▷ Flags
    Object Length: 36
  ▷ SUBOBJECT: Unnumbered Interface ID: 10.0.50.2:2
  ▷ SUBOBJECT: Unnumbered Interface ID: 10.0.50.4:2
  ▷ SUBOBJECT: IPv4 Prefix: 10.0.50.1/32
▷ METRIC object
▷ METRIC object
▷ METRIC object
▷ LABELSET object
```

Figure 13.5: Wireshark Capture of a Path Computation Reply



Tpce	10 ms
Tcc	50 ms
SD	62 ms

Tpce	15 ms
Tcc	104 ms
SD	121 ms

(a) Virtual GMPLS-controlled MPLS-TP Network 1

(b) Virtual GMPLS-controlled MPLS-TP Network 2

Figure 13.6: Service Setup Delay

Part IV

Optical Network Virtualization in Heterogeneous Control Domains

Chapter 14

Dynamic deployment of Virtual Optical Networks with Heterogeneous Control Domains

This chapter proposes a resource broker to dynamically provision multi-domain VON across heterogeneous control (GMPLS, OpenFlow) domains and transport (OPS, EON) technologies. Experimental evaluation has been performed in an international testbed across Spain, UK and Japan.

14.1 Data Center interconnection across multiple domains

Current Data Center (DC) interconnections are based on Ethernet transport services, which are provided through the classical IP/MPLS stack [70]. DC interconnection traffic stands for 7% of global DC traffic and 284 EB per year are expected in 2014 [71]. In consequence, there is a clear need for the development of efficient and dynamic network technologies for Ethernet transport services.

In this context, Optical Packet Switching (OPS) technology is seen as an appealing packet transport solution for offering Ethernet services. On the other hand, Elastic Optical Networks (EON) provides long-reach optical transport for data rates beyond 100Gb/s. Thus, EON will provide the required flexible transport capacity at the backbone networks, while OPS switches, used for intra-DC connections, provide the benefits of statistical multiplexing and connection-oriented packet-based services. Moreover, virtualization enables physical infrastructure providers to partition and compose their physical resources into multiple independent slices (i.e., virtual networks) with each virtual resource exactly mimicking functionality and performance of the real physical resource slices.

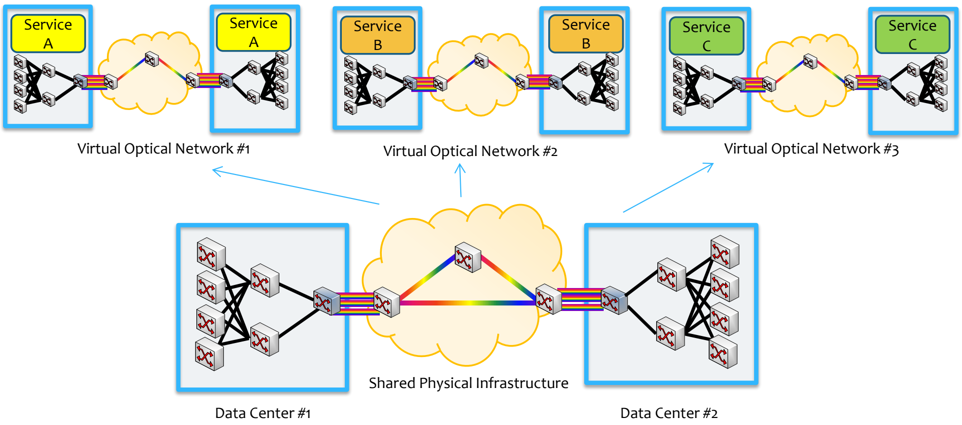


Figure 14.1: Data Center Interconnection with Virtual Optical Networks

Virtual Data Centers (VDC), comprising both virtual networks (interconnecting the distributed DC) and virtual IT resources (compute and storage within DCs), have been proposed for supporting multi-tenancy and application-specific requirements of DCs (see Fig. 14.1) [50]. While IT resources can be easily virtualized, the provisioning of a virtual network for DC interconnection remains a research challenge, due to the fact that a DC interconnection usually comprises heterogeneous optical transport and control plane technologies, which do not naturally interoperate. The authors in [35] have proposed virtual optical network services across multiple domains, but this work did not take into account the inherent heterogeneousness of multiple control domains.

We propose a virtualization mechanism which allows the composition of Virtual Optical Networks (VON) across different transport technologies (i.e., OPS, EON) and control plane technologies (e.g., OpenFlow or GMPLS). The obtained VON domains can be controlled by GMPLS or OpenFlow (depending on the virtualization technology) and a service orchestration mechanism could be used to provide end-to-end connectivity.

14.2 Multi-domain VON System Architecture

The proposed Multi-domain Resource Broker (MRB) system architecture provides a mechanism for virtualizing transport nodes and links. The partitioning of the resources is technology dependent, and to this end, the proposed system architecture incorporates a generic network slicing abstraction mechanism for the different transport infrastructure resources (e.g., OPS, EON).

We propose three different implementations of the network slicing abstraction mechanism, with the purpose to provide a virtualization mechanism in GMPLS-controlled EON, OF-controlled EON and OF-controlled OPS. The virtualization of a GMPLS-controlled EON and of an OF-controlled EON have been addressed in [72] and [50], respectively. To the best of our knowledge the virtualization of an OPS network has not been addressed previously. For simplicity, we assume a virtualization model of an OPS node based on port partitioning.

The MRB controls the VON deployment by means of the different Virtualization Visors (VV) (Fig. 14.2). These VV are responsible for the virtualization of optical infrastructure domains. Each proposed VV partitions each domain resources (i.e. links and nodes) into virtual optical resources. Later, the obtained virtual optical resources are composed into actual VONs, controlled by either a GMPLS or an OpenFlow control plane, assigned by each VV.

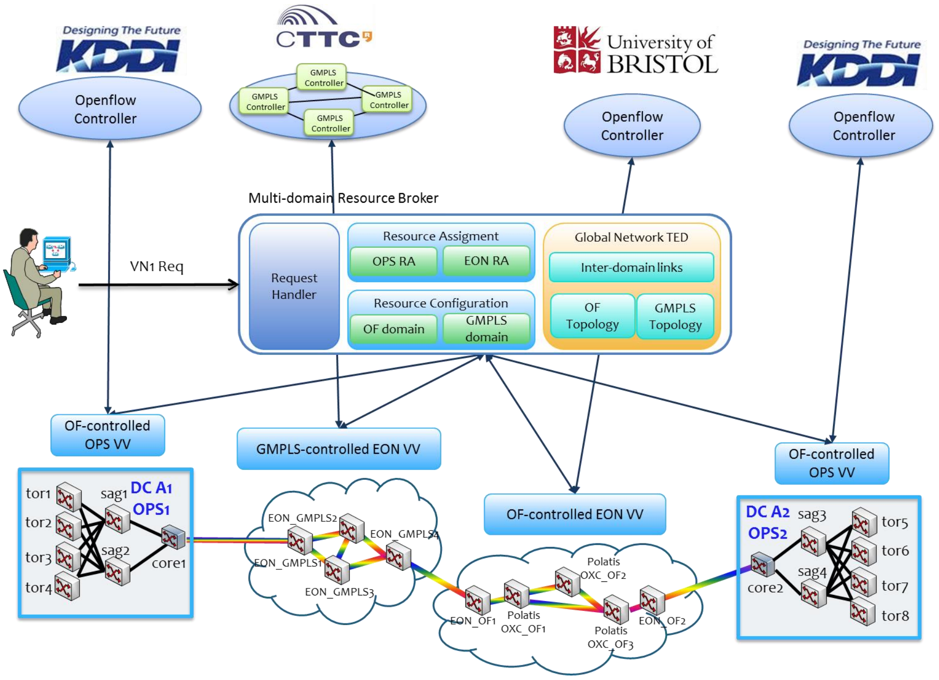


Figure 14.2: Virtualization testbed architecture

The MRB consists of four main components: the request handler, the resource assignment module, the resource configurator module and the global network TED. The request handler accepts VON requests from a client, using incoming TCP sessions for the reliable delivery, and handles these requests asynchronously and dynamically. A VON Request consists of a XML file, describing the requested virtual nodes, the requested virtual links between these nodes and the required minimum guaranteed bandwidth. For each VON setup request a VON identifier (VON-ID) is assigned. The VON-ID will be used by the MRB to map the assigned resources in each domain to the VON request.

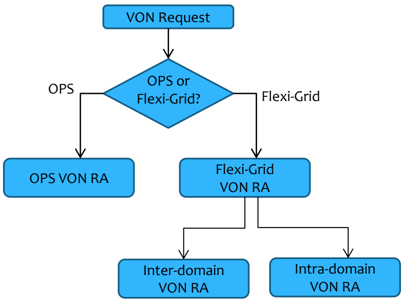


Figure 14.3: Virtual network allocation algorithm

Resource assignment algorithms need to be introduced, focusing on the optimal planning (i.e., off-line) or the dynamic request allocation (i.e., on-line) of VON requests. We have developed an algorithm bundle (Fig. 14.3), including several algorithms designed for different scenarios (i.e., single/multi-domain flexi-grid, OPS). The algorithm bundle reads the information of physical networks (e.g., the network topology and the availability of physical resources of each domain, such as ports or spectrum slots, and the inter-domain connectivity) from the global network TED, which is the infrastructure resource database and contains all the necessary information on the virtualized resources, such as spectrum availability or virtual control resources, as well as the status of the inter-domain links, and is obtained from a local XML file, although a topology discovery mechanism for VV is expected (e.g., topology server, BGP-LS [73]). Taking into account the above inputs and optical constraints (i.e., spectrum continuity), the algorithm bundle will compose a VON that can satisfy the user's request and optimize the physical resource utilization.

A Multi-Domain Shortest Path First-Fit Spectrum Allocation algorithm is used to find a physical path for each virtual link (Fig.1b). To satisfy the Spectrum Continuity Constraint (SCC), an available first-fit Frequency Slot (FS) with the requested bandwidth is located among all the available spectrum slots for the whole multi-domain EON topology. After running the algorithm bundle, all the involved network resources for configuring the flexi-grid equipment are generated as outputs and sent to the corresponding VVs. The OF VV consists on an OpenFlow FlowVisor, while the GMPLS VV consists on a previously proposed Resource Broker for deploying virtual GMPLS-controlled elastic optical networks [72].

14.3 Experimental Assessment

To experimentally evaluate the proposed virtualization architecture, a heterogeneous multi-domain international testbed has been built comprising an EON domain in the High Performance Networks group at University of Bristol (UK), a layer 2 optical packet switched domain in KDDI R&D Labs (Japan) and an EON domain in CTTC (Spain) as shown in Fig. 14.2.

The University of Bristol testbed is comprised of an in-house built 8x8 (4x4 bidirectional) BV-OXC utilizing two BV-WSS switches with internal recirculation fibre loops to emulate multiple nodes; a BV transponder (BV-TX & BV-RX) supporting C-band and 3 OpenFlow-enabled Polatis fibre switches. The CTTC GMPLS control plane platform of the ADRENALINE Testbed includes 14 nodes that run GMPLS Controllers with emulated EON hardware. A packet-based emulated network with a DC network topology (including ToR, aggregation and distribution layers) has been deployed in KDDI. The international connectivity between the MBR and the different VV running on each testbed is provisioned over VPN Tunnels over Internet.

Figure 14.4 shows the message exchanges between the MRB and the different network elements responsible to setup a virtual network. A user creates a virtual network request (VN1) which might imply several domains, and sends it to the MRB via XML interface. The MRB runs the proposed algorithm (Fig. 14.3) and for each required domain contacts the required VV. In the example shown, an OF-controlled EON VV and a GMPLS-controlled EON VV are contacted.

The OF-controlled EON VV corresponds to an optical Flowvisor, so the VV uses the XML-RPC API [33] to create the required slices through the definition of the required flowspaces for each slice [32]. The GMPLS-controlled EON VV is requested via a proprietary interface which has been described in [72]. Once the different VVs notify that the virtual resources for VN1 have been allocated (VN1 ACK), the user is notified with the assigned resources (VN1 ready).

Figure 14.5 shows the results that were obtained to verify the functionalities of the proposed MRB architecture for different network domains (alternating CTTC, University of Bristol and KDDI domains). The VON Domain slice setup time is affected significantly by control plane VPN delay as well as the communication method between each OF agent or GMPLS controller and its corresponding NE. The different VON domain slice setup times can be easily explained by the fact that different

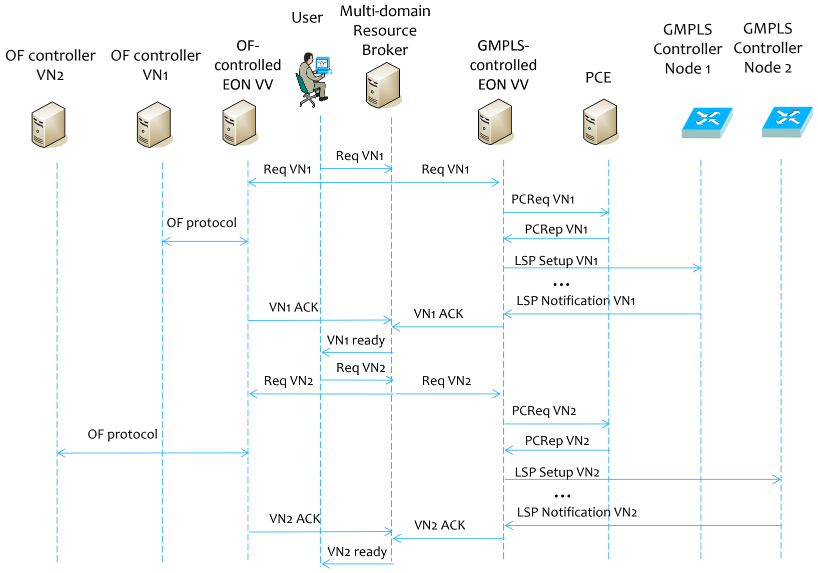


Figure 14.4: Multi-domain resource broker message exchange

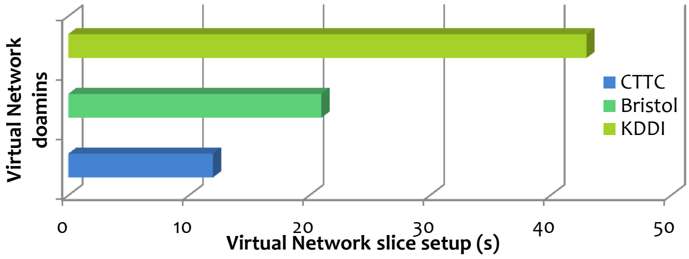


Figure 14.5: Virtual Network Domains slice setup time (s)

VV using different technologies where used in each domain. Also the different intra-domain network topologies explain the obtained results, which are tightly coupled with each intra-domain topology. It is also remarkable the need for a faster interface towards FlowVisor, where several requests for different flowspaces could be grouped allowing faster VON domains slice setup times.

14.4 Conclusions

A multi-domain resource broker has been proposed for providing dynamic Virtual Optical Networks as a service in heterogeneous control domains and transport technologies. The proposed MRB design has been presented, and an experimental evaluation has been performed in an international testbed across Spain, UK and Japan.

Part V

Dissemination and Exploitation Results

Chapter 15

Scientific publications

This chapter presents the scientific publications published, accepted or submitted. Publications are classified in journals, international conferences, national conferences and collaborations.

15.1 Journals

1. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Virtual Optical Network Resource Allocation using PCE Global Concurrent Optimization for Dynamic Deployment of Virtual GMPLS-controlled WSON (accepted), Journal of Optical Communications and Networks, 2013.

2. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Dynamic Virtual GMPLS-controlled WSON Using a Resource Broker with a VNT Manager on the ADRENALINE Testbed, Optics Express Vol. 20, Iss. 28, pp. 29149-29154, Dec 2012.

3. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez GMPLS-enabled MPLS-TP/PWE3 node with integrated 10Gbps tunable DWDM transponders: design and experimental evaluation, Elsevier Journal of Computer Networks, Special Issue High-performance Switching & Routing, vol. 56, pp. 3123-3135, 2012.

15.2 International conferences

1. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Data Center Interconnection Orchestration with Virtual GMPLS-controlled MPLS-TP Networks over a Shared Wavelength Switched Optical Network, in Proc. of 39th European Conference and Exhibition on Optical Communication (ECOC 2013), September 22-26 2013, London, U.K.
2. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Dynamic deployment of virtual GMPLS-controlled elastic optical networks using a virtual network resource broker on the ADRENALINE testbed (Invited), in Proc. of 15th International Conference on Transparent Optical Networks (ICTON), June 23-27, 2013, Cartagena (Spain).
3. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Dynamic Virtual GMPLS-controlled WSON Using a Resource Broker with a VNT Manager on the ADRENALINE Testbed, in Proc. of 38th European Conference and Exhibition on Optical Communication (ECOC 2012), September 16-20 2012, Amsterdam, The Netherlands.
4. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Dynamic Wavelength Allocation Algorithms for On-Demand Deployment of GMPLS-controlled Virtual WSON Using a VON Resource Broker, in Proc. of 17th European Conference on Network and Optical Communications (NOC 2012), June 20-22, 2012, Vilanova i la Geltru, Spain.
5. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Experimental evaluation of a VON Resource Broker and Compositor for deployment and performance of virtual GMPLS control planes, in Proc. of 16th International Conference on Optical Networking Design and Modeling (ONDM), April 17-20, 2012, Colchester, UK.
6. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Experimental demonstration of a Virtual Optical Network Resource Broker and Compositor for dynamic GMPLS WSON infrastructure services,in Proc. of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), March 4-8, 2012, Los Angeles, California, USA.
7. R. Vilalta, R. Muñoz, R. Casellas, R. Martínez, Design and performance evaluation of a GMPLS-enabled MPLS-TP/PWE3 node with integrated 10Gbps tunable DWDM transponders, in Proc. of IEEE Conference on High Performance Switching and Routing (HPSR 2011), July 4-6 2011, Cartagena, Spain. Best Student Paper Award.
8. R. Muñoz, R. Vilalta, R. Casellas, R. Martínez, F.J. Vílchez, First Lab Trial of an integrated MPLS-TP/WSON transport network controlled by a unified GMPLS control plane in the ADRENALINE testbed (Invited Paper), in Proc. of Photonics in Switching (PS) 2012, September 11-14 2012, Ajaccio, Corsica Island, France.
9. R.Muñoz, R. Casellas, R. Martínez, R. Vilalta, J. Vílchez, J. Vázquez, Virtualizing ADRENALINE testbed for deploying dynamic GMPLS-controlled WSON as a service, in Proc. of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), March 6-10, 2011, Los Angeles, California, USA.

15.3 National conferences

1. R. Vilalta, On the Resource Abstraction, Partitioning and Composition for Virtual GMPLS Controlled Multi-Layer Optical Networks, in Proc. of Barcelona Forum on PhD. Research in Information and Communication Technologies, October 15th, 2012, Barcelona (Spain).

15.4 Collaborations

1. R. Casellas, R. Muñoz, R. Martínez, R. Vilalta, Applications and Status of Path Computation Elements (PCE) (accepted), Journal of Optical Communications and Networking, 2013.
2. R. Casellas, R. Martínez, R. Muñoz, R. Vilalta, L. Liu, T. Tsuritani, I. Mortita, Control and Management of Flexi-Grid Optical Networks with an Integrated Stateful PCE and OpenFlow controller, Journal of Optical Communications and Networking, 2013.
3. R. Muñoz, R. Vilalta, R. Casellas, R. Martínez, S. Frigerio, A. Lometti, Design and Experimental Evaluation of Dynamic Inverse-Multiplexing Provisioning in GMPLS-controlled Flexi-Grid DWDM Networks with Sliceable OTN BVTs, in Proc. of 39th European Conference and Exhibition on Optical Communication (ECOC 2013), September 22-26 2013, London, U.K.
4. R. Muñoz, R. Casellas, R. Martínez, R. Vilalta, Control Plane Solutions for Dynamic and Adaptive Flexi-grid Optical networks (invited), in Proc. of 39th European Conference and Exhibition on Optical Communication (ECOC 2013), September 22-26 2013, London, U.K.
5. R. Martínez, R. Casellas, R. Muñoz, R. Vilalta, Experimental Evaluation of Delay-Sensitive Traffic Routing in a Multi-Layer Aggregation Network for Fixed Mobile Convergence, in Proc. of 39th European Conference and Exhibition on Optical Communication (ECOC 2013), September 22-26 2013, London, U.K.
6. R. Martínez, A. Castro, R. Casellas, R. Muñoz, L. Velasco, R. Vilalta, J. Comellas, Experimental Validation of Dynamic Restoration in GMPLS-controlled Multi-layer Networks using PCE-based Global Concurrent Optimization, accepted at Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), March 17-21, 2013, Anaheim, California, USA.
7. M. Channegowda, R. Nejabati, M. Rashidfard, S. Peng, N. Amaya, G. Zervas, D. Simeonidou, Ricard Vilalta, Ramon Casellas, Ricardo Martínez, Raül Muñoz, Lei Liu, Takehiro Tsuritani, Itsuro Morita, Achim Autenrieth, Jörg-Peter Elbers, Pawel Kostecki, Pawel Kaczmarek First Demonstration of an OpenFlow based Software-Defined Network Employing Packet, Fixed and Flexible DWDM Grid Technologies on an International Multi-Domain Testbed, Optics Express, Vol. 21, No. 5, Pages 5487-5498, February 2013.
8. M. Channegowda, R. Nejabati, M. Rashidfard, S. Peng, N. Amaya, G. Zervas, D. Simeonidou, Ricard Vilalta, Ramon Casellas, Ricardo Martínez, Raül Muñoz, Lei Liu, Takehiro Tsuritani, Itsuro Morita, Achim Autenrieth, Jörg-Peter Elbers, Pawel Kostecki, Pawel Kaczmarek, First Demonstration of an OpenFlow based Software-Defined Network Employing Packet, Fixed and Flexible DWDM Grid Technologies on an International Multi-Domain Testbed (Post-deadline Paper) in Proc. of 38th European Conference and Exhibition on Optical Communication (ECOC 2012), September 16-20 2012, Amsterdam, The Netherlands.
9. L. Liu, D. Zhang, T. Tsuritani, R. Vilalta, R. Casellas, L. Hong, I. Morita, H. Guo, J. Wu, R. Martínez, R. Muñoz, Field Trial of an OpenFlow-based Unified Control Plane for Multi-layer Multi-granularity Optical Switching Networks, accepted for publication in IEEE/OSA Journal of Lightwave Technology (JLT), 2012.
10. F. J. Vílchez, R. Muñoz, R. Casellas, R. Martínez, R. Vilalta, Design, Deployment and Experimental Assessment of All-Optical Wavelength Conversion in the GMPLS-controlled ADRENALINE Testbed, in Proc. of 17th European Conference on Network and Optical Communications (NOC 2012), June 20-22, 2012, Vilanova i la Geltru, Spain.

11. R. Muñoz, R. Casellas, R. Martínez, R. Vilalta, A Research and Experimentation Overview on Future Optical Network Control Plane in the ADRENALINE Testbed, (Invited Paper) in Proc. of 16th International Conference on Optical Networking Design and Modeling (ONDM), April 17-20, 2012, Colchester, UK.
12. L. Liu, D. Zhang, T. Tsuritani, R. Vilalta, R. Casellas, L. Hong, I. Morita, H. Guo, J. Wu, R. Martínez, R. Muñoz, First Field Trial of an OpenFlow-based Unified Control Plane for Multilayer Multi-granularity Optical Networks, in Proc. of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC) (Post-Deadline Paper), March 4-8, 2012, Los Angeles, California, USA.
13. R. Muñoz, R. Casellas, R. Martínez, R. Vilalta, Experimental demonstration of a hybrid operated MPLS-TP transport network for dynamic Ethernet service provisioning and restoration, in Proc. of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), March 4-8, 2012, Los Angeles, California, USA.
14. R. Martínez, R. Casellas, R. Muñoz, R. Vilalta, Requirements and Enhancements for the Evolution of the GMPLS Control Plane of the ADRENALINE testbed to support Multi-Layer (MPLS-TP/WSON) Capabilities, (Invited Paper) in Proc. of 13th International Conference on Transparent Optical Networks (ICTON 2011), June 26-30, 2011, Stockholm, Sweden.

Chapter 16

International, european and national R&D projects

Some experimental results obtained in this PhD thesis are used in several National, European and International R&D projects in optical networking. This chapter introduces these R&D projects and how this PhD thesis is related with them.

16.1 International R&D projects

16.1.1 STRAUSS - Scalable and efficient orchestration of ethernet services using software-defined and flexible optical networks

16.1.1.1 Project abstract

The STRAUSS project aims to define a highly efficient and global (multi-domain) optical infrastructure for Ethernet transport, covering heterogeneous transport and network control plane technologies, enabling an Ethernet ecosystem. It will design, implement and evaluate, via large-scale demonstrations, an advanced optical Ethernet transport architecture. The proposed architecture leverages on software defined networking principles, on optical network virtualization as well as on flexible optical circuit and packet switching technologies beyond 100 Gbps.

In particular, the STRAUSS project focuses on the integration and development of a) cost/energy efficient and extremely fast-performing switching nodes, based on variable-capacity and fixed-length optical packet switching technology for access and aggregation networks, and on flexi-grid DWDM optical circuit switching technology for long haul transport; b) highly integrated and scalable software defined optical transceivers supporting bandwidth variable multi-flows for flexible Ethernet transmission; c) a virtualization layer for dynamic and on-demand partitioning of the optical infrastructure offering virtual optical Ethernet transport networks (slices); d) legacy (e.g. GMPLS) and new (e.g. OpenFlow based) control plane approaches for control and management of virtual slices and finally e) a service and network orchestration layer for the interworking and coordination of heterogeneous control plane and transport technologies to offer end-to-end Ethernet transport services.

Outcomes of this project will be experimentally validated by means of demonstrations on large scale testbeds in EU & Japan. STRAUSS will provide technological roadmaps, technical approaches and deployment strategies aiming at shortening innovation and exploitation cycles in the area of future optical Ethernet transport networks for both academia and industry in EU and Japan.

16.1.1.2 PhD Thesis relationship

The work performed in this PhD Thesis has resulted in the design of an architecture for providing a virtualization layer for dynamic and on-demand partitioning of the optical infrastructure offering virtual optical Ethernet transport networks (slices).

Chapter 13 provides an orchestration architecture for providing virtual MPLS-TP networks over a shared WSON infrastructure. Chapter 14 presents the resulting virtualization orchestration architecture and provides experimental evaluation of the proposed architecture.

16.2 European R&D projects

16.2.1 STRONGEST - Scalable Tunable and Resilient Optical Networks Guaranteeing Extremely-high Speed Transport

16.2.1.1 Project abstract

STRONGEST leverages on the definition of innovative architectures for developing a scalable, resilient and cost-effective transport network, offering ultra-high capacity to the end users in the broadband society of the future. The new architectures will take into account the evolution of the access network technologies, in order to ensure transparent core-access integration, but the studies carried out by the project will focus mainly on the metro and core areas, because these are the part of the network where the main scalability issues are foreseen in the next years.

The main objective of STRONGEST is to design and demonstrate an evolutionary ultra-high (Petabit) capacity multilayer transport network, compatible with Gbit/s access rates, based on optimized integration of Optical and Packet nodes, and equipped with a multi-domain, multi-technology control plane. This network will be able to offer:

- High scalability and flexibility
- Guaranteed end-to-end performance and survivability
- Increased energy efficiency
- Reduced total cost of ownership

16.2.1.2 PhD Thesis relationship

The designed and developed GMPLS-controlled MPLS-TP node with 10Gbps tunable transponders is the basis of the described work in Chapter 12.

16.2.2 IDEALIST - Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Networks

16.2.2.1 Project abstract

Traffic demand is increasing dramatically, year on year, with typical growth figures of up to 60% for Internet based traffic. Such traffic increase is impacting on both network costs and power consumption. Moreover, traffic is not only increasing but becoming much more dynamic, both in time and direction. For these reasons, transport network evolution from current DWDM systems towards elastic optical networks, based on flexgrid transmission and switching technologies, could significantly increase both transport network scalability and flexibility. Further benefits come from multilayer interworking mechanisms enabling electronic switching technologies (IP/MPLS, OTN, etc) to directly control the bandwidth of the Bandwidth Variable Transponders (BVT) for optical bandwidth optimization purposes. This then defines the key objective behind IDEALIST: To research in detail a cost and power efficient transport network architecture able to carry a wide range of signal bandwidths, each of which will be varying in real time in direction and magnitude, and some of which will be extremely large and possibly exceeding 1Tb/s. The network architecture proposed by IDEALIST is based on four technical pillars:

- Transport systems enabling flexible transmission and switching beyond 400Gbps per channel.
- Control plane architecture for multilayer and multidomain elastic optical networks.
- Dynamic network resources allocation at both IP and elastic optical layers.
- Multilayer network optimization tools enabling both off-line planning and on-line network re-optimization in elastic optical networks.

The intention is that the IDEALIST network architecture will be easily industrialised. Therefore, feasibility studies and experimental implementation and demonstration of prototypes will be key activities, as well. IDEALIST will also feed the collaboration with other Projects and the submission of contributions to ITU-T, OIF, IETF, thus reinforcing European position in standardization bodies.

16.2.2.2 PhD Thesis relationship

IDEALIST project does not include optical network virtualization among its main objectives. The work performed in a GMPLS/PCE Control Plane for Elastic Optical Networks has been introduced in the Virtual Elastic Optical Network architecture in Chapter 10.

16.2.3 OFELIA - OpenFlow in Europe: Linking Infrastructure and Applications

16.2.3.1 Project abstract

The project OFELIA (http://www.fp7-ofelia.eu) creates a unique experimental facility that allows researchers to not only experiment ”on” a test network but to control and extend the network itself precisely and dynamically. The OFELIA facility is based on OpenFlow, a currently emerging networking technology that allows virtualization and control of the network environment through secure and standardized interfaces The CTTC Optical Networking Area has entered the OFELIA consortium via the second open call, proposing a stand-alone project on the topic of the integration of PCE and OpenFlow : The project aims at functionally enhance the OpenFlow control plane, by introducing new path computation architectures and related new network operating system applications. By formally decoupling the path computation function involved in the provisioning of connectivity services, the architecture will be able to offload computations to (optionally remote) dedicated servers and to perform load balancing, distributing path computation tasks amongst a pool of such servers. The main objective is to integrate the notion of (hierarchical) PCEs in the OpenFlow centralized control model. Special focus will be given at the use of the OpenFlow architecture and protocols for the control of Wavelength Switched Optical Networks and the newly deployed Elastic Optical Networks based on flexi-grid.

16.2.3.2 PhD Thesis relationship

Although this PhD Thesis does not observe Optical Network Virtualization by use of OpenFlow Optical FlowVisor, the work performed in the framework of the Optical Task Force in Ofelia project has led to a deep knowledge of optical openflow and has been incorporated in Chapter 14, which deals with Virtualization Orchestration in Heterogeneous Technology domains.

16.3 National R&D projects

16.3.1 DORADO - Design and evaluation of intelligent optical networks based on hybrid nodes with Ethernet and wavelength switching

16.3.1.1 Project abstract

The aim of the DORADO project is to investigate the Wavelength Switched Optical Networks (WSON), Connectio-oriented Ethernet (COE), and Generalized Multiprotocol Label Switching (GMPLS) technologies as the basis for cost-efficient, high-speed, flexible and dynamic optical transport networks. This will be done from an integrated approach that will include technological aspects of modelling, design, performance evaluation and experimentation in the ADRENALINE Testbed of both optical communication systems and optical networks. The DORADO project is the evolution of the previous fundamental-research project RESPLANDOR (TEC-2006-12910/TCM) in which some research aspects of optical networks applied to GMPLS-enabled WSON infrastructures were addressed. The following list summarizes the main scientific and technical objectives that will be addressed:

- Design and evaluation of spectrally-efficient transmission systems based on optical OFDM. Investigation of all-optical OFDM transmission systems to overcome the limitations induced by electronic processing.

- Design and evaluation of advanced optical performance monitoring techniques for dynamically reconfigurable transparent optical networks based on RF spectrum analysis and asynchronous sampling.
- Centralized and distributed RWA algorithms and reservation protocols along with the required extensions to the current GMPLS RSVP-TE, OSPF-TE, and PCEP protocols for provisioning/recovery of optical services in WSON.
- Grooming and aggregation algorithms, along with the required protocol extensions, for the efficient provisioning/recovery of connection-oriented Ethernet services over WSON.
- Development of an experimentation platform to study the performance of GMPLS-enabled connection-oriented Ethernet transport over WSON with all optical wavelength converters and tuneable dispersion compensator modules.

16.3.1.2 PhD Thesis relationship

The design and evaluation of Virtual Optical Networks Resource Allocation algorithms introduced in DORADO project has been addressed in Chapters 8, 11 and 9.

16.3.2 FARO - Integrating control and transmission technologies for flexible, elastic and spectrum-efficient optical networks

16.3.2.1 Project abstract

Global IP traffic has increased eightfold over the past five years, and will increase fourfold over the next five due to the accelerated proliferation of users, not only with broadband Internet access at home (e.g., VDSL or FTTH) but also with mobile broadband (smartphones and tablets). Moreover, the traffic behaviour is suffering a complete change in its pattern. For the first time, Internet video will reach the 50% of total consumer IP traffic by the end of the current year (2012) and 90% by end 2015 (end of this project). The main implication is that the Internet traffic is evolving from a relatively steady and symmetric stream profile (dominant peer-to-peer/file sharing traffic) to a more dynamic traffic with a highly asymmetric profile. Therefore, future optical transport networks will be required to provide high, flexible, elastic and adaptive bandwidth connectivity between IP routers, based on the traffic demands and network conditions. Current ITU-T DWDM wavelength grid with fixed channel spacing is particularly inefficient for large granularities (e.g., 10 Gb/s and 100 Gb/s) since a whole wavelength is assigned to a lower rate optical path (e.g.,10 Gb/s) that does not fill the entire wavelength capacity. This situation has traditionally forced to aggregate and groom low-bit-rate data flows with electrical TDM crossconnects (e.g. SONET/SDH) or, more recently, through electrical packet switching technology (e.g., MPLS), since no alternative in the optical domain (e.g. all-optical packet switching) was commercially feasible. However, a new architecture introducing elasticity and adaptation in optical networks has been recently proposed to provide flexible, highly-efficient and adaptive optical spectrum management, attaining granular grooming in the optical domain. In elastic optical networks (i.e., with variable channel spacing), the optical spectrum is partitioned into basic fixed-size spectrum slots (e.g., 6.25 GHz or 12.5 GHz). The required spectral resources are dynamically and adaptively allocated by assigning the necessary number of contiguous basic fixed-size spectrum slots according to the traffic demand and the network conditions.

The key enablers of elastic optical networks are the Optical OFDM (O-OFDM) technology and the GMPLS/PCE control plane. O-OFDM is the most promising technology for the design of bandwidth-variable transponders, since it enables software-defined optical transmission (SDOT). That is, the transponder can be adapted to multiple modulation formats or variable bandwidth occupancy by means of electronic digital signal processing. Moreover, the transmission of multiple orthogonal subcarriers provides a high spectral efficiency (subcarriers are overlapped), unique flexibility and adaptive bit-rate/bandwidth (modifying number of subcarriers and modulation formats), and sub/super-wavelength granularity. A GMPLS control plane with PCE allows the automatic provisioning and recovery of flexible connections and its bandwidth modification / re-routing in real time in elastic and adaptive optical networks. The integration of these technologies is the basis to support scalable (beyond 100 Gb/s) and large data rate granularities in an energy and cost-effective manner. However, this new network architecture arises new technical challenges that must be properly investigated at both transmission and networking level by means of not only theoretical, analytical and simulation studies but also from an experimental research to capture the whole set of interdependencies between system components. The aim of the FARO project (Integrating control and transmission technologies for flexible, elastic and spectrum-efficient optical networks) is to investigate, from an integrated approach, O-OFDM transmission and GMPLS/PCE control technologies as the key enablers for elastic optical networks, leveraging the bandwidth/bit rate variable optical transponders and an advanced control plane, for the dynamic provisioning of adaptive connectivity services with recovery, in both single and multi-domain contexts. This macroscopic objective includes the following scientific and technological ones:

1. Identification of a concrete list of functional requirements, use cases and scenarios that drive the main scientific and technological work within the FARO project.
2. Design and implementation of SDOT systems based on tuneable, bandwidth and bit rate variable O-OFDM transponders.
3. Development of resource-efficient techniques for O-OFDM transmission enhancement, including distortionless PAPR mitigation and performance monitoring with reduced overhead.
4. New techniques, algorithms and procedures advancing the state of art of multilayer/RSA algorithms with electrical/optical grooming, the subsequent network optimization, and their dissemination.
5. Implementation and assessment of the applicability of a GMPLS/PCE control plane for multilayer optical networks by means of its deployment, along with the numerical (quantitative) evaluation of its performance.
6. Conception of a virtual elastic optical network resource broker and service composition manager.
7. Experimental validation and evaluation of the integrated elastic optical network prototype.

16.3.2.2 PhD Thesis relationship

The FARO project addresses the conception of a virtual elastic optical network resource broker and service composition manager. The work performed in this project has been addressed in Chapter 10.

Part VI

Conclusions and Future Work

Chapter 17

Conclusions and future work

17.1 Conclusions

This PhD Thesis has disserted about virtual optical networking, which supports dynamic provisioning of dedicated networks over the same network infrastructure. VONs are receiving a lot of attention by network infrastructure providers, due to the fact that they can support the stringent network requirements of the emerging high bandwidth and dynamic applications such as high-definition video streaming (e.g., telepresence, television, remote surgery, etc.), and cloud computing (e.g., real-time data backup, remote desktop, etc.). The dynamic deployment of dedicated infrastructure is known as IaaS. In the proposed view, network service providers can request, on a per-need basis, a dedicated VON for each application and have full control over it.

This PhD Thesis has introduced optical network virtualization technologies which allow the partitioning and composition of the network infrastructure (i.e., physical optical nodes and links) into independent virtual resources, adopting the same functionality as the physical resource. The composition of these virtual resources (i.e. virtual optical nodes and links) allows the deployment of multiple VONs. As has been detailed, a VON must be composed of not only a virtual transport plane but also of a virtual control plane, with the purpose of providing the required independent and full control functionalities (i.e., optical connection provisioning, traffic engineering, protection/restauration, etc.).

As previously detailed, the results of this PhD thesis have been disseminated through several scientific publications (3 journals, 9 international conferences, 1 national, and 14 collaborations) and international (STRAUSS), european (IDEALIST, STRONGEST, OFELIA) and national R&D projects (FARO, DORADO).

The first objective of this PhD Thesis consisted on the introduction of the concept of a resource broker for dynamic WSON infrastructure services, whose task is to dynamically deploy VONs from service provider requests and serves as interface between service and infrastructure providers. The introduction of a VON resource broker implies resource management and allocation algorithms for optimal usage of the physical infrastructure, which have been proposed in this PhD Thesis. The introduction of optical network virtualization for EONs has also been considered in this objective.

Firstly, a network virtualization architecture has been presented for deploying GMPLS-controlled WSON networks as a service. Experimental evaluation of the deployed virtual networks has been carried out in the ADRENALINE Testbed, assessing the feasibility of the proposed architecture with a low impact on the VON performance in terms of connection provisioning and recovery delay.

Later, a VON resource broker has been proposed as an interface between service providers and infrastructure providers to deploy virtual GMPLS-controlled WSON infrastructure services. Experimental evaluation carried out in the ADRENALINE testbed has shown the feasibility of deploying independent instances of virtual GMPLS-controlled WSON, providing low delays for VON setup and tear down. The SPUB-FF algorithm has been presented and evaluated its performance in terms of blocking rate of VON requests and VON setup delay.

Four different algorithms for VON resource allocation have been presented (i.e., TWA-FF, TWA-RND, WAWC-FF, WAWC-RND). These algorithms have been analyzed by obtaining the blocking rate of VON requests on three proposed scenarios (i.e., European, NSF, full-meshed network topologies). The proposed algorithms are able to allocate VON request which are connected subgraph of the underlying topology, meaning that each virtual link is directly mapped to a physical link. The proposed algorithms take into account the WCC and the usage of WC to satisfy it. The obtained results show that in the presence of WCs in the physical optical nodes, and the usage of an algorithm that takes them into account, really improves the blocking rate of VON requests.

To overcome the limitation of VON request being a connected subgraph of the underlying network topology, several VON RA algorithms (i.e., SPUB-FF, SPCWA-FF, SPCWA-RND and WTMHSP) have been proposed and compared in terms of VON blocking rate and VON setup delay. These algorithms have been experimentally evaluated in a system architecture including the proposed VON resource broker with VNTM, which deploys virtual GMPLS-controlled WSON infrastructure services over the same physical optical network infrastructure. The obtained results show how the proposed WTMHSP algorithm does lower the blocking rate of VON requests in the analyzed NSFNet scenario, while not increasing the complexity of the VON resource broker, minimizing the total number of hops (a virtual link can span more than one hop in the physical infrastructure) for the requested virtual optical links (GCO).

The VON resource broker has been expanded to deploy virtual GMPLS-controlled EON topologies by establishing elastic optical connections over a GMPLS-controlled physical optical network infrastructure and offering them as virtual TE links. Experimental evaluation carried out in the ADRENALINE Testbed has shown the feasibility of deploying independent instances of virtual GMPLS-controlled EON, providing the blocking rate of VEON requests.

Finally, it has been studied the measurement of the impact of deploying several virtual VON GMPLS control planes on its performance (i.e. setup delay and blocking probability). The virtualization of a GMPLS control plane has resulted in higher setup delay and higher blocking probability in lightpath provisioning. The obtained results have showed the trade off between the diminution of blocking rate of VON requests, through the provision of more virtual control resources for the virtual control planes, and the deterioration of the virtual control plane performance.

The second objective of this PhD Thesis was to design a system architecture and deploy virtual MPLS-TP networks provided over a shared WSON. With this purpose, this PhD Thesis has also focused on the development of an MPLS-TP node which has been deployed for the proposed virtual MPLS-TP network over a shared WSON.

Firstly, the architecture of a GMPLS-enabled MPLS-TP/PWE3 node with 10Gbps tunable DWDM transceivers has been presented and the performance of the implemented software MPLS-TP/PWE3 label forwarding engine has been evaluated in two representative multi-layer scenarios with different traffic grooming strategies.

Later, virtual MPLS-TP networks over a shared WSON have been proposed. Each virtual MPLS-TP network can provide the required flexibility for each emerging cloud application. Experimental assessment carried out in the ADRENALINE Testbed has shown the feasibility of deploying independent instances of GMPLS control plane for each deployed virtual MPLS-TP network under dynamic virtual network requests, providing low setup delays.

The third objective was the composition of multiple virtual optical network with heterogeneous control domains. The concept of a multi-domain resource broker has been introduced and a multi-domain optical network virtualization architecture has been validated.

17.2 Future work

Although the main proposed objectives have been dealt with, this PhD Thesis is a mirror of the continuous work performed by research institutions in optical network virtualization. As it is an on-going research, future results are expected in the proposed topics.

In the objective of virtual GMPLS-controlled optical networks further research in resource allocation algorithms is expected, focusing on EONs. One of the main drawbacks of EONs is the fragmentation of the optical spectrum and usually requires a replanning tool which re-allocates the requested optical connections. VEON resource allocation algorithms will require to overcome such limitations.

Further research will include in VON resource allocation algorithms not only the allocation of network infrastructure, but also of other types of physical infrastructure, such as Computation or Storage Resources, offering end-to-end provisioning DataCenters as a Service (DCaaS).

Novel architectures for Data Center interconnection are expected. Although MPLS-TP networks over WSON are being deployed, Data Center interconnections are expected to be controlled with a Software Defined Networking approach. This may include the requirement of OpenFlow agents at Data Center networking, and virtual network orchestration.

Software Defined Networking has became an enabler for Network Virtualization, since separation of data and control planes makes it easier to manage a logical network topology. The fast introduction of open SDN-related projects such as OpenDaylight controller will result in faster R&D in how networks are virtualized.

Bibliography

[1] M. Jinno *et al.*, "IP Traffic Offloading to Elastic Optical Layer Using Multi-flow Optical Transponder," in *Proc. ECOC 2011, Mo.2.K.*

[2] T. Takeda, "Network Design and Control for Router Offloading Networks," in *iPOP 2012, 1-1.*

[3] R. Ramaswami, K. N. Sivarajan, and G. H. Sasaki, *Optical Networks. A Practical Perspective.* 2010.

[4] ITU-T, "Architecture of optical transport networks," tech. rep., ITU-T Recommendation G.872, 2012.

[5] W. S. Hu and Q. J. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photonics Technology Letters*, vol. 10, no. 7, pp. 970–972, 1998.

[6] P. Roorda and B. Collings, "Evolution to colorless and directionless roadm architectures," in *Proc. of OFC/NFOEC*, 2008.

[7] B. Niven-Jenkins, D. Brungard, M. Betts, N. Sprecher, and S. Ueno, "Requirements of an MPLS Transport Profile," tech. rep., IETF RFC 5654, September 2009.

[8] L. Andersson, L. Berger, L. Fang, N. Bitar, and E. Gray, "MPLS Transport Profile (MPLS-TP) Control Plane Framework," tech. rep., IETF RFC 6373, Sept. 2011.

[9] ITU-T, "Spectral grids for WDM applications: DWDM frequency grid," tech. rep., ITU-T Recommendation G.694.1, 2012.

[10] R. Muñoz, R. Casellas, R. Martinez, L. Liu, T. Tsuritani, and I. Morita, "Experimental Evaluation of Efficient Routing and Distributed Spectrum Allocation Algorithms for GMPLS Elastic Networks," in *in Proc. of ECOC 2012.*

[11] R. Casellas, R. Martínez, R. Muñoz, L. Liu, T. Tsuritani, and I. Morita, "An integrated stateful pce / openflow controller for the control and management of flexi-grid optical networks," in *Proc. of OFC/NFOEC*, 2013.

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

[13] E. Mannie *Ed.*, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," tech. rep., IETF RFC 3945, Oct. 2004.

[14] N. McKeown *et al.*, "Openflow: Enabling innovation in campus networks," *ACM Communications Review*, 2008.

[15] ITU-T, "Generic functional architecture of transport networks," tech. rep., ITU-T Recommendation G.805, 2000.

[16] J. Lang *Ed.*, "Link Management Protocol (LMP)," tech. rep., IETF RFC 4204, Oct. 2005.

[17] Berger, L. and Bryskin, I. and Papadimitriou, D. and Farrel, A., "RSVP-TE extensions in support of end-to-end generalized multi-protocol label switching (GMPLS) recovery," tech. rep., IETF RFC 4873, May 2008.

[18] J. Moy *Ed.*, "OSPF Version 2," tech. rep., IETF RFC 2328, April 1998.

[19] A. Farrel, J.-P. Vasseur, and J. Ash, "A Path Computation Element (PCE)-Based Architecture," tech. rep., IETF RFC 4655, Aug. 2006.

[20] R. Casellas, R. Martinez, R. Muñoz, L. Liu, T. Tsuritani, I. Morita, and M. Tsurusawa, "Dynamic virtual link mesh topology aggregation in multi-domain translucent WSON with hierarchical-PCE," *Optics Express Journal*, vol. 19, no. 26, 2011.

[21] Y. Lee, G. Bernstein, and W. Imajuku, "RFC 6163: Framework for GMPLS and PCE Control of wavelength switched optical networks (WSON)," April 2011.

[22] R. Martınez, R. Casellas, R. Muñoz, and A. Bukva, "Evolution of the GMPLS Control Plane of the ADRENALINE testbed: Overview towards Multi-Layer (Ethernet/WSON) Capabilities," in *in Proceedings of RT-Multilayer Workshop)*, June 2010.

[23] R. Muñoz, R. Vilalta, R. Casellas, R. Martínez, and F. Vílchez, "First lab trial of an integrated mpls-tp/wson transport network controlled by a unified gmpls control plane in the adrenaline testbed (invited paper)," in *Proc. of Photonics in Switching (PS)*, 2012.

[24] O. N. Foundation, "Software-defined networking: The new norm for networks." White paper, 2012.

[25] N. Gude, T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown, and S. Shenker, "Nox: towards an operating system for networks," *ACM SIGCOMM Computer Communication Review*, 2008.

[26] D. Erikson, "Floodlight Java based OpenFlow Controller," tech. rep.

[27] "OpenDaylight - An Open Source Community and Meritocracy for Software-Defined Networking," tech. rep., Linux Foundation, 2013.

[28] S. Das, "Extensions to the openflow protocol in support of circuit switching," tech. rep., 2010.

[29] V. Gudla, S. Das, A. Shastri, G. Parulkar, N. McKeown, L. Kazovsky, and S. Yamashita, "Experimental Demonstration of OpenFlow Control of Packet and Circuit Switches," in *in Proc. of OFC 2011.*

[30] M. Channegowda, R. Nejabati, M. Rashidfard, S. Peng, N. Amaya, G. Zervas, D. Simeonidou, R. Vilalta, R. Casellas, R. Martinez, R. Muñoz, L. Liu, T. Tsuritani, I. Morita, A. Autenrieth, J. Elbers, P. Kostecki, and P. Kaczmarek, "First Demonstration of an OpenFlow based Software-Defined Network Employing Packet, Fixed and Flexible DWDM Grid Technologies on an International Multi-Domain Testbed," *Optics Express*, vol. 21, no. 5, pp. 5487–5498, 2013.

[31] L. Liu, R. Muñoz, R. Casellas, T. Tsuritani, R. Martinez, and I. Morita, "OpenSlice: an OpenFlow-based Control Plane for Spectrum Sliced Elastic Optical Path Networks," in *in Proc. of ECOC 2012.*

[32] R. Sherwood, G. Gibby, K. Yapy, G. Appenzellery, M. Casado, N. McKeowny, and G. Parulkar, "FlowVisor: A Network Virtualization Layer," tech. rep.

[33] S. Azodolmolky, R. Nejabati, S. Peng, A. Hammad, M. Channegowda, N. Efstathiou, A. Autenrieth, P. Kaczmarek, and D. Simeonidou, "Optical flowvisor: An openflow-based optical network virtualization approach," in *Proc. of OFC/NFOEC*, 2012.

[34] N. Feamster, L. Gao, and J. Rexford, "How to lease the internet in your spare time," *SIGCOMM Comput. Commun. Rev.*, vol. 37, pp. 61–64, January 2007.

[35] Y. Wang, Y. Jin, W. Guo, W. Sun, and W. Hu, "Virtualized optical network services across multiple domains for grid applications," *Communications Magazine, IEEE*, vol. 49, pp. 92 –101, may 2011.

[36] N. M. K. Chowdhury and R. Boutaba, "A survey of network virtualization," *Computer Networks*, vol. 54, no. 5, 2010.

[37] W. Wei *et al.*, "Optical orthogonal frequency division multiple access networking for the future internet," *Journal of Optical Communication Networks, Vol. 1, No. 2*, vol. 1, no. 2, 2009.

[38] R. Nejabati, E. Escalona, S. Peng, and D. Simeonidou, "Optical network virtualization," in *Optical Network Design and Modeling (ONDM), 2011 15th International Conference on*, pp. 1 –5, feb. 2011.

[39] A. Farrel and I. Bryskin, *GMPLS. Architecture and Applications.* 2006.

[40] E. Grasa, G. Junyent, S. Figuerola, A. Lopez, and M. Savoie, "Uclpv2: A network virtualization framework built on web services," *IEEE Communications Magazine*, vol. 46, no. 3, 2008.

[41] J. Wu, M. Savoie, S. Campbell, and H. Zhang, "Layer 1 Virtual Private Network Management by Users," *IEEE Communications Magazine*, vol. 44, no. 12, pp. 86–93, 2006.

[42] S. Figuerola *et al. IEEE/OSAJournal of Optical Communications and Networking*, vol. 1, no. 2, pp. A247–A257, 2009.

[43] A. Pages, J. Perelló, S. Spadaro, and G. Junyent, "Strategies for virtual optical network allocation," *IEEE Communications Letters*, vol. 16, no. 2, 2012.

[44] A. Pages, J. Perello, S. Spadaro, J. Garcia-Espin, J. Riera, and S. Figuerola, "Optimal allocation of virtual optical networks for the future internet," in *in Proc. of 16th International Conference on Optical Network Design and Modeling (ONDM), 2012.*

[45] S. Peng *et al.*, "Performance modelling and analysis of dynamic virtual optical network composition," *Proc. of Optical Network Design and Modeling*, 2012.

[46] S. Peng, R. Nejabati, S. Azodolmolky, E. Escalona, and D. Simeonidou, "An impairment-aware virtual optical network composition mechanism for future internet," *Opt. Express*, vol. 19, Dec 2011.

[47] J. Perello, P. Pavon-Marino, and S. Spadaro, "Cost-efficient Virtual Optical Network Embedding for Manageable Inter-Data Center Connectivity," *ETRI Journal*, vol. 35, no. 1, pp. 142–145, 2013.

[48] J. van der Ham *et al.*, “Using the network description language in optical networks,” in *Integrated Network Management, 2007. IM ’07. 10th IFIP/IEEE International Symposium on*, pp. 199 –205, 2007.

[49] S. Peng, R. Nejabati, S. Azodolmolky, E. Escalona, and D. Simeonidou, “An impairment-aware virtual optical network composition mechanism for future internet,” in *37th European Conference and Exposition on Optical Communications*, 2011.

[50] S. Peng, R. Nejabati, M. Channegowda, and D. Simeonidou, “Application-aware and adaptive virtual data centre infrastructure provisioning over elastic optical ofdm networks,” in *Optical Communications (ECOC), 2013 39th European Conference on*, september 2013.

[51] E. Escalona, S. Peng, R. Nejabati, D. Simeonidou, J. A. Garcia-Espin, J. Ferrer, S. Figuerola, *et al.*, “Geysers: A novel architecture for virtualization and co-provisioning of dynamic optical networks and it services,” in *Future Network Summit*, 2011.

[52] S. Azodolmolky, R. Nejabati, E. Escalona, R. Jayakumar, N. Efstathiou, and D. Simeonidou, “Integrated openflow - gmpls control plane: An overlay model for software defined packet over optical networks,” in *Proc. of ECOC*, 2011.

[53] C. Pinart, “Anticipation of Traffic Demands to Guarantee QoS in IP/Optical Networks,” *Future Internet*, vol. 2, no. 3, pp. 417–430, 2010.

[54] F. Galán Márquez and R. Muñoz, “An automatic model-based reconfiguration and monitoring mechanism for flexible gmpls-based optical networking testbeds,” in *Optical Network Design and Modeling*, 2007.

[55] R. Vilalta, R. Muñoz, R. Casellas, and R. Martínez, “Experimental demonstration of a virtual optical network resource broker and compositor for dynamic gmpls wson infrastructure services,” in *OFC/NFOEC*, 2012.

[56] R. Vilalta, R. Muñoz, R. Casellas, and R. Martínez, “Dynamic virtual gmpls-controlled wson using a resource broker with a vnt manager on the adrenaline testbed,” *Proc. of ECOC*, 2012.

[57] R. Muñoz, R. Casellas, R. Martínez, R. Vilalta, J. Vílchez, and J. Vázquez, “Virtualizing adrenaline testbed for deploying dynamic gmpls-controlled wson as a service,” in *OFC/NFOEC*, 2011.

[58] I. Nishioka and D. King, “ Use of the Synchronization VECtor (SVEC) List for Synchronized Dependent Path Computations,” tech. rep., IETF RFC 6007, Sept. 2010.

[59] Y. Lee *et al.*, “Path Computation Element Communication Protocol (PCEP) Requirements and Protocol Extensions in Support of Global Concurrent Optimization,” tech. rep., IETF RFC 5557, Jan. 2009.

[60] R. Casellas, R. Martínez, R. Muñoz, and S. Gunreben, “Enhanced backwards recursive path computation for multi-area wavelength switched optical networks under wavelength continuity constraint,” *J. Opt. Commun. Netw.*, vol. 1, Jul 2009.

[61] R. Casellas, R. Muñoz, *et al.*, “Gmpls/pce control of flexi-grid dwdm optical networks using co-ofdm transmission [invited],” *Journal of Optical Communication Networks*, 2012.

[62] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku, and Y. Takigawa, “Dynamic multilayer routing schemes in GMPLS-based IP+ optical networks,” *Communications Magazine, IEEE*, vol. 43, no. 1, pp. 108–114, 2005.

[63] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. F. Kaashoek, “The click modular router,” *ACM Trans. Comput. Syst.*, vol. 18, pp. 263–297, August 2000.

[64] A. Malis and M. Townsley, “Pseudowire Emulation Edge-to-Edge (PWE3) Fragmentation and Reassembly,” tech. rep., IETF RFC 4385, Aug. 2006.

[65] E. Rosen, A. Viswanathan, and R. Callon, “Multiprotocol Label Switching Architecture,” tech. rep., IETF RFC 4385, Jan. 2001.

[66] S. Bradner and J. McQuaid, “Benchmarking Methodology for Network Interconnect Devices,” tech. rep., IETF RFC 4385, Jan. 1999.

[67] B. Chen and R. Morris, “Flexible control of parallelism in a multiprocessor pc router,” in *Proceedings of the 2001 USENIX Annual Technical Conference (USENIX ’01)*, (Boston, Massachusetts), pp. 333–346, June 2001.

[68] Intel Corporation, *IRQBALANCE - handholding your interrupts for power and performance*, 2006.

[69] A. Bianco, R. Birke, D. Bolognesi, J. Finochietto, G. Galante, M. Mellia, M. Prashant, and F. Neri, “Click vs. linux: two efficient open-source ip network stacks for software routers,” in *High Performance Switching and Routing, 2005. HPSR. 2005 Workshop on*, pp. 18 – 23, May 2005.

[70] P. Pang, “Efficient inter-data center transport within sdn framework,” in *Proc. of iPOP*, 2012.

[71] Cisco, “Cisco global cloud index: Forecast and methodology, 2011-2016,” tech. rep., 2012.

[72] R.Vilalta, R. Muñoz, R. Casellas, and R. Martínez, “Dynamic Deployment of Virtual GMPLS-controlled Elastic Optical Networks Using a Virtual Network Resource Broker on the ADRENALINE Testbed,” in *in Proc. of 15th International Conference on Transparent Optical Networks (ICTON 2013), 23-26 June 2013, Cargagena (Spain).*

[73] H. Gredler, J. Medved, S. Previdi, A. Farrel, and S. Ray, “North-Bound Distribution of Link-State and TE Information using BGP,” tech. rep., IETF draft-ietf-idr-ls-distribution-03, May 2013.