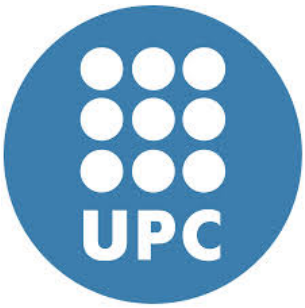
RESILIENCE MECHANISMS FOR CARRIER-GRADE NETWORKS

Universitat Politecnica de Catalunya
Departament d' Arquitectura de Computadors

Thesis presented in Fulfilment of the requirements for the degree of Doctor for the Universitat Politecnica de Catalunya
Research Group: CRAAX



PhD Student: Wilson Ramírez

Advisor: Xavier Masip-Bruin

Co-Advisor: Marcelo Yannuzzi

September, 2014

To my parents and to Ana Ramírez...

Acknowledgements

Es muy difícil, por no decir imposible describir como y cuales personas me han ayudado para finalizar con éxito mi doctorado el cual es el motivo de esta Tesis. Esto de por sí ya sería un documento más extenso que esta misma Tesis. Es por esto que tratare de ser breve. Primeramente debo de agradecer a mi familia: mi padre Dr. Wilson Ramírez, madre Dra. Lourdes Almonte, tíos Danilo, Oscar y Nélcida Almonte, por el apoyo incondicional brindado, en el ámbito moral, económico, así como de cualquier otra índole de la cual pudiera necesitar. Sin duda, a pesar de las distancias físicas, durante estos años su cariño me ha hecho valorar más el valor de una familia. Sin su apoyo este documento de Tesis no tuviera sentido. Más aun, el deber de hacerlos sentir orgullosos fue mi "bandwidth" durante estos últimos años.

Por otro lado, debo (naturalmente) agradecer a mis directores de Tesis, Xavier Masip-Bruin y Marcelo Yannuzzi, así como a debo de agradecer a Víctor López, Eva Marin-Tordera y demás compañeros de laboratorio. Gracias por: 1) Aceptarme en su grupo de Investigación (CRAAX); 2) las conversaciones e interacciones; 3) la guía ofrecida; y, 4) el largo tiempo dedicado en leer mis artículos. Sin ellos no hubiese podido desarrollar mis ideas o a ver mas allá de las palabras las ideas de otros; pero más importante aun saber como presentarlas.

He de no olvidar, dar gracias a los revisores externos asignados a esta tesis así como a (algunos) los revisores anonimos los cuales sus aportaciones y correcciones mejoraron sin duda los resultados mostrados en esta tesis. Finalmente, pero no menos importante, debo de agradecer a Donald Membreño, Luis Sánchez, Arisleidy Mesa, por su apoyo de amor y amistad. Aunque estas personas no estuvieron directamente vinculadas en esta Tesis. Sin su amistad este documento de Tesis no tuviera el mismo grado de calidad. Y es que esta Tesis no representa solo un logro académico sino también una superación y formación personal.

Barcelona, September 2014

Wilson Ramírez

Abstract

In recent years, the advent of new Future Internet (FI) applications is creating ever-demanding requirements. These requirements are pushing network carriers for high transport capacity, energy efficiency, as well as high-availability services with low latency. A widespread practice to provide FI services is the adoption of a multi-layer network model consisting in the use of IP/MPLS and optical technologies such as Wavelength Division Multiplexing (WDM).

Indeed, optical transport technologies are the foundation supporting the current telecommunication network backbones, because of the high transmission bandwidth achieved in fiber optical networks. Traditional optical networks consist of a fixed 50 GHz grid, resulting in a low Optical Spectrum (OS) utilization, specifically with transmission rates above 100 Gbps. Recently, optical networks have been undergoing significant changes with the purpose of providing a flexible grid that can fully exploit the potential of optical networks. This has led to a new network paradigm termed as Elastic Optical Network (EON).

A multi-layer Carrier-Grade Network (CGN) demands scalable and efficient protection schemes with the ability to recover from failures in an agile manner. In light of this, the throughput advantages of EON features have been evaluated aiming at reducing the Protection Cost (P_{cost}).

Recently, a new protection scheme referred to as Network Coding Protection (NCP) has emerged as an innovative solution to proactively enable protection in an agile and efficient manner by means of throughput improvement techniques such as Network Coding (NC). It is an intuitive reasoning that the throughput advantages of NCP might be magnified by means of EON features.

The goal of this thesis is three-fold. The first, is to study the advantages of NCP schemes in planning scenarios. For this purpose, this thesis focuses on the performance of conventional protection schemes such as DP in comparison with NCP assuming both a fixed as well as a flexible spectrum grid. In addition, this thesis studies NCP schemes under the context of a multi-layer network model since this is a widely adopted network model.

However, conversely to planning scenarios, in dynamic scenarios the accuracy of Network State Information (NSI) is crucial since inaccurate NSI might substantially affect the performance (blocking probability and P_{cost}) of an NCP scheme. The accuracy of NSI is highly sensitive to the frequency of its dissemination. The second contribution of this thesis is to study the performance of protection schemes in dynamic scenarios considering inaccurate NSI. For this purpose, this thesis explores prediction techniques in order to mitigate the negative effects of inaccurate NSI.

Abstract

The first two goals of this thesis focus on the study of distributed (source) protection schemes considering realistic network scenarios. These scenarios are based on both single and multi-layer networks with IP/MPLS and Optical technologies, assuming the conventional host-oriented communication model, which is widely deployed in the current Internet and Routing architectures.

On the other hand, Internet users are continuously demanding seamless connectivity with mobility features. These requirements cannot be supported by the current IP-based addressing scheme supporting the whole routing architecture because of its well-known limitations, mainly driven by the unstoppable growth of Internet users and its host-oriented design.

The host-oriented communication model embeds several issues, which are hindering its deployment in future Internet architectures such as the so-called Internet of Things (IoT). Fortunately, there is a new trend in network research referred to as ID/Locator Split Architectures (ILSAs) which is a non-disruptive technique that can be adopted to mitigate the issues related to the current host-oriented communication model. Moreover, a new routing architecture referred to as Path Computation Element (PCE) has emerged as a centralized scheme with the aim of overcoming the well-known issues of distributed routing schemes.

Undoubtedly, routing and protection schemes (including the one proposed in this thesis) need to be enhanced to fully exploit the advantages provided by new network architectures such as ILSAs and PCE schemes. In light of this, the third goal of this thesis introduces a novel PCE-like architecture termed as Context-Aware PCE. In a context-aware PCE scenario, the driver of a path computation is not a host/location, as in conventional PCE architectures, rather it is an interest for a service defined within a context.

Contents

List of Figures

List of Tables

Abbreviations

AS	Autonomous System
ASON	Automatically Switched Optical Networks
ASPL	Average Shortest Path Length
ATM	Asynchronous Transfer Mode
AVND	Average Node Degree
APC	Average Protection Cost
CAPEX	Capital Expenditure
CGE	Carrier Grade Ethernet
CGN	Carrier-Grade Network
CR	Connection Request
DCN	Data Center Networks
DHT	Distributed Hast Table
DNS	Domain Name System
DP	Dedicated Protection
DPP	Dedicated Path Protection
EID	Endpoint Identifier
EON	Elastic Optical Network
ETL	Expected Traffic Loss
FPBR	Finer Prediction Based Routing
FI	Future Internet
FF	First-Fit
GIS	Geographic Information System
GMPLS	General Multi Protocol Label Switching
HSRP	Hot Standby Routing Protocol
HPBR	Hybrid Prediction based Routing
ILSA	ID/Locator Split Architecture
IoT	Internet of Things
ITRs	Ingress Tunnel Routers
LCP	Least-Congested Path
LL	Least Loaded
LCB	LISP Control Box
LRP	LISP Redundancy Protocol
LOC	Locator
2ML	Multi Layer
MLR	Multi-Layer Resilience
MLRBU	Multi-Layer Resilience Bottom-UP
MLRTD	Multi-Layer Resilience Top-Down

MTD	Multi-layer Topology Discovery
MTE	Multi-Layer Traffic Engineering
NC	Network Coding
NCP	Network Coding Protection
NE	Network Element
NMS	Network Management System
NSI	Network State Information
OAM	Operational and Maintenance
OPEX	Operational Expenditures
OEO	Optical-Electrical-Optical
OFDM	Orthogonal Frequency Division Multiplexing
OS	Optical Spectrum
PBR	Prediction Based Routing
PCC	Path Computation Client
PCE	Path Computation Element
PCReq	Path Computation Request
PCRep	Path Computation Reply
PCRUpd	Path Computation Update
P_{cost}	Protection Cost
PHR	Programmable Hybrid Routers
PLI	Physical Layer Impairments
PNCP	Predictive Network Coding Protection
QoS	Quality of Service
RR	Random Wavelength Assignment
RI	Routing Inaccuracy
RWA	Routing and Wavelength Assignment
RLOC	Routing Locator
RSA	Routing Spectrum Assignment Algorithms
SDNs	Software Defined Networking
SDH	Synchronous Digital Hierarchy
SNMP	Simple Network Management Protocol
SLR	Single-Layer Resilience
SLA	Service Level Agreement
SOA	Semiconductor Optical Amplifier
SP	Shared Protection
SONET	Synchronous Optical Network
UMLR	Uncoordinated Multi-Layer Resilience
VNTM	Virtual Network Topology Manager
WCC	Wavelength Continuity Constraint
WDM	Wavelength Division Multiplexing
WRN	Wavelength Routed Network
WR	Wavelength Router
WS	Wavelength-Selective
WSAR	Weighted Selective Adaptive Routing

1 Introduction

This introductory chapter begins by providing a quick overview on the evolution of transport and network technologies in the last 30 years. Then, it continues by discussing the motivations driving this work and the objectives of this thesis. The rest of the chapter concludes with an overview of the structure of this manuscript.

1.1 Evolution of Carrier-Grade Networks

Substantial modifications have been made on the transport and network technologies used by Carrier-Grade Networks (CGNs), mainly driven by the ever-emerging requirements of new applications, such as life video or real time gaming, and services such as cloud computing or network resources virtualization.

The overall evolution of both transport and network technologies is depicted in Fig. 1.1. As it can be observed, from earliest 80's Synchronous Digital Hierarchy (SDH) and Synchronous Optical Network (SONET) have been the standards mostly used in transport networks (SONET and SDH are very similar, but SONET has been mostly used in North America, whereas SDH has been used outside North America), reasoned by the set of key features SONET/SDH brought to packet-based technologies, such as Frame Relay or Asynchronous Transfer Mode (ATM), as follows:

- High transmission rates.
- OAM support.
- Low Recovery time (50 ms).
- Grooming of multiple technologies [1] .

However, despite these features SONET/SDH has several issues that hinder its deployment in current transport networks, such as: (1) Shortest-Path algorithms may not be always used

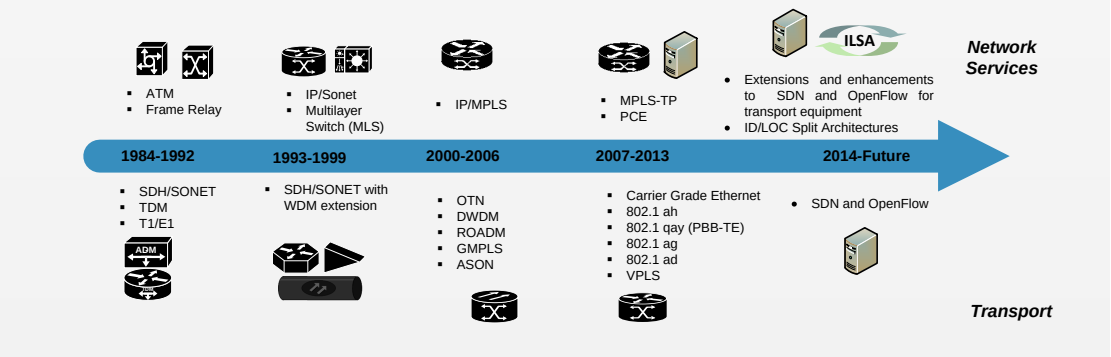


Figure 1.1: Evolution of network and transport layers.

for lightpath provisioning; (2) the protection schemes for SONET/SDH demand the utilization of dedicated links –turning into a Capital Expenditure (CAPEX) increase–, and; (3) SONET/SDH nodes are not capable to set-up or torn-down lightpaths dynamically.

On the other hand, regarding the network layer evolution, service providers started to deploy Frame Relay and ATM technologies in order to offer their services to potential customers. Nevertheless, in early 2000 these two technologies were eventually replaced by IP/MPLS services due to several well-known issues, e.g., routing performance, interoperability or operational costs.

Motivated by the shortcomings of SONET/SDH, in the mid 90's, Wavelength Division Multiplexing (WDM) emerged as the preferred transport technology among network carriers. WDM came up as a suitable solution for using the vast amount of bandwidth provided by optical fiber technologies. As a matter of fact, nowadays research studies claim that the limit of fiber capacity is near 100 Tbps [2]. Such amount of bandwidth is incredibly higher than the bandwidth offered for any other transmission media. In fact, recent optical transmission testbeds show that reaching an optical bandwidth capacity above 1 Tbps is not an utopia [3],[4].

Despite the benefits on the transmission capacity offered by WDM technologies, during the 90's decade, optical nodes had a major limitation related to their ease of configuration. Consequently, great research efforts have been made in optical networking since the early 2000's, in order to endow the optical nodes with enough flexibility to be remotely configured. This effort led to the creation of Reconfigurable Optical Add Drop Multiplexers (ROADMs) that enable the remote lightpaths management.

In order to fully exploit the dynamic operation of ROADMs certain control capacities are required. This motivated the advent of control plane technologies such as Generalized Multi Protocol Label Switching (GMPLS) and Automatically Switched Optical Networks (ASON), both aiming to provide optical equipments with the required functionalities to support a variety of network features such as traffic engineering, recovery capabilities, among others.

Also in the mid 2000's new standards emerged in order to endow transport technologies such as Ethernet with new network features, turning Ethernet into the so-called Carrier Grade Ethernet (CGE)[5]. One of the main building blocks of CGE are the standards IEEE 802.1ad [6], and 802.1ah [7], which increase the level of broadcast segmentation provided by the VLAN stacking employed by the conventional Ethernet. As a result, there is an improvement on the scalability, traffic engineering, and security features. On the other hand, the standards IEEE 802.1ag the IEEE 802.1qay (PBB-TE) endow CGE with resilience capabilities [8], [9].

Also in that decade, IP/MPLS started to become an essential part of the network core of CGNs. This motivated the network community to start focusing on providing MPLS with added-value features such as survivability and maintenance tasks among others. In this regard, the MPLS Transport Profile (MPLS-TP) emerged as an initiative of both the Internet Engineering Task Force (IETF) and the Telecommunication Standardization Sector (ITU-T), to expand MPLS for supporting features that are commonly demanded by transport technologies [10].

In the last years, the control plane of both network and transport layers started to undergo significant changes. New network architectures and third-party systems such as Path Computation Element (PCE), and Software Defined Networking (SDN), were conceived with the aim of offering new network features, such as advanced path computation and customized configuration tasks. These network features are essential to enhance the resilience of CGNs.

Moreover, in recent years there is a trend in network research referred to as service or context-oriented communications. This trend consists in adopting a service/context-oriented communication model instead of the conventional host/location-oriented model. This new network paradigm is motivated by the well-known limitations of the host/location-oriented model, as well as by the poor benefits offered by the diverse set of ad-hoc solutions that have been proposed to address these limitations.

A context-aware communication model comes up as an alternative to the traditional "OSIfied IP networks", raising two conceptual trends in network research: 1) Clean-Slate architectures, that is solutions decoupled from the traditional OSI layered structure (for example adopting a context-aware communication model), and; 2) Non-disruptive approaches, that is solutions "friendly" to the current layered structure (but still offering context-aware communication capabilities) such as ID/LOC Separation Architectures (ILSA) schemes. Both approaches have become the target for numerous research efforts in the recent years. Although some contributions may be found in the literature working in both trends, probably the right (or more commonly used) approach would be the one best meeting the pragmatic aspect of operational networks. In that sense, it is important to notice that network carriers are very reluctant to adopt clean-slate architectures, mainly due to the difficulty of migration tasks, and the potential disruption on the provided services that this migration could drive. Thus, based on this pragmatic feeling, it seems that non-disruptive approaches, such as ILSA schemes, come up to be more appealing (at least easy to deploy) than Clean-Slate architectures.

Table 1.1: Requirements of CGNs.

Requirements
High bandwidth consumption: above 100 Gbits at the aggregation/core level.
Energy Efficiency: minimize OEO conversions.
Low provisioning times: connection lightpath are set-up and torn-down on a short-term basis.
High Performance RWA algorithms: minimize blocking probability while reducing the amount of signaling overhead.
Efficient Protection schemes: minimize both protection cost and recovery time.
Enhanced Network Features: mobility without communication disruption (Full Mobility), TE engineering and resilience.

ILSA schemes deal with both the depletion (exhaustion) of addresses and the semantic overload (double functionality) of addresses problems by assigning an independent set of addresses for identification and location functions.

1.2 Thesis Motivation

As described in Chapter 1.1 both network and transport layers have undergone substantial changes mainly motivated by requirements such as: 1) transport capacity, 2) energy efficiency, 3) high-available low latency services provisioning, and 4) short-term configuration/provisioning tasks, among other requirements, see Table 1.1[11].

For instance, nowadays most of the CGNs are used as commodity for DCs, where traffic is nearly four times global Internet traffic and it is expected to increase 50% in the next two years [12]. This traffic increase expectation is mainly caused by: i) applications requiring massive bandwidth and Quality of Service (QoS) guarantees, such as video content distribution (YouTube, or Netflix), and; ii) applications generating bursty traffic loads, such as big data analytics (e.g., Map Reduce), search (e.g., Google), Social Networking (e.g., Facebook, Twitter), etc. Hence, CGNs must be designed and optimized to deal with huge volumes of highly demanding traffic in a cost/energy efficient way.

Motivated by the need for handling huge volumes of traffic while simultaneously reducing the power consumption, a widespread practice is to adopt an all-optical routing model for the transport layer. Under this model, the transport layer is interconnected by means of WDM fiber-optical links. In a WDM network –also referred to as Wavelength Routed Networks

(WRNs)–, Wavelength Router (WR) nodes are capable of routing traffic in the optical domain without Optical-Electrical-Optical (OEO) conversion at intermediate nodes. To this end, highly efficient Routing and Wavelength Assignment (RWA) algorithms, jointly with control plane technologies capable of provisioning and torn-down lightpaths on a short-term basis such as ASON, GMPLS or SDN, are required.

All-optical routing is highly demanded to reduce both power consumption and communication latency. As a matter of fact, all-optical WRNs are becoming a widespread practice, conversely to electrical and opaque networks, due to the advantages related to both transport capacity and energy efficiency. This is noticeable by the continuously growing deployment of optical commodity switches based on WDM technologies, replacing electrical switches at the core level in DCN scenarios [13].

It is widely accepted and broadly demonstrated (unfortunately even in real cases) that in a CGN using WDM technologies as the transport medium, a link failure might lead to a significant loss of traffic. Therefore, CGNs must be endowed with resilience mechanisms in order to withstand and recover from failures in an agile and efficient manner. Undoubtedly, these resilience mechanisms might potentially leverage the flexibility provided by both the new control plane technologies as well the emergent network paradigms such as PCE and ILSAs.

In planning scenarios, a resilience mechanism is commonly evaluated by its recovery time and P_{cost} (amount of resources allocated to protection), whereas in dynamic scenarios, another criterion to be considered is the blocking probability. The blocking probability of a resilience mechanism is mainly affected by the so-called Wavelength Continuity Constraint (WCC), which states that a lightpath can be solely established if the same wavelength is available on the path selected from the source to the destination WR pair.

In order to meet the WCC constraint, the accuracy of the NSI, particularly the wavelength availability per link, is significantly important. An incorrect selection of wavelengths might increase the amount of blocked connections, because inaccurate (or outdated) NSI will not represent a real “picture” of the current network topology state.

The negative effect added by inaccurate NSI is referred to as the Routing Inaccuracy (RI) problem [14]. The RI problem has been widely studied in unprotected scenarios. It is an intuitive observing that these negative effects are more harmful in protected scenarios where resilient lightpaths are demanded, i.e., two link-disjoint lightpaths must be computed per Connection Request (CR).

In a protected scenario, two major resilience approaches can be adopted: 1) Proactive protection, consisting in the simultaneous allocation of network resources for both primary and backup paths, and; 2) Reactive protection, consisting in the allocation of network resources for a backup path solely when the primary path is affected by a failure [15], [16].

Proactive protection schemes, such as Dedicated Protection (DP), enable protection against link failures in an agile manner, but unfortunately require a vast amount of network resources. Conversely, reactive protection schemes such as Shared Protection (SP) are more efficient in managing the network resources devoted to protection in comparison with proactive protection schemes, but the demanded recovery time is much higher.

From the perspective of a CGN designer, it would be optimal to combine the advantages of both proactive and reactive protection schemes. Driven by this necessity, a novel protection strategy referred to as Network Coding Protection (NCP) has recently emerged as a promising solution offering protection in an agile and cost-efficient manner. NCP strategies leverage the use of throughput improvement techniques such as Network Coding (NC) jointly with a proactive protection scheme. For more information related to NC operation,the reader is referred to the Appendix Overview of Network Coding.

NCP has been widely studied in network research at the network planning phase, where CR demands are known beforehand [17], [18]. However, to the best of our knowledge, there is not any study dealing with NCP under dynamic traffic considering inaccurate NSI. A rationale extending the focus of this thesis is to fill this gap by extensively studying the behavior of NCP in a source routing scenario under dynamic traffic and hence considering inaccurate NSI caused, for example, by a periodic updating policy. This thesis shall show that NCP yields a lower P_{cost} because of its efficient usage of network resources in comparison with conventional proactive protection schemes such as DP. However, it might be more susceptible to inaccurate NSI due to the routing constraints that must be met in order to obtain the benefits of NC.

Moreover, the advent of new network paradigms such as ILSA schemes or PCE is encouraging network researches to explore innovative opportunities to enhance network features such as traffic engineering and resilience. Further investigation is yet required to combine both network paradigms.

1.3 Objectives of this Thesis

The main conceptual objective of this thesis is to provide solutions aiming at optimizing overall network resilience. To this end, the work done and the presented contributions fall into the technical areas of network routing and protection. The main objective is then split into two technical objectives as follows:

1. Technical Objective 1 (TO1): Proposing and validating innovative NCP-based strategies to enhance network resilience considering both planning and dynamic scenarios.

2. Technical Objective 2 (TO2): Evaluating the benefits of new network paradigms such as PCE and context-aware communications to enhance network routing and protection performance.

Activities in TO1 are separated in planning (static) and dynamic scenarios. For planning scenarios, this thesis ends up providing a techno-economic analysis comparing proposed NCP-based solutions versus conventional proactive protection solutions such as DP protection. For this purpose, the scenario evaluated is based on the realistic multi-layer network of the Spanish Backbone topology. Moreover, motivated by the need to reduce network resources consumption, this thesis also evaluates the benefits of NCP in EONs, setting the seed in this flexible-grid scenario for further analysis (to the best our knowledge, this is the first study related toe NCP in a flexible-grid setting). In addition, this thesis proposes an innovative NCP scheme called DPNC+. DPNC+ is devised to reduce the network resources required to enable link protection (P_{cost}) in multi-layer networks, being the first NCP scheme that leverages cross-layer information. Since cross-layer information is very important for the design of multi-layer protection schemes, this thesis provides an algorithm to dynamically discover the cross-layer connections –connection between an IP/MPLS router and a WR–. The main advantage of this algorithm is that it is vendor-agnostic.

Regarding dynamic scenarios, this thesis proposes a dynamic NCP scheme called Predictive Network Coding Protection (PNCP). PNCP is able to reduce the total amount of network resources required to protection (P_{cost}) in comparison with conventional dynamic protection schemes while also mitigating the negative effects of inaccurate NSI. This can be achieved by combining prediction techniques (successfully evaluated in unprotected scenarios) with the throughput advantages of NC [19].

Finally, in TO2, this thesis aims at evaluating new routing paradigms such as centralized routing architectures, e.g., PCE and new communication models, and ILSA schemes. In light of this, this thesis proposes a new PCE scheme so-called context-aware PCE. A context-aware PCE leverages ILSA schemes in order to enhance network features such as traffic engineering and resilience.

1.4 Thesis Structure

The rest of thesis is organized in four chapters. Chapter II and Chapter III are devoted to the study of routing and resilience in single-layer and multi-layer CGN networks respectively. Then, Chapter IV presents the future challenges and trends related to routing and resilience. Finally, Chapter V concludes this thesis. In the following paragraphs the organization of this thesis is described in more detail.

Chapter II.

Section 2.1 and 2.2. These two sections plunge into the issues affecting the routing and wavelength assignment problems.

Section 2.3. In this section, distinct types of protection schemes are introduced for Optical networks considering fixed and a flexible spectrum grid. In particular, this section emphasizes

on the benefits of NCP. For this purpose, a novel technique of NCP called multiple-coding is described. NCP enhanced with multiple coding is evaluated against other proactive protection schemes in terms of P_{cost}, as well as both Capital Expenditures (CAPEX) and Operational Expenditures (OPEX). Moreover, this section distills the technical issues concerning the deployment of an NCP scheme

Section 2.4. This section discusses the RWA problem in online scenarios considering the negative effects of inaccurate NSI on the RWA algorithms performance.To this end, this section proposes two novel schemes, namely Hybrid Prediction based Routing (HPBR) and Finer Prediction Based Routing (FPBR), devised to mitigate the negative effects of inaccurate NSI in unprotected scenarios. Both HPBR and FPBR leverages prediction techniques with the aim of improving routing performance while reducing the amount of signaling required. This section extends the study related to the negative effects of inaccurate NSI to protected scenarios. Moreover, this section also introduces an innovative scheme namely PNCP. The novelty of PNCP is that it combines the advantages of both NCP and Predictive routing with the aim of reducing the P_{cost} as well as the negative impact of inaccurate NSI in terms of blocking.

Chapter III.

Sections 3.1 and 3.2. These two sections present in a nutshell distinct recovery schemes for multi-layer CGNs, as well as the challenges for managing resilience in multi-layer CGNs.

Sections 3.3. This section presents a protection scheme namely DPNC+ devised for multi-layer networks is proposed. DPNC+ is an NCP-based scheme that leverages cross-layer information.

Section 3.4. This section provides an insightful discussion of how cross-layer information is useful for managing resilience in multi-layer CGNs. Moreover, an algorithm facilitating a dynamic discovery of the multi-layer network topology is proposed.

Chapter IV.

Sections 4.1 and 4.2. In these two sections the future challenges as well new trends for routing and resilience are described.

Sections 4.3. This section studies the resilience capabilities required by an ILSA scheme.

Sections 4.4 and 4.5. In these two sections is introduced and validated a novel PCE scheme called context-aware PCE.

Chapter V.

This chapter reviews and summarizes the proposed ideas of this Thesis. Moreover, it suggests avenues for future work.

2 Routing and Resilience in Carrier-Grade Networks

This section deeply discusses both routing and resilience schemes for CGNs. First, the RWA problem for planning and online unprotected scenarios is analyzed in a comprehensive manner. Then, the RWA problem is extended for protected scenarios, presenting two novel proactive protection schemes based on NCP. Finally, this thesis distils the RWA problem in protected scenarios under the presence of inaccurate NSI.

2.1 Routing and Wavelength Assignment in WDM Networks

WDM technologies are becoming a widely used commodity for the transport layer of CGNs due to their vast transmission capacity and low power consumption. This is noticeable by the continuously growing deployment of optical commodity switches based on WDM technologies, replacing electrical switches at the core level in DCN scenarios [13].

In a WDM network – also referred to as WRNs–WR nodes are capable of routing data in the optical domain without OEO conversion at intermediate nodes –this is known as all-optical WRNs. In order to successfully establish a connection between two WRs, a lightpath consisting in both a route and an optical wavelength must be properly selected on a short-term basis by means of a RWA algorithm. In planning scenarios, the set of CRs are known in advance. Therefore, the RWA problem consists in establishing a lightpath for all the CRs (minimizing blocking) while allocating the minimum amount of network resources (optical wavelengths). This problem can be formulated as a Mixed-Integer Linear Problem (MILP) which complexity is NP-Complete [20].

On the other hand, in dynamic (online) scenarios, a CR arrives at a WR in a random manner (with certain inter-arrival time) and it remains in the network during a certain amount of time (with certain holding time). Typically, it is assumed that the inter-arrival time follows a Poisson distribution whereas the holding time is exponentially distributed.

The RWA problem in dynamic scenarios must be solved locally by each WR –in case that it is assumed source or destination based routing—or by a dedicated lightpath computation

Table 2.1: List of Routing Heuristics.

Routing Heuristics	
Fixed-Alternate	[23],[24],[25].
Adaptive Routing	[26],[27],[28].

entity such as a PCE. Notice that source-based RWA algorithms combined with a distributed control plane scheme such as GMPLS or ASON is the option commonly recommended [21], [22].

In order to minimize complexity, the RWA problem can be either jointly solved or splitted into two sub-problems: 1) **the Routing sub-problem,** and 2) the **Wavelength Assignment sub-problem**; then each sub-problem can be independently solved. Heuristics are commonly used for both the routing and the wavelength assignment sub-problems.

There are three major heuristics for addressing the routing sub-problem:

1. **Fixed-Routing:** consisting in a pre-computed single (candidate) route for each source-destination pair.

2. **Fixed-Alternate Routing:** for this approach multiple candidate routes for each source-destination pair are pre-computed offline.

3. **Adaptive Routing:** under this approach, a route is selected online according to a cost function which uses (commonly) global NSI such as available bandwidth, number of optical fibers, etc.

In Table 2.1 a wealth of routing heuristics are listed.

Both Fixed and Fixed-Alternate routing approaches yield a lower path computation time as well as a less signaling overhead compared with adaptive routing since routes were pre-computed offline.

Nevertheless, a handicap related to the pre-computation of candidate routes is the difficulty in finding the set of optimal candidate routes under the absence of online NSI. Moreover, the set of candidate routes must be recomputed whenever the network topology have changed, e.g., link or node failures.

Once a route is selected, a heuristic is used to select a wavelength. Conversely to planning scenarios, in dynamic scenarios the number of optical wavelengths as well as optical fibers on a link is fixed. Hence, the main goal of a wavelength assignment heuristic is to minimize blocking. In Table 2.2 distinct wavelengths assignment heuristics for single and multi-fiber networks are listed.

Table 2.2: List of Wavelength Assignment Heuristics.

Wavelength Assignment Heuristics	
Single Fiber	Random Wavelength Assignment (RR), First-Fit (FF), Least-Used[29], Most-Used [30], Relative Capacity Loss [31], Wavelength Reservation, Protecting Threshold [32].
Multiple Fiber	Min-Product, Max-Sum, Relative Capacity Loss, Least-Loaded [33].

Once a lightpath is computed a reservation protocol is required in order to allocate the selected optical wavelength on the most suitable route. There are four main approaches of resource reservation schemes.

1. ***Parallel Reservation***. This scheme assumes that each WR has NSI related to the whole network topology (global NSI). Once a lightpath is computed, the source WR sends reservation messages to each WR on the selected route in order to allocate the selected wavelength on the links forming the selected route [34].

2. ***Source-initiated Reservation***. Under this approach, once a lightpath must be setup,a source WR sends a reservation message along the selected route reserving some optical wavelengths along this route. Once the reservation message reaches the destination WR, this one selects a wavelength that has been successfully reserved and sends a confirmation request to the source WR, releasing other reserved wavelengths [35].

3. ***Destination-initiated Reservation.*** Under this approach, the source WR selects a route, then it sends a reservation message along this route to collect NSI related to the available wavelengths along this route, i.e., NSI is collected on the fly. Once the reservation message reaches the destination node, this one selects an optical wavelength based on the NSI collected [35].

Since NSI is collected on the fly, destination-initiated reservation yields a higher performance compared with parallel and source-initiated reservation schemes. However, destination-initiated reservation schemes require a higher time to provision a lightpath due to the propagation delay of the fiber links. On the other hand, parallel reservation schemes are cumbersome due to signaling overhead issues.

Source-initiated reservation schemes do not have the disadvantages related to signaling overhead and lightpath set-up of both parallel and destination initiated schemes. As a matter of fact, a source-initiated reservation scheme is the option commonly recommended for distributed control planes such as ASON.

2.2 Challenges for RWA Algorithms

As mentioned in section 2.1, the WCC substantially impacts the blocking probability of a RWA algorithm under inaccurate NSI.

There are several sources that may potentially cause inaccurate NSI such as: 1) the information aggregation caused by a hierarchical network design; 2) the non-negligible delay propagation; 3) the failure of control messages;) 4) frequent CRs arrivals; and 5) the trigger policy that determines when NSI should be disseminated, which often follows a periodic behavior [36],[37].

The traditional main vehicles to offset the negative effects caused by inaccurate NSI are supported by: i) the use of multi-fiber systems; ii) enhancing WR nodes with wavelength conversion capabilities– WCRs; and, iii) decreasing the update time [38], [39].

On one hand, the simplest strategy (from a CAPEX perspective) to mitigate the negative effects of inaccurate NSI is to reduce the update time. However, this might be cumbersome because of signaling overhead (update NSI messages) concerns that lead to scalability issues. Moreover, it is worth mentioning that even with unrealistic updating periods, i.e., flooding update messages per network state change, NSI might still be inaccurate due to non-neglected propagation delays [19].

On the other hand, the use of both multi-fiber systems and WCRs tend to be overlooked because of the added high costs. In addition, WCRs deployment is also avoided because of technical difficulties and power consumption. In light of this, there are several studies available in the literature dealing with both the WCRs placement with a limited range of wavelength conversion –sparse and limited WI networks–, and the deployment of non-uniform multi-fiber systems, –the number of fibers on the network links are different–, aiming at reducing both costs and technical difficulties [40]. However, the deployment of WCRs and multi-fiber systems may not completely guarantee NSI accuracy.

The potential inaccuracy of the NSI is substantially affecting the routing decision process, this is a well-known problem referred to as the RI problem [14]. There are several studies available in the literature dealing with the RI problem in unprotected scenarios. It is institutive that the negative effects of inaccurate NSI are more harmful in protected scenarios, where for each primary lightpath a link-disjoint backup lightpath must be computed.

In this thesis it is considered that further investigation is needed aiming at designing RWA algorithms capable of handling the negative effects of the RI problem in protected scenarios. In addition to mitigate the RI problem, protection RWA algorithms must be efficient (low P_{cost}) and fast recovery times.

2.3 Offline RWA in Protected Scenarios

The deployment of RWA in protected scenarios has the same constraint (WCC) as unprotected scenarios. However, a new constraint must be added, this is that a primary and backup lightpath must be link-disjoint. In the following sections it is discussed the issues related to the deployment of NCP in planning scenarios with fixed-grid. Moreover, it is introduced the benefits of multiple-coding related to the Protection Cost.

2.3.1 Network Coding Protection in WDM Networks with Fixed-Spectrum

In recent years, NCP has started to gain momentum in network research driven by its fast recovery times and low P_{cost}. NCP is a proactive protection scheme that combines NC and a proactive protection schemes such as DP. The pioneer work found in [41] introduced the benefits of NC to improve network throughput specifically in multicast and wireless network scenarios. On the other hand, DP schemes are one of the most widespread protection strategies used due to: (1) simplicity, (2) low recovery time, and (3) near hit-less recovery features. Nevertheless, DP schemes are severely limited by bandwidth availability.

Among the list of studies related to NCP it can be mentioned studies such as [17], proposing the use of NC combined with a 1+N protection strategy on p-cycles. Other studies such as [42],[18] proposed network coding combined with a DP scheme, hereinafter referred to as DPNC. In this thesis it is proposed an NCP scheme based on a DP strategy referred to as DPNC* scheme which uses multiple-coding to reduce the P_{cost}. DPNC* is able to operate in both IP/MPLS and Optical network layers.

For the purpose of illustrating the basic operation of DPNC* in a planning scenario, it is considered the network topology depicted in Fig. 2.1a showing a single layer connected digraph, $G(V,E)$, where V can be either the set of WRs or IP/MPLS routers and E is set the of optical links or IP/MPLS virtual connections.

To protect the traffic sent along links $e_{4,5}$, $e_{1,5}$ and $e_{3,5}$ ($T_{4,5}$, $T_{1,5}$ and $T_{3,5}$) using a DPNC*, $T_{4,5}$, $T_{1,5}$ and $T_{3,5}$ are simultaneously sent along link-disjoint paths which are ($e_{4,7}$, $e_{7,3}$, $e_{3,2}$, $e_{2,5}$), ($e_{1,2}$, $e_{2,5}$) and ($e_{3,2}$, $e_{2,5}$), respectively.

In addition, in order to benefit from NC features, node 3 codes the traffic $T_{3,5}$ (not shown in Fig. 2.1a) and $T_{4,5}$, producing $T''_{3,2}$, and node 2 codes all protected traffic (as well as already coded data), i.e., $T'_{1,5}$+ $T''_{3,2}$, producing $T''_{2,5}$, and then sends $T''_{2,5}$ to node 5 along the link $e_{2,5}$. Under this configuration, whenever there is a failure affecting only one of the protected links, for instance, $e_{1,5}$, node 5 can recover (in an agile manner) the affected traffic by decoding $T'_{2,5}$, e.g., $T''_{2,5} + T_{3,5} + T_{4,5} = T'_{1,5}$. Thus, all protected traffic is aggregated (coded) into a single data stream $T''_{2,5}$, resulting in a lower P_{cost}. Indeed, the main advantage of a DPNC scheme resides on the coding of traffic. By coding traffic (by means of a simple linear coding strategy, i.e., Exclusive-Or) is possible to transmit several data streams while using the same amount of

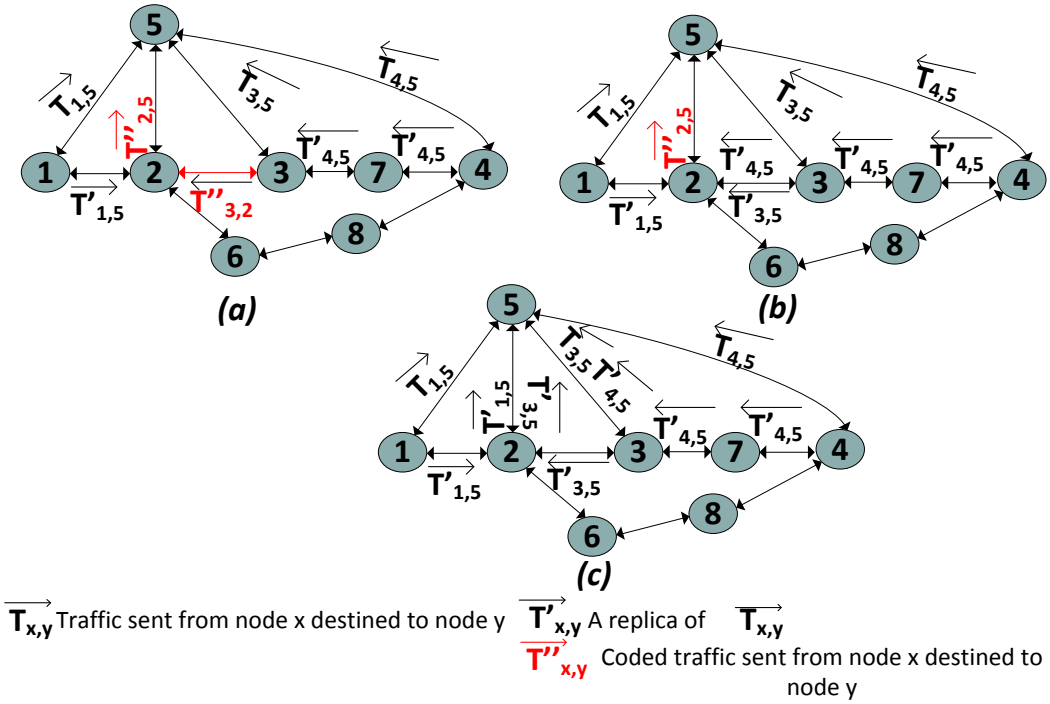


Figure 2.1: Protection strategies: a) DPNC*; b) DPNC; c) DP.

network resources required to transmit solely a single data stream.

Consider that the P_{cost} for protecting links $e_{4,5}$, $e_{1,5}$ and $e_{3,5}$ using DPNC* with multiple coding is: P_{cost} ($e_{4,5}$, $e_{1,5}$,$e_{3,5}$) = $5U$ –count the number of $T'_{x,y}$ and $T''_{x,y}$–, being U a network resource unit, such the number of optical wavelengths allocated for link protection. This cost is lower than the obtained by a conventional DP scheme ($7U$), and DPNC which is $6U$ due to the inability of multiple-coding, see 2.1c and Fig. 2.1b respectively. Therefore, it can be stated that NC with the capacity to code data already coded significantly reduces the P_{cost}.

In the scenario described in this section the following settings are considered:

1. The input traffic for all links has the same bitrate and requires the allocation of $1U$.
2. The objective is to protect all links.
3. The finest traffic granularity to be protected is the whole traffic sent along a link.
4. For simplicity, all coding operations are based on the exclusive-or over $GF(2)$, i.e., the Galois field of two or more data streams.
5. The NCP schemes described on this thesis are based on systematic coding. For more information concerning other coding strategies the reader is referred to [43].
6. A standard fixed spectrum grid of $50\,GHz$.

Table 2.3: List of Symbols and Terminology for Section 2.3.1

W	The set of protection groups.
$P_{j,k}$	Set of protection paths for protection group j, where $j \in W$ and $k \in j$.
$P'_{j,k}$	Pre-eliminary set of protection paths for protection group j.
D_{in}	Function that given a node returns its indegree.
V'	Set of nodes with an indegree greater than 2.
U	Represents the cost to send traffic along a link.

7. Links with common terminal vertices are protected. It is worth mentioning that links with different terminal vertices can be also protected. However, in this thesis it was considered that the protection of links with common terminal vertices minimizes the complexities of NC operations, i.e., minimize coding operations as well as NSI related to the coded data streams[1].

8. Single link failures are assumed since they are the most frequent type of failures in communication networks. It is worth mentioning that the paths $e_{2,5}$ and $e_{3,2}$, $e_{2,5}$ are referred to as coding paths. A coding path is a path that conveys coded (protected) traffic. In addition, note that $T''_{2,5}$ encodes the already coded traffic ($T''_{3,2}$), this is the concept of multiple-coding introduced in [45], which minimizes the P_{cost}. Moreover, Table 2.3 summarizes the list of symbols used in this section.

The goal of DPNC* is to maximize the amount of coded data as long as it reduces the P_{cost}. For instance, more coded traffic leads to the allocation of less optical wavelengths, see Equation 2.1.

$$min \sum_{j \in W} \sum_{k \in j} |P_{j,k}| \tag{2.1}$$

The set of protection paths for a protection group j can be defined as shown in Equation 2.2. This is the set of coding paths (common links) and the no-common links.

$$P_{j,k} = \left(P'_{j,k} / \left(\bigcap_{k \in j} P'_{j,k} \right) \right) \bigcup \left(\bigcap_{k \in j} P'_{j,k} \right) \tag{2.2}$$

The operation of DPNC* is shown in Algorithm 1. The main aim of DPNC* is to avoid

[1] It is worth mentioning that there are studies available in the literature that deal with NCP with different destinations [44].

the forwarding of coded traffic by the terminal vertices of the links jointly coded (protected), i.e., only links or paths with common terminal vertices are protected. Moreover, candidate protection paths are selected as long as the cost of a candidate path is not more than x times (we set x to 4) the cost of the shortest-path for the same pair of endpoints. This is done in order to reduce the time complexity of DPNC*.

Algorithm 1 Overview of DPNC*.

Input: $(G(E,V), layerTechnology)$, **Output**: (P_{cost})

$P_{cost} = 0$ {Initialize the total protection cost}
W =Create protection groups according to the network layer technology$\left(layerTechnology\right)$.
for i in W **do**
  $G'(E,V) = G(E,V)$
  for j in i **do**
    Remove each $j \in i$ from $G(E,V)$ {Remove primary links.}
  $Backup_i$ =Compute a set of candidate protection paths$(G(E,V))$.
  δ =Create subgroups formed by a single protection path belonging to each set (protection paths/link)$\left(Backup_i\right)$.
  for k in δ **do**
    $\alpha_k = \cap_{n=1}^{|k|} \delta_k$ {find common links among the protection paths, this implies that along these links traffic is suitable for coding.}
    $P_{cost}^{'} = \emptyset$ {Initialize the protection cost set of each protection subgroup.}
    if $\alpha_k != \emptyset$ **then**
      $\beta_k = \delta_k \setminus \alpha_k$ {find no common links.}
      $P_{cost}^{'}.add\left(Cost(\alpha_k) + Cost\left(\beta_k\right)\right)$ {Compute the protection cost and add it to the set P'_{cost}}
    else
      $P_{cost}^{'}.add\left(Cost(\delta_k)\right)$
  $P_{cost} = P_{cost} + min\left(P_{cost}^{'}\right)$
  $G(E,V) = G'(E,V)$ {add primary links.}

Furthermore, it is created what is called protection groups (the data streams sent along links with common terminal vertices which are suitable for NCP) following two strategies. In case that the network to be protected uses optical technologies the goal of DPNC* is to maximize the size of protection groups. This is handy for topologies with a low Average Node Degree (AVND). DPNC* attempts to find a balance between the size of protection groups and the use of conventional DP for those data streams that cannot be coded (protected with NCP) for topologies with a high AVND, as it is the usual case with optical technologies the goal of DPNC* is to maximize the size of protection groups.

On the other hand, based on the conditions assumed in this section, it can be deduced that the number of links that can be protected using NCP based on a DP scheme for single failure scenarios can be computed as shown in Equation 2.3.

$$\sum_{i' \in V'} D_{in}\left(i'\right) - 1 \tag{2.3}$$

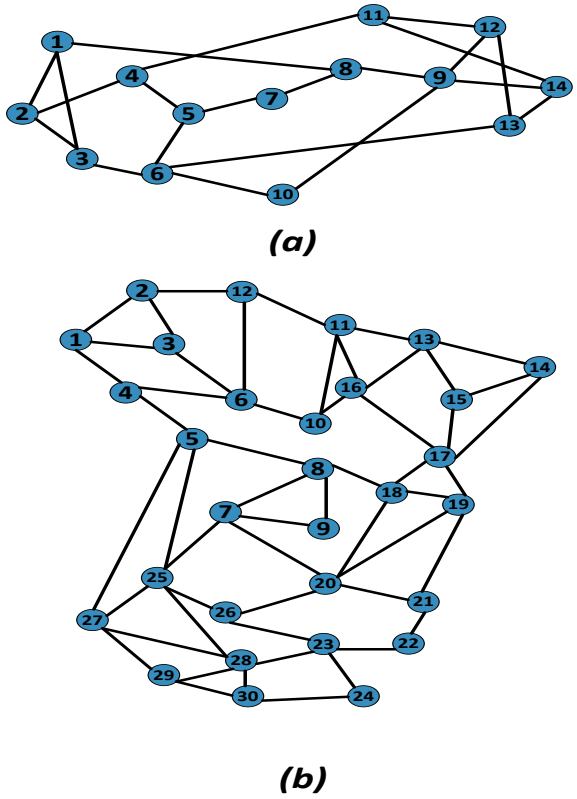


Figure 2.2: Evaluated Network Topologies assuming a planning scenario with a fixed-grid.

Therefore, it will be impossible to protect the traffic sent along all links of a network using NCP based on a DP scheme. As a consequence, for those cases conventional DP is used instead, i.e., DP is used for small protection subgroups, $|j| < 3$.

2.3.2 Evaluation of Protection Schemes in WDM Networks with Fixed-Spectrum

For the purpose of evaluating the performance of the proposed NCP scheme, namely DPNC*, against other proactive protection solutions, in this thesis it is used the well-known programming language python [46] and the library NetworkX [47] (a tool for creating and manipulating graphs and networks) to build the simulation testbed. The performance of the proposed protection scheme is evaluated on the two network topologies shown in Fig. 2.2, the well known NSFNET topology (14 nodes, 21 links), and a model of the real Spanish Backbone topology, hereinafter referred to as TID topology.

The performed trials assumed the following conditions: 1) routes are computed through the shortest-path routing algorithm using hops as the metric; and 2) a link is considered "congested" whenever more than half of its capacity is used for link protection. It is worth mentioning that when a link is congested it cannot be used to allocate protected traffic.

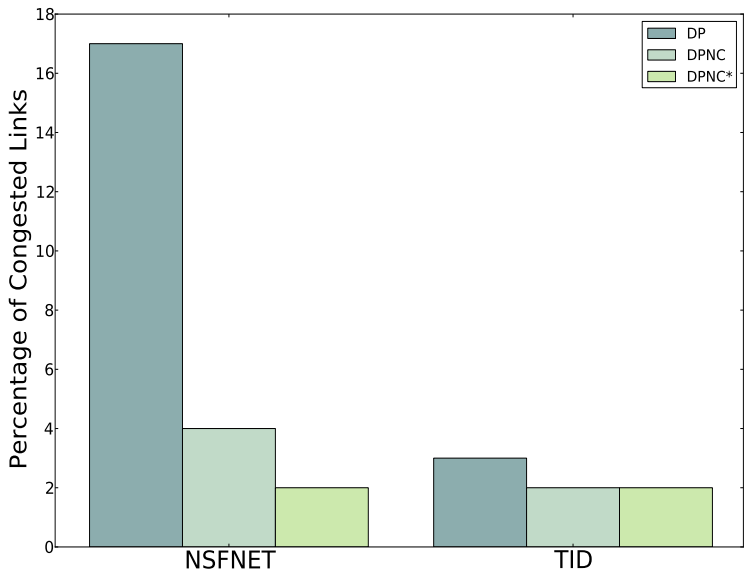


Figure 2.3: Total of Congested Links.

It must be noticed that the proposed protection scheme is designed to be employed at the network planning phase; hence, traffic is not generated dynamically, and rather it is assumed a certain fixed link capacity to cope with the expected traffic demand and the traffic to be protected. Moreover, the evaluation is performed under the premise that the network topology is static, i.e., it maintains its structure during time. Therefore, a modification in the network topology will require a new network planning for defining the protection levels.

Evaluation results for both network topologies are depicted in Fig. 2.3 and Fig. 2.4. The metrics used in the evaluation are: 1) the percentage of congested links – links which more than 50% of their capacity is allocated to protected traffic; and 2) the P_{cost}—the total amount of network resources allocated to link protection.

It can be observed that for the NSFNET topology the performance obtained with DPNC* is higher than the one for both DP and DPNC schemes. On one hand, only 1% of links were congested when using DPNC*, compared with 17% using DP and 5% using DPNC. On the other hand, the protection cost when using DPNC* is $138U$ compared with $140U$ and $154U$ when using DPNC and DP respectively, where U represents the cost to send traffic along a link.

Finally, for the TID topology, DPNC* shows a P_{cost} reduction of 12% ($231U$) compared to DP protection ($264U$), while 2% reduction ($259U$) is obtained with DPNC. However, the advantages of multiple-coding related to the percentage of congested links are not so noticeable since the connectivity of the TID topology is high.

In addition to the evaluation results obtained in Fig. 2.3 and Fig. 2.4, Table 5 summarizes

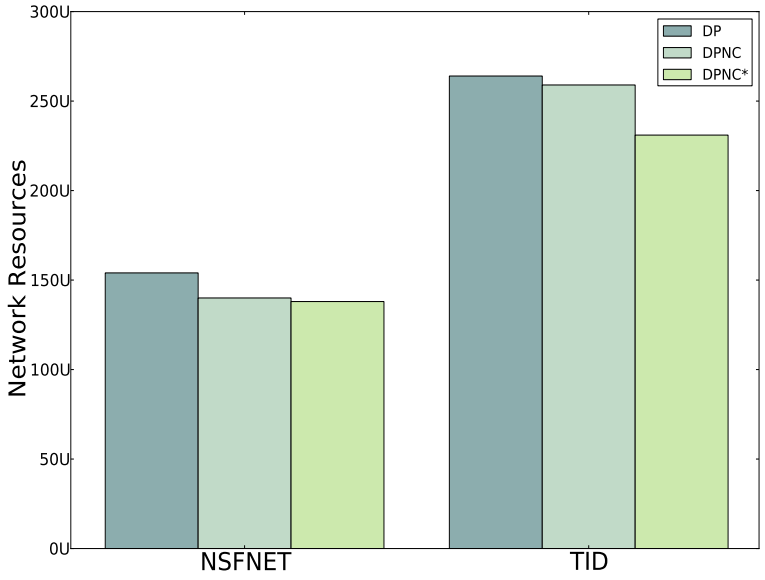


Figure 2.4: Total Protection Cost.

Table 2.4: Percentage of P_{cost} over the total network capacity.

Protection schemes	**Evaluated network topologies**		
	B1 topology	NSFNET topology	Telefonica I+D topology
DP	28%	36%	23%
DPNC	27%	33%	22%
DPNC*	26%	32%	20%

the percentage of P_{cost} over the total network capacity available for each network topology evaluated. It can be observed that the proposed protection strategy, DPNC* shows a more efficient utilization of the available network resources.

2.3.3 NCP in WDM Networks with Flexible-Spectrum

In the previous section we evaluate the benefits in terms of P_{cost} brought by NCP schemes assuming a fixed-grid spectrum. In this section we study the performance of NCP schemes considering a flexible-grid spectrum.

WDM networks commonly use a 50 GHz fixed-grid [48] that may result in an inefficient utilization of the OS. To overcome this issue, EONs have been proposed with the aim of enabling enough flexibility to adapt the transponders bit rate to heterogeneous line rates [49]. It has been already demonstrated in the literature that an efficient utilization of the OS leads

Table 2.5: OS utilization for Flexible and Fixed 50 GHz grid WDM solution

Demand bit rate (Gbps)	Modulation Format	Reach (Km)	EON	Fixed-Grid
Br <40	QPSK (1 subcarrier)	2000	12.5	50
$40 \leq Br \leq 100$	16-QAM (1 or 2 subcarriers)	500	12.5-25	50
$100 < Br < 200$	16-QAM (3 or 4 subcarriers)	500	37.5-50	100
$200 \leq Br \leq 300$	32-QAM (4 or 5 subcarriers)	250	50-62.5	150
$200 \leq Br \leq 300$	64-QAM (3 or 4 subcarriers)	125	37.5-50	150

to a reduction in both equipment power consumption (OPEX cost), as well as equipment installation, i.e., transponders or optical fiber (CAPEX cost) [50].

The major building blocks of EONs are the Orthogonal Frequency Division Multiplexing (OFDM) and the Coherent Detection techniques, both combined with the exploitation of distinct modulation schemes. It is worth mentioning that a flexible-grid configuration is not limited to an unique multiplexing technique such as OFDM.

There are several studies in network research discussing the benefits of EONs with regard to energy efficiency compared to traditional fixed-grid optical networks [51], [52]. Other studies available in [53], [54] focus on combining EON techniques with protection schemes for the purpose of reducing power consumption.

It is an intuitive reasoning that a substantial reduction on the P_{cost} can be achieved by combining the flexibility of EONs with NC techniques. However, these benefits remain unaddressed. In light of this, in this thesis we derive a mathematical formulation for the deployment of DPNC* enhanced with the flexibility inherent to the OS utilization enabled by EON. The proposed NCP scheme is referred to as E-DPNC*.

In the previous section it was highlighted the operation of a DPNC* scheme in fixed-grid scenarios. Nevertheless, in flexible grid setups there are several issues that need to be considered when deploying a DPNC* scheme. Consider that in EONs the OS allocation depends on the modulation format used, e.g., BPSK, QPSK, 16, 32 or 64-QAM. The modulation format is selected according to the transmission rate as well as the distance length (transparent reach), i.e., maximum distance without OEO conversion. On this basis, the computation of link-disjoint paths must take into account both the transmission rate of the demands sent along primary links to be protected as well as the (geographic) distance length of the backup paths.

Table 2.5 shows the OS utilization for both fixed and flexible (EON) grid WDM solutions

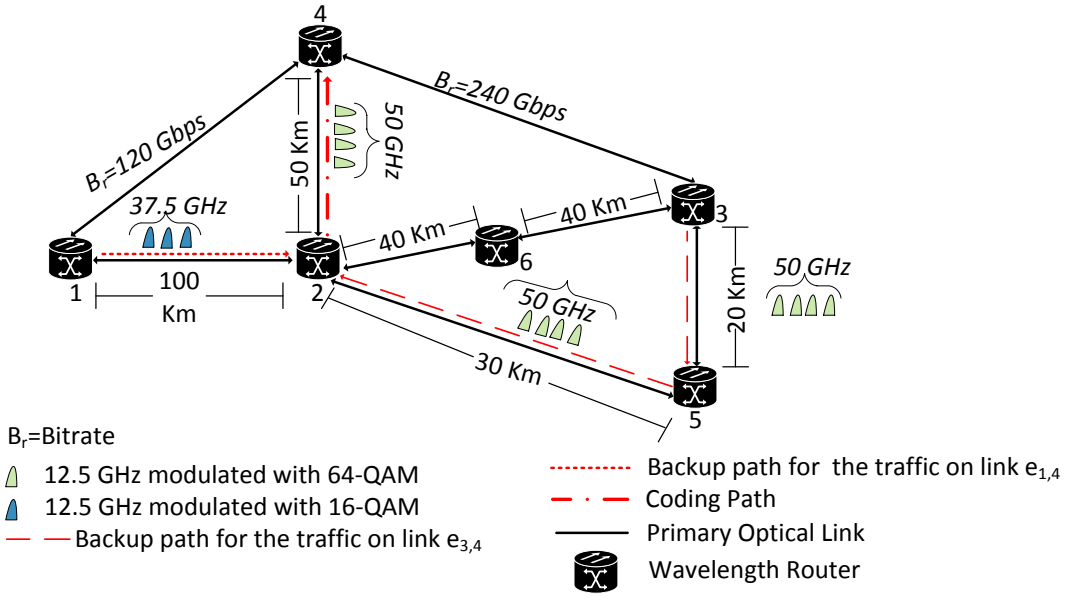


Figure 2.5: Link Protection using an E-DPNC* scheme in an EON scenario.

according to the transparent reach (the optical signal transmission without OEO conversion) as well on the Bitrate (Br) of the data stream to be transmitted [54]. Notice that an additional 10 GHz guard band (not shown in Table 2.5) must be allocated to avoid the negative effects of adjacent channel interference. To illustrate the operation of DPNC* scheme in EON scenarios we consider the network topology shown in Fig. 2.5. It is also worth noticing that in this flexible grid scenario, the P_{cost} refers to the amount of OS required to enable link protection.

In order to protect the traffic sent along the primary links $e_{1,4}$ and $e_{3,4}$ –which are modulated using a 16-QAM and 64-QAM schemes respectively– with DPNC*, the following configuration is enabled: 1) The backups paths $e_{1,2}$, and $e_{3,5}$, $e_{5,2}$ are provisioned, and; 2) the path $e_{2,4}$ is configured as the coding path. With this configuration the P_{cost} –according to the spectrum slicing shown in Table 6 for EONs— is 175 GHz:37.5 GHz allocated to path $e_{1,2}$; 100 GHz allocated to path $e_{3,5}$, $e_{5,2}$, and 50 GHz allocated to the coding path ($e_{2,4}$). This configuration is the most suitable according to both the geographic distance of the backup paths, and the demanded bit rate traversing the primary links. Otherwise, if the backup path $e_{3,6}$, $e_{6,2}$ is configured instead of backup path $e_{3,5}$, $e_{5,2}$, the P_{cost} would reach almost 190 GHz because the transparent reach for a optical signal modulated with a 64-QAM scheme is 125 Km (kilometers). Note that the distance length of path $e_{3,6}$, $e_{6,2}$, is 130 Km, hence, a 32-QAM modulation scheme must be used.

On the other hand, in case that a fixed-grid solution is used for the scenario depicted in Fig. 2.5 the P_{cost} would be 550 GHz. Nevertheless, a higher P_{cost} (650 GHz) is obtained if conventional DP is used instead of DPNC*. Therefore, DPNC* combined with the features of EONs (hereinafter referred to as E-DPNC*) provides a significant P_{cost} reduction in EON

scenarios.

Operation of E-DPNC* in single link failure scenarios

The goal of E-DPNC* is to link enable protection in such a way that the coding of traffic is maximized, while simultaneously minimizing the P_{cost}. Notice that the deployment of NCP in EON scenarios is a sub-problem of the 1+1 DP formulation proved to be NP-Complete by authors in [55].

In the following lines it is described mathematical model for the deployment of E-DPNC* in single link-failure scenarios. The notation used for this model is depicted in Table 2.6.

Table 2.6: List of Symbols and Terminology for section 2.3.3.

Symbol	Meaning
$G(V,E)$	Directed graph representing a EON scenario, where E is the set of optical links and V is set of WRs.
W	Set of Protection Groups.
g	A Protection Group, where $g \in W$.
s	Protection Subgroup, where $s \in g$.
$\delta(j,d)$	The OS required to route a traffic demand along a given link
$\iota(j)$	The length of link j (kilometers).
$D_1(d)$	$D_1(d) = 1$ if demand $d < 40$ Gbps, otherwise 0.
$D_2(d)$	$D_2(d) = 1$ if demand $40\,Gbps \leq d \leq 100\,Gbps$, otherwise 0.
$D_3(d)$	$D_3(d) = 1$ if demand $100\,Gbps < d < 200\,Gbps$, otherwise is 0.
$D_4(d,\iota(j))$	$D_4(d) = 1$ if demand $200\,Gbps \leq d \leq 300\,Gbps$ and the length of link j is $>$ 125 and $\leq$ 250, otherwise is 0.
$D_5(d,\iota(j))$	$D_5(d) = 1$ if demand $200Gbps \leq d \leq 300\,Gbps$ and the length of link j is $\leq$ 125, otherwise is 0.
x_j^d	$x_j^d = 1$ if link j is the primary link for demand d, 0 otherwise.
y_j^d	$y_j^d = 1$ if link j belongs to the protection path of demand d, otherwise is 0.
$y^{1,d}_j$	$y'^d_j = 1$ if link j belongs to the second protection path of demand d, otherwise is 0.
z^s	$z^s = 1$ if protection subgroup s is protected, 0 otherwise.
R_d	Receiver Node of demand d.
S_d	Source Node of demand d.
L'	The set of affected failure links.
Γ_j^s	$\Gamma_j^s = 1$ if link j belongs to protection subgroup s, otherwise $\Gamma_j^s = 0$.
L'	The set of failed links.
D'	The set of failed demands.

The objective function is defined as given a graph $G(V,E)$(representing a flexible-grid optical network) minimize the P_{cost}. The mathematical model is as follows.

$$min \sum_{k \in g} z^s \times \left[\sum_{j \in E} \sum \delta(j,d) \times y_j^d - \sum_{d \in s} max\left(\bigcup \delta(j,d) \times y_j^d \right) \right] \tag{2.4}$$

Equation (2.4) defines the objective function as the OS needed to route the traffic demand along a protection path, minus the OS saved by the throughput improvement achieved by means of NC features.

The mathematical constraints are the following.

$$x_j^d + y_j^d \leq 1 \; \forall s \in g, d \in s, j \in E \tag{2.5}$$

Equation (2.5) ensures link-disjointness between a primary link and its backup path.

$$\sum_{s \in g} Z^s = 1 \tag{2.6}$$

Equation (2.6) defines that solely one protection subgroup belonging to a specific protection group will be protected.

$$\sum_{\forall (v,u) \in E} x_{(v,u)}^d - \sum_{\forall (u,v) \in E} x_{(u,v)}^d = \begin{cases} 1, & if \; v = R_d \\ -1, & if \; v = S_d \\ 0, & otherwise \end{cases} \tag{2.7}$$

$$\sum_{\forall (v,u) \in E} y_{(v,u)}^d - \sum_{\forall (u,v) \in E} y_{(u,v)}^d = \begin{cases} 1, & if \; v = R_d \\ -1, & if \; v = S_d \\ 0, & otherwise \end{cases} \tag{2.8}$$

Equations (2.7) and (2.8) formulate the flow conservation constraints for the primary links

and the backup paths respectively.

$$\delta(j,d) = 12.5D_1(d)+25D_2(d)+37.5D_3(d)+50D_4(d,l(j))+62.5D_5(d,l(j)), \forall d \in s, j \in E \quad (2.9)$$

Equation (2.9) captures the strategy for OS assignment.

Operation of E-DPNC* in multiple link failure scenarios

Despite the fact that multiple simultaneous link failure scenarios are not as common as single link failure scenarios, they must be also addressed by a protection scheme since their direct impact on traffic losses might be highly significant. This is proven by several works already available in the literature related to multiple link failure scenarios caused by natural disasters, power outages or even by malicious attacks [56], [57], [58].

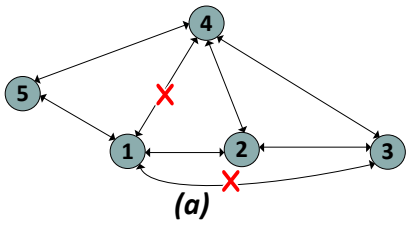
An NCP scheme may not be feasible under certain type of multiple link failure scenarios, i.e., two or more links fail simultaneously. By feasible, it is referred to the cases where the P_{cost} lower than the one obtained using a DP scheme. In this thesis three types of multiple link failure scenarios (A, B, C) are distinguished assuming in all of them that the network topology maintains its connectivity. Figure 2.6 depicts all three scenarios as well as the primary ($T_{x,y}$) and backup ($T'_{x,y}$) data streams sent for every source-destination pair.

In Failure Scenario A (see Fig. 2.6a) we consider that two or more primary links fail simultaneously, e.g., $e_{1,4}$ and $e_{1,3}$. In this scenario, an NCP scheme can efficiently recover the affected traffic traversing the failed optical links, as long as: i) the protection paths of the failed links are link disjoint, and; ii) no more than two primary links with the same terminal vertex simultaneously fail. Failure Scenario A is formulated as follows.

$$\sum_{j \in L'} r_j^s \leq 1 \; \forall s \in g \quad (2.10)$$

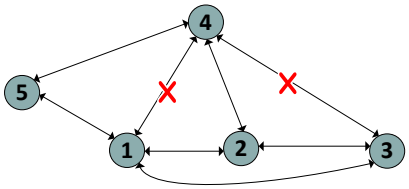
$$\sum_{d \in D'} x_j^d + y_k^d = 1 \; \forall j,k \in L', k \neq j \quad (2.11)$$

In Failure Scenario B (see Fig. 2.6b) we consider that two or more links fail simultaneously, assuming that; i) at most two primary links with the same terminal vertex fail, e.g., $e_{1,4}$ and $e_{3,4}$, and; ii) the terminal vertex indegree of the failed links is greater than one. We also assume that a failed link does not belong to the protection path of another failed link. This scenario is



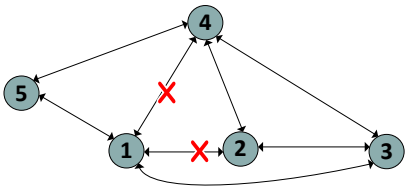
(a)

*	Source				
Destination	**1**	**2**	**3**	**4**	**5**
1	0	$T'_{2,4}$	0	0	$T'_{5,4}$
2	$T'_{1,4}$	0	$T'_{3,4}$	0	0
3	0	0	0	0	0
4	$T_{1,4}, (T'_{2,4} \oplus T'_{5,4})$	$T_{24}, (T'_{1,4} \oplus T'_{3,4})$	$T_{3,4}$	0	$T_{5,4}$
5	0	0	0	0	0



(b)

*	Source				
Destination	**1**	**2**	**3**	**4**	**5**
1	0	$T'_{2,4}, T'_{3,4}, T'_{5,4}$	0	0	$T'_{5,4}$
2	$T'_{1,4}, (T'_{5,4} \oplus T'_{2,4} \oplus T'_{3,4})$	0	$T'_{3,4}$	0	0
3	0	0	0	0	0
4	$T_{1,4}, (T'_{2,4} \oplus T'_{5,4})$	$T_{2,4}, (T'_{1,4} \oplus T'_{3,4}), T'_{5,4}$	$T_{3,4}$	0	$T_{5,4}, (T'_{1,4} \oplus T'_{2,4})$
5	$T'_{2,4} \oplus T'_{1,4}$	0	0	0	0



(c)

*	Source				
Destination	**1**	**2**	**3**	**4**	**5**
1	$T'_{1,4}$	$T'_{2,4}$	0	0	$T'_{5,4}$
2	0	0	$T'_{3,4} \oplus T'_{1,4}$	0	0
3	$(T_{1,4} \oplus T'_{3,4}), (T'_{2,4} \oplus T'_{5,4})$	$T'_{1,4}$	0	0	0
4	$T_{1,4}, (T'_{2,4} \oplus T'_{5,4})$	$T_{2,4}, (T'_{1,4} \oplus T'_{3,4})$	$T_{3,4}$	0	$T_{5,4}$
5	0	0	$T'_{3,4}$	0	0

$T_{u,v}$, $T'_{u,v}$ Traffic Demand, and Replica of a Traffic Demand sent along link u,v respectively

* Traffic Matrix defining the traffic demands sent along primary links and backup paths

$\oplus$ Exclusive-Or operation

Figure 2.6: Multiple link failure scenarios: 1) Scenario A; 2) Scenario B; 3) Scenario C.

formulated as follows.

$$\sum_{j \in L'} r_j^s \leq 2 \; \forall s \in g \tag{2.12}$$

$$\sum_{d \in D'} x_j^d + y_k^d = 1 \; \forall j,k \in L', k \neq j \tag{2.13}$$

It is worth mentioning that in Failure Scenario B, DPNC* might recover affected traffic, but a different the strategy related to both selection of protection paths, and decoding operations, is required see Fig. 2.6b. However, The P_{cost} used might be high in comparison with conventional DP.

In Failure Scenario C (see Fig. 2.6c), two or more links fail simultaneously, e.g., $e_{1,4}$ and $e_{1,2}$, assuming that: i) neither of the failed links have a terminal vertex in common, and; ii) at most one of the failed links is part of the backup path of another failed link, i.e., link $e_{1,2}$ is part of the backup path of link $e_{1,4}$, which is ($e_{1,2}$,$e_{2,4}$). In this scenario a DPNC* scheme can efficiently recover the affected traffic, but a second protection backup path might be required. This scenario is formulated as follows.

$$\sum_{j \in L'} r_j^s \leq 2 \; \forall s \in g \tag{2.14}$$

$$\sum_{d \in D'} x_j^d + y_k^d = 1 \; \forall j,k \in L', k \neq j \tag{2.15}$$

Equation (2.5) must be modified to meet Failure Scenario C as shown in Equation (2.16):

$$x_j^d + y_j^d + y_j^{1,d} \leq 1 \; \forall s \in g, d \in s, j \in E \tag{2.16}$$

Moreover, the following constraints must be added to achieve the proper deployment of

E-DPNC*.

$$\sum_{j \in E} y^{1,d}_{j} + y^{1,d2}_{j} > |s| \; \forall s \in g, d \in s, d \neq d_2 \tag{2.17}$$

Equation (2.17) defines the conditions to enable traffic coding for the second protection path. Equation (2.18) defines the flow conservation for the second protection path.

$$\sum_{\forall (v,u) \in E} y^{1,d}_{(v,u)} - \sum_{\forall (u,v) \in E} y^{1,d}_{(u,v)} = \begin{cases} 1, & if \; v = R_d \\ -1, & if \; v = S_d \\ 0, & otherwise \end{cases} \tag{2.18}$$

On the other hand, notice that Equation (2.16) may not be fulfilled, since it depends on topology characteristics such as edge connectivity, i.e., maximum number of link-disjoint paths. Thus, Equation (2.16) should not be considered and the following constraints must be added to the E-DPNC* problem.

$$x^{d}_{j} + y^{d}_{j} = 1 \; \forall s \in g, d \in s, j \in E \tag{2.19}$$

$$x^{d}_{j} + y^{1,d}_{j} = 1 \; \forall s \in g, d \in s, j \in E \tag{2.20}$$

$$\frac{\sum_{j \in E} \left(y^{d}_{j} + y^{1,d}_{j} \right)}{\sum_{k \in E} \left(y^{d}_{k} + y^{1d}_{k} \right)} \geq |\varphi| \; \forall s \in g, d \in s \tag{2.21}$$

Equation (2.19) and (2.20) define link-disjointness between primary paths and protection paths. Finally, Equation 2.21 defines the so-called protection grade, which is the link-disjoint degree between the protection path, and second protection path for a certain protection subgroup. The protection grade was set to 0.33 –according to [59]– for all protection subgroups, hence, a protection path and a second protection path must be at least 33% link-disjoint.

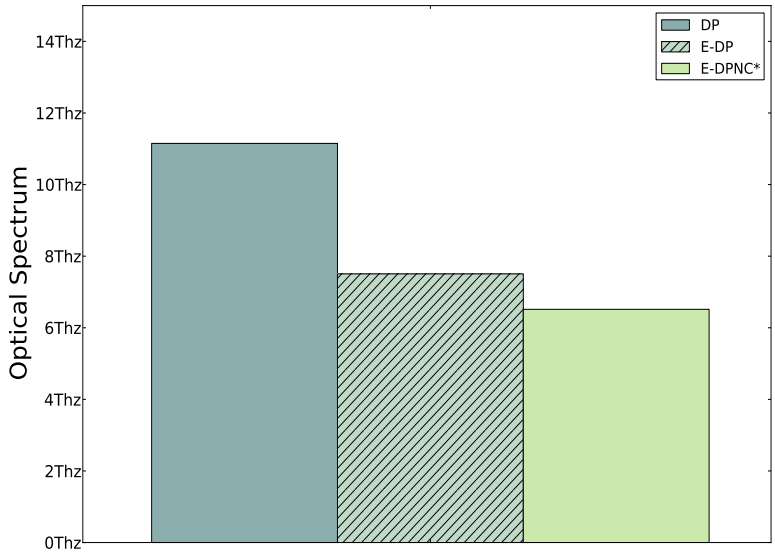


Figure 2.7: P_{cost} for single link failure scenarios.

2.3.4 Evaluation of Protection Schemes in Optical Networks with Flexible-Spectrum

In this section we compare the performance of an E-DPNC* scheme with both DP for a fixed-grid (assuming 50 GHz channels), and E-DP for a flexible spectrum configuration respectively. The evaluation environment was built using the Python graph library NetworkX. In this performance evaluation we consider the Telefonica Spanish backbone optical topology, with a traffic matrix reaching a total traffic volume of 1.20 Tbits [59], see Fig 2.1b.

Figure 2.7 and Fig. 2.8 depict the total P_{cost} for single, and multiple link failure scenarios (Failure Scenario C), where theP_{cost} is the amount of optical spectrum required to enable link protection. In addition, Fig. 2.9, depicts the P_{gain} for both E-DP and E-DPNC* schemes on single and multiple link failure scenarios. The P_{gain} computed as shown in Equation (2.22), defines the improvement in the P_{cost} for a protection scheme compared to a conventional dedicated protection scheme, where $P_{cost}(DP)$ and $P_{cost}(NCP)$ are the P_{cost} for a DP and an NCP scheme respectively.

$$P_{gain} = \frac{P_{cost}(DP) - P_{cost}(NCP)}{P_{cost}(DP)} \tag{2.22}$$

Note that a P_{cost} reduction of 32.6 and 20.5 % can be obtained with E-DP compared with DP for single and multiple link failure scenarios respectively. Nevertheless, a higher P_{cost} reduction is obtained when NC techniques are used, that is, 42% and 38% respectively.

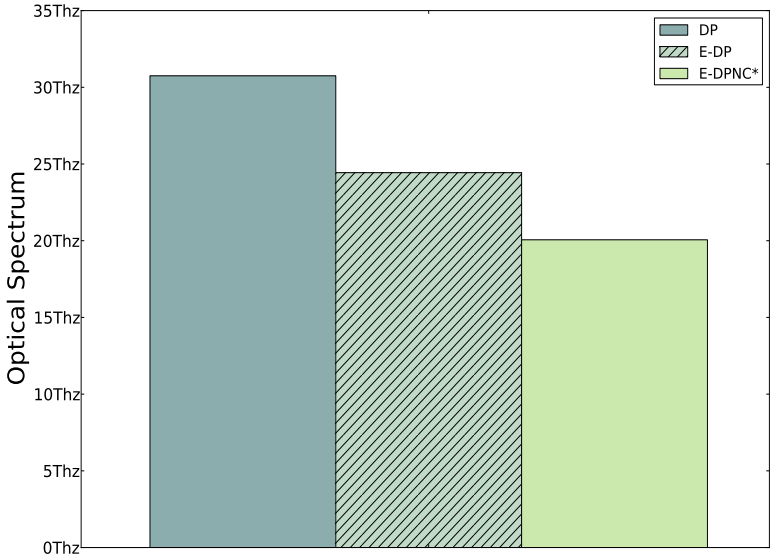


Figure 2.8: P_{cost} for multiple link failure scenarios.

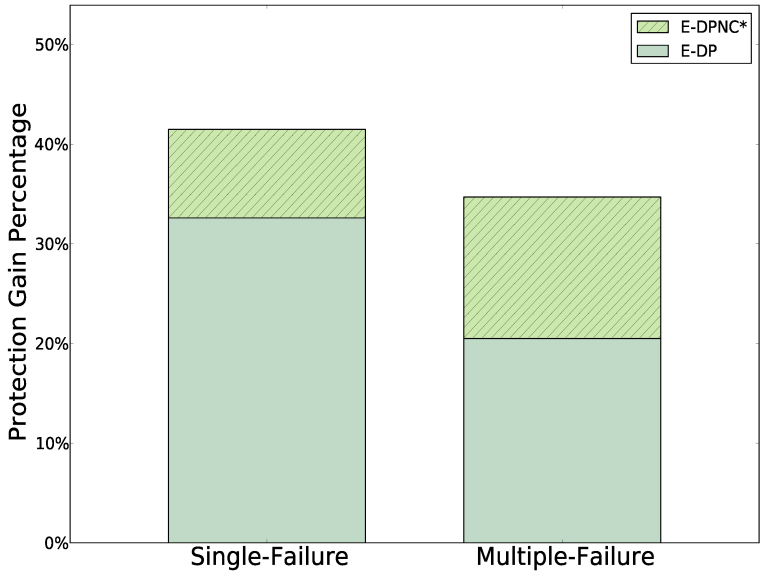


Figure 2.9: P_{gain} for single and multiple link failure scenarios.

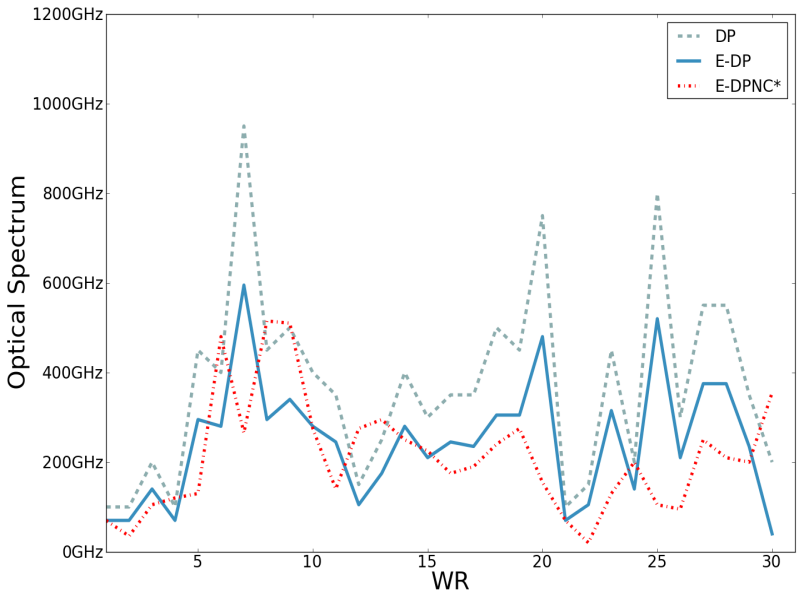


Figure 2.10: Comparison of the P_{cost} per WR in a single link failure scenarios.

On the other hand, with E-DPNC*, the P_{gain} is 41.5% and 34.7% for single and multiple link failure scenarios respectively. Therefore, the usage of NCP techniques significantly increases the P_{gain} for the E-DP scheme on a 8.9% for single link failures and a 14.2% for multiple link failures.

Moreover, Fig. 2.10 shows the P_{cost} per WR for each of the evaluated protection schemes on single failure scenarios. The P_{cost} for a specific WR strongly depends on the routing metrics used as well as on the traffic matrix. Furthermore, for the specific case of a E-DPNC* scheme, a intuitive thought is to assume that nodes with a high degree centrality (highly connected) [60], and their neighbors (direct connected nodes) are also high connected, tend to have a high P_{cost}, such as WRs 9, 10 and 6, see Fig. 2.10. Therefore, whether a WR x is highly connected and its neighbors are too, x will be probably selected as coding node, i.e., a node that codes (mix) traffic; hence, it will allocate optical resources to protected traffic. In a similar manner, Fig. 2.11 depicts the P_{cost} per WR for multiple link failure scenarios (Failure Scenario C).

Based on the numerical results presented in this section, it can be concluded that a proactive protection scheme such as E-DPNC*, that combines the advantages provided by the flexible-grid of EON and network coding related to network throughput improvement, outperforms conventional proactive protection schemes using either a fixed or a flexible grid in both single link and multiple link failure scenarios.

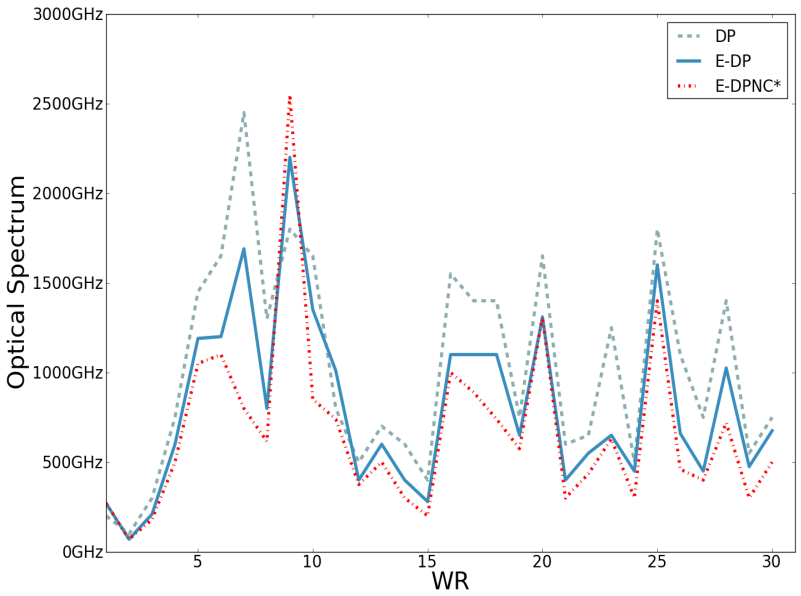


Figure 2.11: Comparison of the P_{cost} per WR in a multiple link failure scenario.

2.3.5 Techno-Economic Analysis of NCP schemes

Most of the studies related to NCP evaluate the P_{cost} in a technology agnostic manner, i.e., the specific topology technology issues, either IP/MPLS or Optical, are not considered. In fact, there is limited information in the research literature regarding the performance of NCP deployed on Optical and IP/MPLS topologies, and the advantages that NCP may bring to a network provider concerning its CAPEX and OPEX. In order to provide some lights on this issue, in this section we conduct a novel techno-economic study with the aim of evaluating both CAPEX and OPEX required by proactive protection schemes, with and without NC features, and deployed either at the IP/MPLS or at the Optical layer of a multi-layer CGN with a fixed grid.

It must be highlighted that the intention of this techno-economic study is to adopt the NCP strategies proposed in this thesis and in [18] in order to perform an extensive evaluation regarding the impact of NCP schemes (specifically NCP based on a DP scheme) on CAPEX and OPEX of a network provider. To this end, in this thesis we consider the following: 1) network layer technology (IP/MPLS or Optical); and 2) NCP schemes deployed solely at the IP/MPLS or at the Optical layer.

It can be stated that an NCP scheme might be deployed as a protection scheme either at the IP/MPLS or at the Optical layer since NC operations can be executed in the optical or in the electrical domain. Therefore, this flexibility with regard to NC operations could be exploited by distinct types of multi-layer recovery schemes, such as Top-Down, Bottom-Up or Integrated

Table 2.7: Building blocks of the multi-layer network model.

Component Type	Cost
4x 100GE line cards	36
Short-Reach Transceiver	1
All optical NC	3
WDM Transponder	15
Amplifiers: (A_p, A_b)	0.8
AWG (40 channels)	0.9
Interleaver (80 channels)	0.5
W_{SS} 1×9 (including splitter and filter)	4

approaches [61]. However, despite of this network layer agnosticism, there are several issues to be considered before deciding on the most "suitable" network layer to deploy an NCP scheme. This can encompass several metrics. For instance, notice that recovery actions executed at the IP/MPLS layer have a high granularity level, i.e., distinct protection paths can be selected per IP flow –even thought protection per MPLS label may required relevant configuration efforts. Conversely, a high granularity level cannot be achieved by a recovery action executed at the Optical layer, because the traffic is more aggregated at this layer, i.e., wavelength granularity, several IP flows may be aggregated into a single wavelength.

Another metric to be considered is the recovery time. Recovery actions executed at the Optical layer have a coarser-granularity. This implies a lower recovery time compared to recovery actions executed at the IP/MPLS layer, because recovering the traffic affected by a failure on an optical link may lead the simultaneous recovery of multiple IP flows, since along an lightpath is aggregated multiple IP flows.

This section focuses on both CAPEX and OPEX as the metrics to decide the most suitable layer to deploy an NCP scheme. To this end, we assume that the evaluated protection schemes are deployed either at the IP/MPLS layer or at the Optical layer; hence, cross-layer information is not required. Moreover, all cost values used in this thesis are normalized to the cost of a 10 Gbps transponder, i.e., 1 cost unit= cost of a 10 Gbps transponder [62]. The network components assumed in the techno-economic study and their respective costs are summarized in table 2.7. It is worth mentioning that the assumptions regarding all-optical coding devices are referred to a future scenario in which these components will be commercially available. However, based on the strong research efforts available in the literature (see Section 2.3.6) it is realistic to assume that NC components will be available soon. The cost of All optical NC features is cost computed based on the NC architecture shown in [63]

Moreover, since the costs of both IP and Optical technology equipments tend to decrease, it is reasonable to predict the cost evolution of the NEs over a period of time by means of

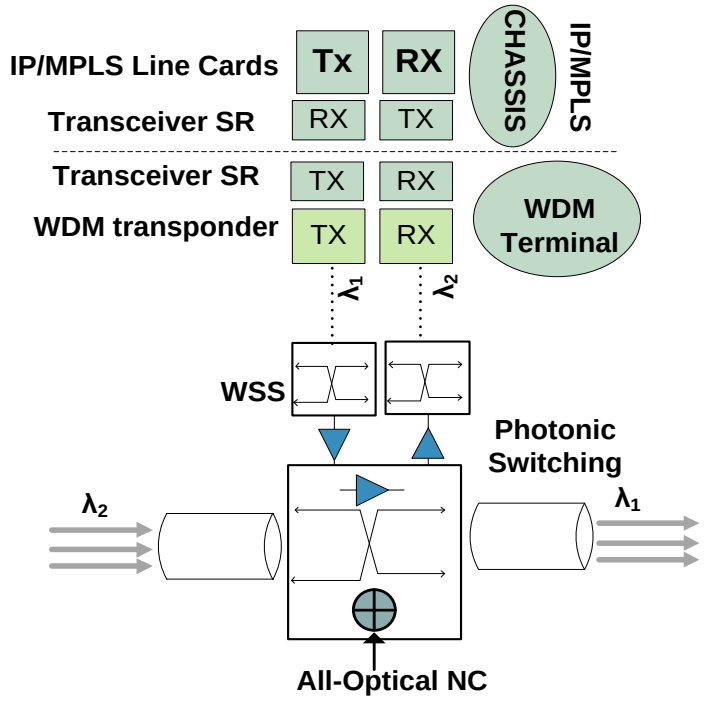


Figure 2.12: Multi-layer node architecture.

forecasting price models. Otherwise, it would not be fair to compare CAPEX or OPEX in different time periods. Forecasting models are traditionally used as network planning tools to estimate (predict) the cost evolution of technology equipments. For instance, the study available in [64] presents a cost prediction model that leverages learning curves and logistic functions. Driven by the accuracy of this prediction model, in this thesis a derivation of this model is employed in order to estimate the cost of IP/MPLS and Optical equipments. The used model includes parameters such as equipment cost in a reference year, relative accumulated production volume sold at the reference year, etc. These parameters must be adapted for each of the network equipment that will be modeled.

On the other hand, regarding the multi-layer nodes architecture, we assume a separate multi-layer network model as shown in Figure 2.12. The rationale driving this assumption is that an integrated architectural model is not a mature technology. There are several issues that need to be addressed for an integrated model such as multi-vendor interoperability.

Figure 2.13 shows the multi-layer Spanish backbone topology used for the techno economic study presented in this section. Moreover, the list of symbols and terminology are described in table 2.8 .

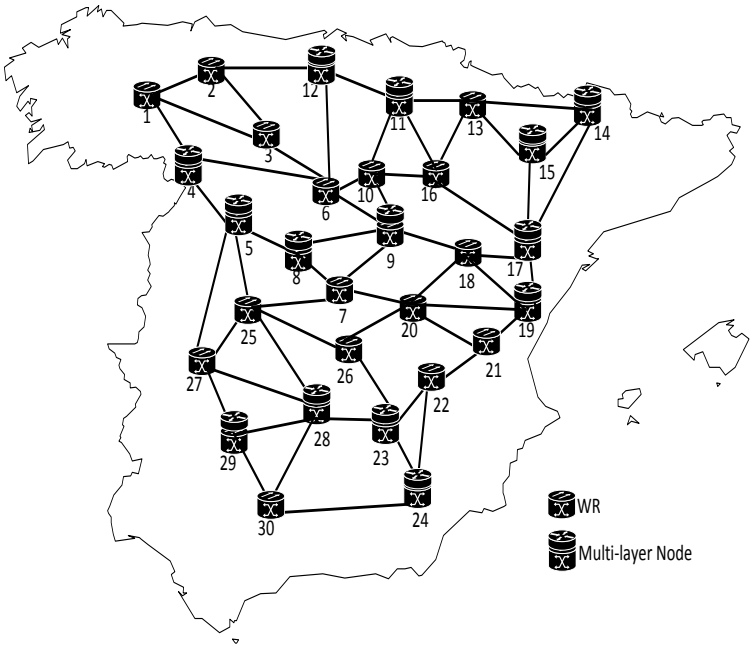


Figure 2.13: Multi-layer Spanish backbone topology.

On one hand, the following assumptions apply to the settings for the IP layer network model.

- Each IP router embeds 2 IP/MPLS router 4x100 GE line cards, since the maximum link degree of the IP topology shown in Fig. 2.13 is 6.
- The cost of each card is 36 cost units.
- The cost of short-reach transceiver is 1 cost unit.
- A 50% traffic increase per year [59].

The total IP network capacity is $C \times E$= 100 × 84=8.4 Tbps, where C and E are the capacity and total amount of IP/MPLS interfaces respectively.

- It was not considered both CAPEX and OPEX concerning electrical NC features, since they do not have a significant impact in comparison with the cost related to data streams transmission.
- It was used a protection-threshold policy, which defines the percentage of the total network capacity that is allocated for protection. In the particular testing scenario evaluated, a 50% protection-threshold was used. Whenever the protection-threshold is exceeded it is necessary to invest in new network equipment. e.g., IP/MPLS Router Line Cards.

Table 2.8: List of Symbols and Terminology for Section 2.3.5

d	A WR degree
W_{ss}	Wavelength Selective Switch.
A_{WG}	Arrayed Wavelength Grating (optical multiplexers)
A_b	Boost amplifier.
A_p	Pre amplifier.
I	Interleaver.
G_X	The cost of All-Optical NC features, which is zero when conventional DP is used.

- It is considered that the 100 Gbps IP/MPLS line cards have a power consumption of 351W [65].

On the other hand, the following assumptions apply for the Optical layer network model.

- 50 GHz fixed-grid WRs.
- The traffic sent along each optical link demands the allocation of 5 optical wavelengths.
- A single fiber system, i.e., one optical fiber per link.
- The WR type is a 80 channel OXC with a link degree equal to 5.
- The total capacity of the optical network is $30 \times 80 = 2400$ channels, where 30 is the number of WRs of the Spanish backbone topology and 80 is the number of channels supported by each WR.
- The cost of a short-reach transceiver, e.g., gray, is 1 cost unit.
- The traffic sent along each optical links have a 50% increase per year, i.e., year-0 = 5 optical channels, year-1 nearly 8 optical wavelengths and so on.
- With the aim of providing realistic results the multilayer cost model presented in [62] is extended to cover the cost of WRs with all-optical coding functionalities. For this purpose, it was assumed all-optical XOR logic gates using Semiconductor Optical Amplifiers (SOAs) based on Cross-Phase Modulation (XPM) with integrated interferometers. This type of XOR gate is widely used because of its low power consumption, high operations speed –over 40 Gbps–, and its support of 3R functions [63]. The cost of an all-optical XOR gate is 3 cost units. In addition, all nodes, both Multi-Layers (MLs), i.e., IP/MPLS routers connected to WRs, and WRs (without cross-layer connections) have All-Optical NC features, even though not all nodes code traffic.

- The cost of a $50\,GHz$ fixed-grid ROADM/OXC node with a capacity of 80 channels is obtained using the Equation (2.23).
- WDM transponders with 100 Gbps and 2000 km of distance reach. The cost of the used WDM transponders is 15 cost units.
- The power consumption for all-NC features is negligible due to its low consumption in comparison with other features, e.g., data transmission, processing.

$$C_{oxc} = d(W_{ss} + A_b + A_p) + 2dI + 4dA_{wg} + 2G_x \tag{2.23}$$

Techno-Economic Study of the IP/MPLS Layer

In this section numerical results related to the evaluation of proactive protection schemes and their impact on both CAPEX and OPEX of a network provider are introduced.

The evaluated proactive protection schemes are: DP, DPNC and DPNC*. It was assumed that these protection schemes are deployed at the IP/MPLS layer. The metrics to be evaluated are: 1) The IP/MPLS P_{cost} (the total amount of IP bandwidth required to enable link protection); and 2) the CAPEX and OPEX required to enhance the ML nodes for DP, DPNC and DPNC* schemes, deployed at the IP layer on the network topology shown in Fig. 2.13

In Fig. 2.14 we show the percentage of IP/MPLS P_{cost} for a time window of 4 years. As it can be observed, in year 1 a DP scheme already exceeds the (50%) protection-threshold level, whereas for DPNC and DPNC* schemes is shifted till two years later. However, the P_{cost} of DPNC* is less compared to DPNC and DP schemes.

On the other hand, Fig. 2.15 depicts the cost evolution for the IP/MPLS Router Line Cards. The IP/MPLS Router Line Cards cost evolution has also been estimated with the forecasting model introduced in the previous section. Notice that for a DP scheme, in year 1, the IP/MPLS P_{cost} already exceeds by 11% the protection-threshold. Therefore, it is required the investment of a 16% in terms of capacity (assuming a 5% safe margin) or an investment of 322 cost units in order to not exceed the protection-threshold. However, no investment is necessary in year 1 for both DPNC and DPNC* schemes, rather in both cases this would be only required in year 3, measured in terms of 164 and 125 cost units for both schemes respectively.

Based on the results depicted in Fig. 2.15, it can be stated that with NCP based on a DP strategy (DPNC or DPNC*) the CAPEX required to add new IP/MPLS Router Line Cards is delayed two years. Moreover, DPNC* requires 72% less CAPEX compared with the required by a DP scheme in the first year, and nearly 5% less than a DPNC scheme in the last year.

Finally, to properly evaluate the impact of the evaluated protection schemes on the OPEX, it was conducted a power cost analysis. In light of this, Fig. 2.16 shows the overall power consumption of the three evaluated schemes for a 4 year period. As expected, the reduction in the network capacity allocated for NCP schemes protection has a direct impact on the power

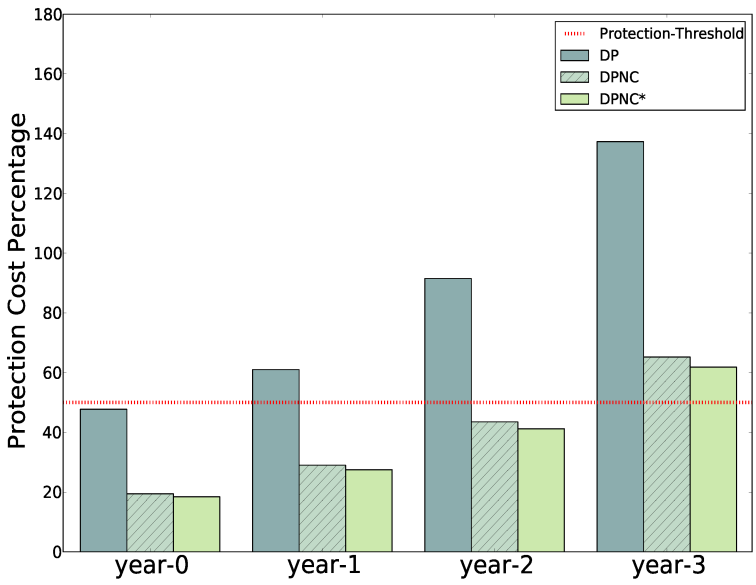


Figure 2.14: IP/MPLS P_{cost} over the total network capacity.

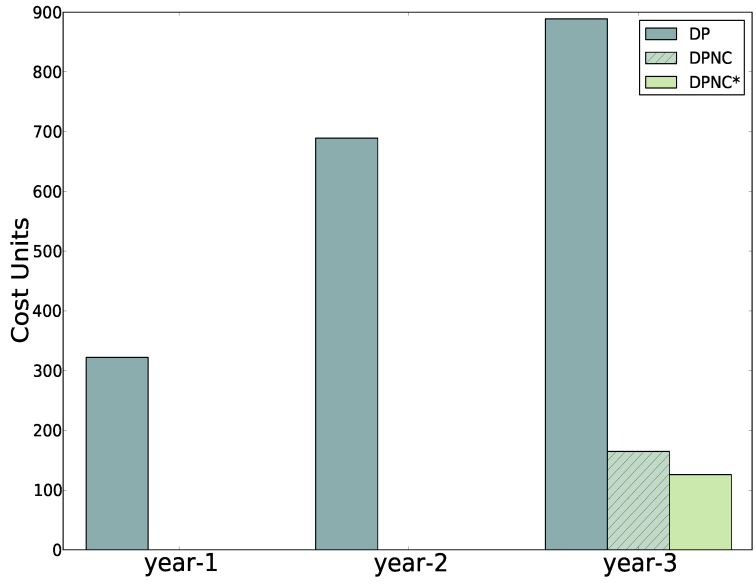


Figure 2.15: CAPEX of the IP/MPLS layer.

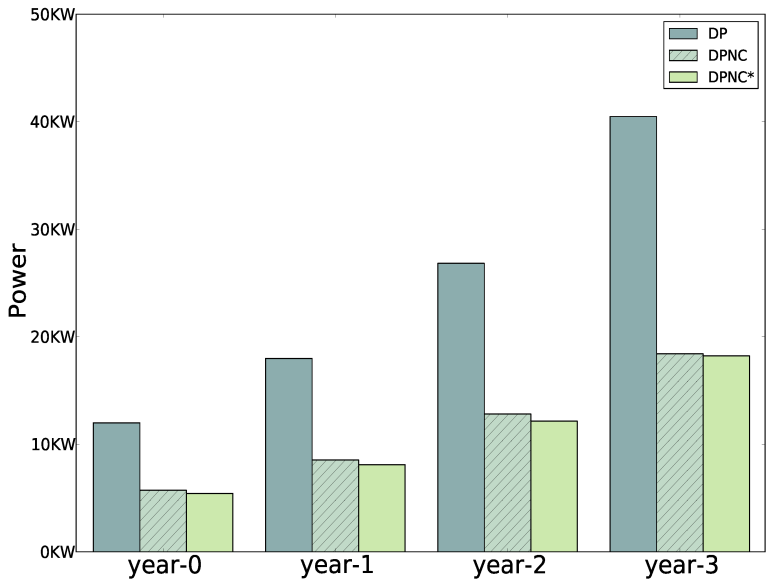


Figure 2.16: OPEX of the IP/MPLS layer

consumption, since less transceivers and less IP/MPLS line cards are required. Notice that the Power Consumption of a DPNC* scheme is 52% power compared with a DP scheme.

Techno-Economic Study of the Optical Layer

This section focuses in the techno-economic study of proactive protection schemes deployed at the Optical layer. To this end, two evaluation tests are performed: 1) The percentage of network resources (optical wavelengths) allocated to protect the traffic sent along optical links (Optical P_{cost}); and 2) The CAPEX related to the Optical layer, i.e., the cost required by transceivers and WDM transponders.The first evaluation test assesses the Optical P_{cost} as shown in Fig 2.17. From the results depicted in this figure it can be concluded that when using DPNC* the Optical P_{cost} is reduced 13% and 3% compared with DP and DPNC respectively.

The second evaluation test depicts the CAPEX required by each evaluated scheme, see Fig.2.18. From the results depicted in this figure it can be concluded that DPNC* requires the lowest CAPEX compared with DP and DPNC schemes. Notice, that DPNC* requires 49% less CAPEX compared with the DP scheme in year 1; despite of the cost required to enable NC features.

To the best of our knowledge, this section introduces the first techno-economic study evaluating NCP schemes in multi-layer scenarios. Based on the obtained results, it can be concluded that by means of the multiple-coding features, both CAPEX and OPEX can be substantially reduced independently of the network layer where DPNC* is to be deployed.

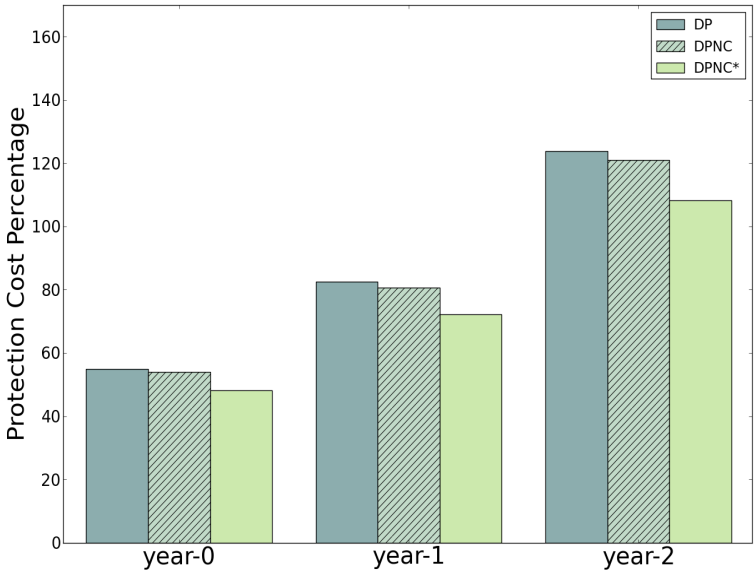


Figure 2.17: Optical P_{cost} over the total of network resources.

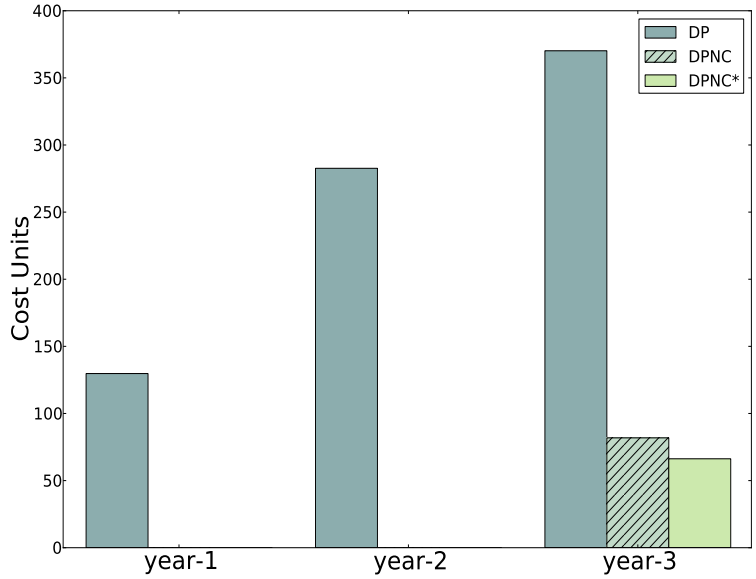


Figure 2.18: CAPEX of the Optical layer.

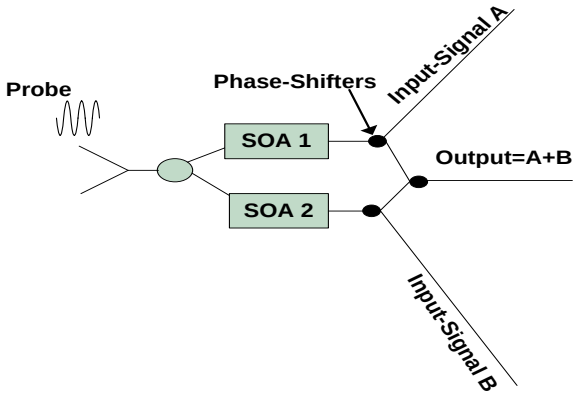


Figure 2.19: All-optical XOR architecture.

Indeed, using DPNC* an average CAPEX reduction of 60.5% can be achieved independently of the network layer technology. For instance, 49% and 72% of CAPEX reduction is obtained when deploying DPNC* at the Optical and the IP/MPLS layer respectively. On the other hand, a 52% of OPEX reduction is obtained at the IP/MPLS layer.

Therefore, since NC operations are supported at both IP/MPLS and Optical layers, the network layer where an NCP scheme will be deployed, may be selected according to the specific requirements of a network operator.

2.3.6 Implementation Issues with regard to NCP

In this section we discuss the technical issues related to the implementation of NCP schemes. Conventional protection schemes such as DP have been widely and successfully deployed in real optical network scenarios [66]. In light of this, it would be reasonable to suppose that the deployment of NCP schemes based on a DP strategy seems feasible in the coming years.

A key issue regarding the deployment of NCP schemes refers to the execution of NC (XOR) operations; in particular considering that the deployment of optical NC operations is more complex in comparison with electrical NC. In the optical domain, the implementation of all-optical XOR gates is widely studied in network research [67], [63]. Typically, the building components of All-Optical XOR gates are Semiconductor Optical Amplifiers (SOAs), see Fig. 2.19, supported by the fact that SOAs offer low-power consumption, easy deployment and short-latency.

The execution of All-optical NC operations can be done at line speed for transmission above 10 Gbps and up to 100Gbps with modulation schemes such as QPSK. Therefore, from a practical perspective, the deployment of NCP schemes in a near future seems feasible. It

has been already shown [63] that the practical implementation of optical XOR operations for optical signals with different modulation schemes such as BPSK and QPSK is also possible under test lab scale. However, the all-optical XOR of other modulation schemes needs further study.

On the other hand, another issue to be considered for the deployment of NCP schemes, specifically to Optical technologies such as Optical Burst Switching (OBS), is the utilization of all-optical buffers. By means of all-optical buffers it is possible to cope with the delay added by the coding of data streams with different bitrates. For more information related to all-optical buffers the reader is referred to [68], [69].

Finally, another issue hindering the deployment of NCP schemes is related to signal features such as phase tracking. The degradation of signal phase stability caused by the long fiber signal transmissions as well as by Physical Layer Impairments (PLI) factors has a strong impact on the performance of coherent signal distribution [70]. This affects both coding and decoding functions require for the correct operation of an NCP scheme. Long optical fiber transmissions are typical in multi-domain scenarios. The deployment of NCP schemes in multi-domain scenarios is out of the scope of this thesis and it is left as a future research trend.

2.4 Online RWA

This section is devoted to the study of online RWA algorithm under the presence of inaccurate NSI. An online or dynamic scenario is the one where a Connection Request (CR) arrives in a random manner. In the rest of this document the words dynamic and online scenario are used interchangeably.

In online scenarios, lightpaths are continuously setup and tear down on a short-term basis. This dynamism might severely affect the performance of an RWA algorithm. This negative effect is mainly motivated by two issues: 1) the connection setup delay; and 2) the inaccuracy of the NSI.

The study of RWA in the context of online scenarios is gaining momentum in network research, motivated by the fact that nowadays FI applications, such as Video on Demand (VoD) or DCNs, demand huge bandwidth and seamless connectivity in an agile manner, which undoubtedly can overload the network with a high volume of Connection Requests (CRs). The following subsections focus on the study of RWA algorithms on online scenarios with inaccurate NSI.

2.4.1 The Routing Inaccuracy Problem

The availability and accuracy of NSI have a profound impact on both performance and scalability of source-RWA algorithms. Indeed, inaccurate NSI might result in sub-optimal path selections that potentially lead to an increase of the blocking probability. The main sources of

inaccurate NSI are: 1) the NSI aggregation caused by a hierarchical network design, which is a common case in DCN scenarios in order to improve scalability; 2) the non-negligible delay propagation, which are commonly generated by Physical Layer Impairments (PLI) factors; 3) the failure of control messages, which can be caused by malfunctioning nodes, and; 4) the updating policy that determines when the NSI should be disseminated, which often follows a periodic behavior.

The accuracy of NSI is important to meet the WCC, which states that a lightpath can be solely established if the same wavelength is available on the path selected from the source to the destination WR pair. Therefore, the accuracy of NSI, particularly reporting about overall wavelength availability is significantly important, since suboptimal lightpath selections might increase the amount of blocked connections.

Among the main vehicles to offset the negative effects caused by inaccurate NSI we can mention the following: i) the use of multi-fiber systems; ii) enhancing WR nodes with wavelength conversion capabilities–Wavelength-Convertible Routers (WCR); and, iii) decreasing the time interval to disseminate NSI, hereinafter referred to as the update time.

The simplest one (from a CAPEX perspective), is reducing the update time. However, this can be cumbersome because of signaling overhead (update NSI messages) concerns that lead to scalability issues. Moreover, it is important to notice that even with unrealistic update times, i.e., flooding update messages per network state change the NSI might still be inaccurate.

On the other hand, the use of both multi-fiber systems and WCRs tend to be overlooked because of the added high costs, technical difficulties and power consumption. Nevertheless, the deployment of multi-fiber systems is a more widespread practice in comparison with WCRs.

A pioneer work related to the study of the RI problem in CGNs using optical technologies (WRNs) for the transport medium can be found in [38], where authors analyze the performance of conventional RWA algorithms under inaccurate NSI. This study successfully positions the RI problem in WRNs by effectively demonstrating that in highly dynamic large scenarios under inaccurate NSI the performance of conventional RWA algorithms significantly decreases, i.e., the blocking probability increases. This study also shows that RWA algorithms such as First-Fit (FF) combined with shortest-path routing, that are considered optimal under accurate NSI conducts suboptimal performance in comparison with other schemes such as Random Wavelength Assignment (RR), i.e., FF is suboptimal under inaccurate NSI. More recent studies propose analytical models for evaluating the performance of source and destination based RWA algorithms under inaccurate NSI caused by propagation delay [71].

Another important study dealing with the RI problem can be found in [14], where the authors focus on the RI problem for IP networks considering delay and bandwidth constrained applications. To this end, the authors propose probabilistic models to express the inaccuracy of NSI. More recent works such as [72] propose a so-called Bypass-Based Routing. The Bypass-

Based Routing is based on a innovative metric utilized to model uncertainty combined with a mechanism that bypasses congested links –whenever a path selection is sub-optimal.

The proposals dealing with the RI problem in WRNs can be categorized into three main approaches.

1. ***Update Policies***. These proposals do not focus on the design or improvement of RWA algorithms, rather propose novel updating policies aiming at minimizing the signaling overhead while attempting to maintain accurate NSI [73], [74]. Conversely to the conventional periodical dissemination of NSI, an updating policy triggers NSI updates according to a certain policy, e.g., an NSI update is disseminated when the residual capacity (the available wavelengths on a link), is less than a certain threshold. A handicap of this approach is its disruption in the current flooding NSI mechanisms.

2. ***RWA based on global NSI.*** These proposals do not attempt to minimize the signaling overhead. Instead, they assume a global NSI scenario –an unrealistic assumption for source RWA using distributed control planes– and propose to enhance WRs with rerouting capabilities. For instance, in case a lightpath cannot be established, an intermediate WR selects an alternative lightpath, i.e., connection reattempts are performed [75], [72]. A handicap of this type of approach is that it may lead to high connection setup times due to reattempts. This issue should be avoided in CGNs demanding stringent constraints such as provisioning times within hundreds of milliseconds.

3. ***RWA based on local NSI***. Under this approach the dissemination of NSI is confined solely to network topology changes. Common examples of these schemes are Predictive RWA algorithms which rely on prediction techniques –successfully used in the computer architecture field [76], supported by predictive counters used to model lightpaths availability. Some contributions such as [19], [77] propose Predictive RWA algorithms that minimize both signaling overhead and blocking probability.

As mentioned in this section, a possible strategy that can be adopted in order to offset the effects of inaccurate NSI caused by periodic updating policies is to decrease the update time. Nevertheless, this strategy might be cumbersome and can lead to signaling overhead issues.

In order to illustrate the negative effects on the signaling overhead caused by low update times, consider the network scenario shown in Fig. 2.20. This scenario represents a smaller fragment of a full-mesh Clos fabric network topology often used within DCNs, hereinafter referred to as DCN topology, implementing a flooding update mechanism for NSI dissemination, as is the case for the widely used OSPF-TE protocol [78]. It is worth mentioning that all WRs within a spine are connected –not shown in Fig. 2.20.

According to a flooding update policy, when a lightpath is provisioned between the WR pairs S and D using wavelength λ_1, WR A as well as other WRs within spine 2, will receive 4 update messages sent by spine 1. Then, each WR in spine 2 will reflood all received update

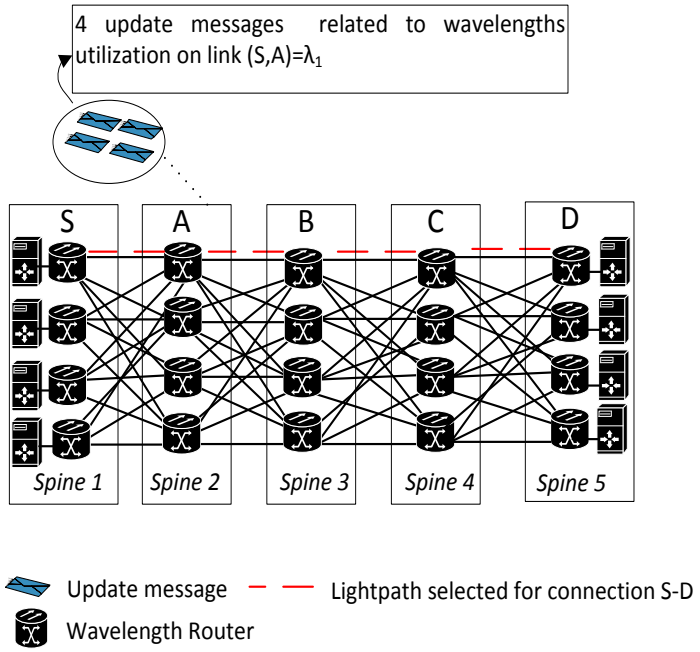


Figure 2.20: Signaling overhead issues related to the update time.

messages to WRs in spines 3, 4, and 5. Therefore, a total of 16 update messages will be generated in the network. This might be cumbersome in highly dynamic large scenarios due to signaling overhead issues, where connections are provisioned and released in a short-term basis. Conversely, low flooding rates (high update times) might not reflect the actual (real) network state, what unquestionably drives higher blocking probability rates due to the inaccuracy related to NSI. Moreover, regardless the updating policy strategy to be deployed there is not any guarantee about the degree of NSI accuracy; hence motivating a non-negligible ratio of connection blocking that cannot be ignored.

On the other hand, another source of inaccurate NSI is the aggregation caused by a hierarchical network design. Hierarchical topologies are becoming a widespread practice in DCN design because of scalability issues related to signaling overhead and network size increase [79]. However, as shown in Fig. 2.21, the aggregation imposed by a hierarchical topology design has a collateral (negative) effect, namely the NSI is non-complete because it does not contain low-granularity information about wavelengths availability, i.e, wavelengths availability per link. For instance, in the scenario shown in Fig. 2.21, WR *S* does not have the wavelength availability for all links within a network segment, e.g., Segment 2.

Other sources introducing inaccuracy on the NSI are non-negligible delay propagation, and the failure of control (OSPF-TE and RSVP-TE) messages. The former factor might have an impact on both source and destination-based routing and it is mainly caused by PLI factors jointly with network topology characteristics such as large diameter –it is worth mentioning that large networks do not necessary have a large diameter. The latest factor can be caused

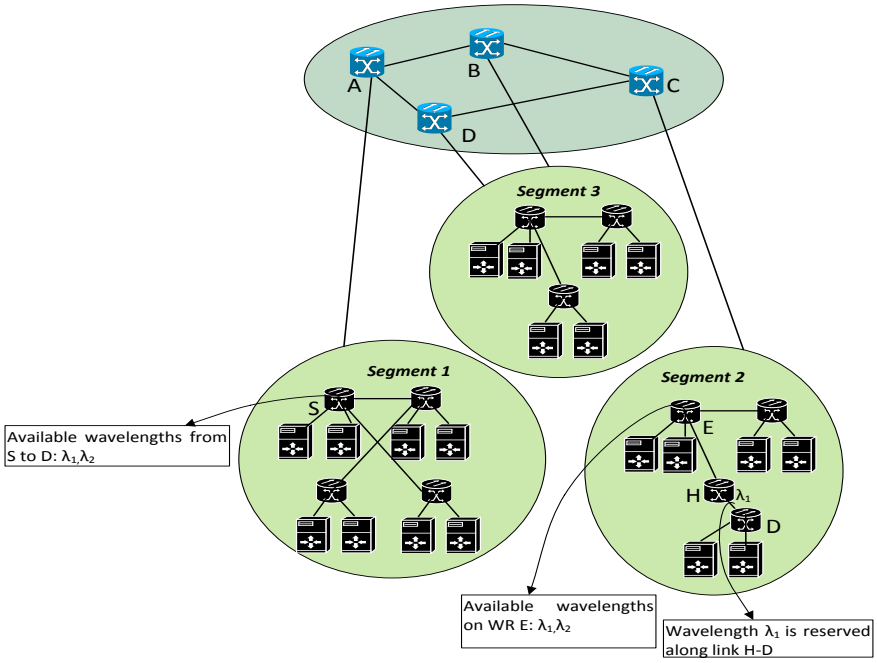


Figure 2.21: Negative effects of inaccurate NSI due to the aggregation imposed by a hierarchical network design.

by a malfunctioning WR. Albeit this rarely occurs, it can potentially have an impact on both distributed and centralized RWA algorithms.

On the other hand, it is worth mentioning that although the RI problem also affects centralized control architectures such as a PCE or SDN, these architectures are out of the scope of this thesis with regard to the RI problem.

In fact, the main reason positioning the focus on source-RWA approaches is the recent concern in speeding up the packet-optical integration in DCNs by adopting centralized control architectures such as OpenFlow for the packet domain; whereas the optical domain remain under distributed control architectures such as GMPLS or ASON (since OpenFlow requires to be further enhanced to support optical capabilities), being source-RWA the option commonly preferred by this type of control planes [80], [81]. For more information the reader is referred to [82].

2.4.2 Hybrid Prediction based Routing

In this section, it is proposed a source-based RWA algorithm for addressing the routing inaccuracy problem in WRNs **without considering the WCC**. Therefore, the proposed RWA algorithm can be easily extended for IP/MPLS networks since the only constraint related to the blocking of a lightpath is the available bandwidth. The symbols used in this section are listed in Table 2.9.

Table 2.9: List of Symbols and Terminology for Section 2.4.2

Symbols	Meaning
$G(V, E)$	Directed graph where V is the set of WRs and E is the set of optical links.
$p_i^{(s)}$	Predictive counter of link i locally computed by a WR s, where $s \in V$, $i \in E$, and $p_i^{(s)} \in \{0,1,2,3\}$.
$\mathscr{L}_j^{1(s)}$	Availability of route j locally computed by a WR s for a Moderate-Dynamic scenario.
$\mathscr{L}_j^{2(s)}$	Availability of route j locally computed by a WR s for a Highly-Dynamic scenario.
$v_i^{(s)}$	Vulnerability degree of link i locally computed by a WR s.
$b_i^{(s)}$	The residual bandwidth of link i locally computed by a WR s.
ϵ	Predefined threshold defining the degree of inaccuracy tolerated by HPBR.
λ	A wavelength unit.
b_{req}	Optical bandwidth demanded by a CR.
$C_j^{1(s)}$	The cost of route j locally computed by a WR s for a Moderate-Dynamic scenario.
$C_j^{2(s)}$	The cost of route j locally computed by a WR s for a Highly-Dynamic scenario
N_j	Number of hops along route j.
$V_j^{(s)}$	Vulnerability of route j locally computed by WR s.

The proposed RWA algorithm is referred to as Hybrid Prediction-based Routing (HPBR). HPBR also exploits the use of prediction techniques but differs from the predictive RWA algorithms proposed in [19], [83], [77] in several aspects.

First, the previus RWA algorithms assume coarse-granularity counters (prediction counter

per route), i.e., a two-bit counter (assigned locally by each node) for each route. HPBR adopts a finer granularity approach. To this end, each WR keeps track of links availability by means of a two-bit counter–prediction counter per link. Secondly, HPBR is not bounded to a unique routing metric; it dynamically selects the most suitable metric according to the network scenario.

HPBR makes use of two-bit counters to predict routes availability. A counter value is increased as follows. $p_i^s = p_i^s + 1$ only when a CR along link i is blocked, i.e., during the lightpath reservation the connection cannot be provisioned due to the lack of bandwidth, and $p_i^s < 3$. Conversely, a counter value is decreased as follows. $p_i^s = p_i^s - 1$ only when a CR along link i is successfully provisioned and $p_i^s > 0$.

The rationale driving the adoption of two-bit predictive counters is because two-bit counters are enough to keep historical behavior of routes or in the case of HPBR of the links. Otherwise, a low counter value is unable to properly model routes availability, whereas higher values add a high degree of hysteresis which might cause sub-optimal paths selections –driven by the inertia generated by high counter values.

Whenever a route is evaluated, the availability of this route is locally computed by each WR according to the offered load conditions. For moderate offered loads (Moderate-Dynamic scenarios) the availability of a route is computed as shown in Equation (2.24), where a high value of $L_j^{1,s}$ means that route j may be unavailable, the contrary occurs with low values.

$$L_j^{1,s} = \sum_{i \in j} p_i^s v_i^s \tag{2.24}$$

Moreover, the vulnerability degree (v_i^s) is a concept introduced by authors in [83] and used to model whether an optical link may lead to a connection blocking. The vulnerability degree of a link is computed as shown in Equation (2.25).

$$v_i^s = \begin{cases} 1, & if\ and\ only\ if \left(1 - \frac{b_{req}}{b_i^{(s)}}\right) < \epsilon \\ 0, & otherwise \end{cases} \tag{2.25}$$

The vulnerability degree of a link is selected according to the pre-defined threshold (ϵ), so-called blocking factor. Thus, the blocking factor parameter must be properly set according to the offered load conditions.

In addition, notice that a predictive counter value is only considered for computing the route availability whenever the link is considered vulnerable. Therefore, even though the value of a predictive counter of a link is greater than zero, if this link is not vulnerable its predictive

counter does not affect the availability of a route.

On the other hand, for high offered loads (a Highly-Dynamic scenario) the availability of a route is computed as shown in Equation (2.26). Notice that in this case, the vulnerability degree is not taken into account for computing the availability, $L_j^{1,s}$. Moreover, the hysteresis degree of a link counter is decreased by means of changing the update policy of the two-bit counters: $p_i^s = p_i^s + 2$ solely when a connection along link i is blocked and $p_i^s < 2$; whereas $p_i^s = p_i^s - 2$ solely when a connection along link i is successfully provisioned and $p_i^s > 0$.

$$L_j^{2,s} = \sum_{i \in j} p_i^s \tag{2.26}$$

To clearly explain the reasons that motivated the adoption of a fine-granularity approach (two-bit counter per link)for the predictive counters conversely to previous proposals (two bit counter per route), such as Prediction-based Routing (PBR) and Fuzzy-based Routing (FRA), let's consider the topology depicted in Fig. 2.22a where source-based routing is assumed and the offered load is moderate. A connection request (CR_1) needs to be provisioned demanding the allocation of 4λ with WRs S and D as endpoints. As a consequence, WR S selects the path $S-1-2-D$ for CR_1. However, this path cannot be provisioned due to lack of bandwidth along this route. This occurs because of the inaccuracy of the NSI in WR S, which reflects b_{1-2}^s with a residual capacity of 8λ. However, the real residual bandwidth of link $1-2$ is 3λ, less than the bandwidth requested by CR_1, i.e., the local NSI stored in WR S is inaccurate. If predictive counters per route (coarse-granularity counters) are used, then the predictive counter of route $S-1-2-D$ is increased.

Consider now the scenario depicted in Fig. 2.22b, where a second connection request (CR_2) reaches WR S with similar characteristics (5λ) on the requested bandwidth and WRs (S and D as endpoints) as CR_1. In order to provision CR_2, WR S avoids selecting route $S-1-2-D$, due to its predictive counter value and attempts to provision this connection along path $S-1-2-3-D$. However, CR_2 cannot be provisioned because link $1-2$ does not have enough bandwidth to allocate CR_2 as it was the case for CR_1. Therefore, coarse-granularity predictive counters do not capture the unavailability of link $1-2$ and this is why route $S-1-2-D$ is shown as available.

Fortunately, contrary to coarse-granularity predictive counters, the use fine-granularity predictive counters can predict the unavailability of link $1-2$. In light of this, consider the scenario depicted in Fig. 2.22c. When CR_2 reaches WR S, this one captures the unavailability of route $S-1-2-3-D$. This is because the counter of link $1-2$ (a link part of route $S-1-2-3-D$) is not 0; hence, route $S-4-5-6-D$ is selected since all links belonging to this route have their predictive counter values on 0. Therefore, based on this illustrative example it can be stated that the use of fine-granularity predictive counters reduces the blocking probability.

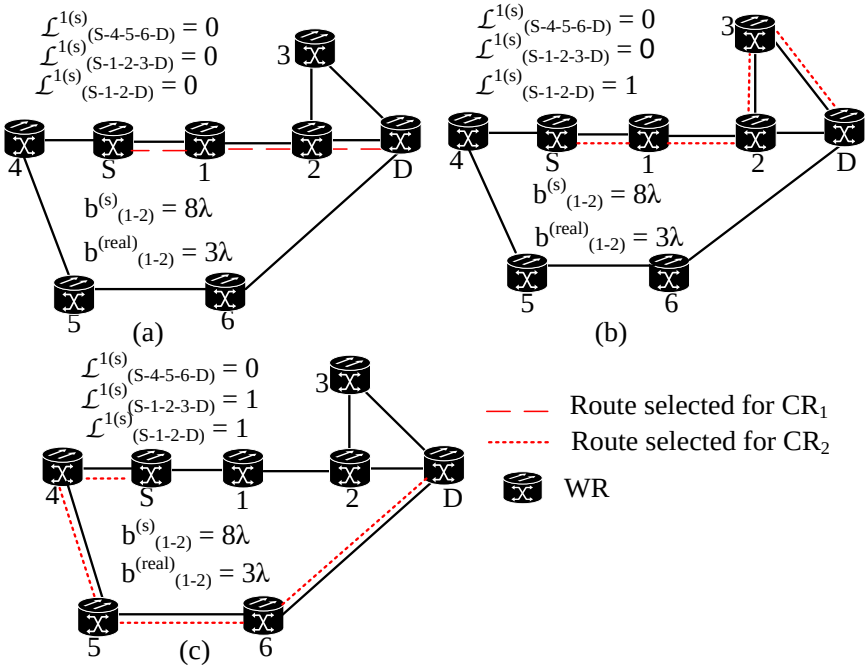


Figure 2.22: Negative effects of inaccurate NSI: a) and b) Coarse-granularity predictive counters; c) Fine-granularity predictive counters.

As such, HPBR uses fine-granularity predictive counters jointly with two cost metrics in order to select a route. On one hand, the first cost metric, computed as shown in Equation (2.27), is an enhancement of the metric presented by [83]. This cost metric can be categorized as a dynamic metric since has high dependency on global NSI (hence, potentially inaccurate) such as residual bandwidth and link vulnerability. Therefore, this cost metric is used for Moderate-Dynamic scenarios, since the NSI is less susceptible to inaccuracy under load offered loads. Notice that b_{min} is the minimum bandwidth available on the links forming route j, where $b_{min} = min\, b_i^s \; \forall i \in j$. The parameter b_{min} is locally computed by each source WR.

$$C_j^{1,s} = \frac{L_j^{1,s}}{N_j - 1} + N_j \left(\frac{1}{b_{min}} \right) \tag{2.27}$$

Moreover, since the availability of a route (left term of Equation (2.27)) usually has more weight than the other parameters, we propose to divide it by the route length in order to balance the weight of each parameter. In addition, in case that a route vulnerability is 0, HPBR does not rely solely on the predictive counter value to compute $C_j^{1,s}$, see the right term of Equation (2.27). This issue was not addressed by authors in [83].

On the other hand, the other cost metric used by HPBR, cf. Equation (29), can be catego-

rized as a quasi-permanently metric since it has a low dependency of variable NSI, i.e., it avoids both vulnerability and bandwidth parameters. This cost metric is used for Highly-Dynamic scenarios.

$$C_j^{2,s} = \frac{L_j^{2,s}}{N_j - 1} + N_j \tag{2.28}$$

In order to select a route for a given CR, HPBR evaluate all candidate paths using both cost metrics shown in Equations (2.27) and (2.28). To reduce path computation complexity, the candidate paths (each source WR has exactly 4 candidate paths per source-destination pair) were computed offline using Dijkstra algorithm with the number of hops as the cost metric. The mechanism used by HPBR to select either $C_j^{1,s}$ or $C_j^{2,s}$ depends on the offered load. This mechanism is discussed in the following lines as well as the rationale driving the use of an hybrid cost metric approach.

Based on extensive simulations, the obtained results lead to consider that for Moderate-Dynamic scenarios, a dynamic metric as the one shown in Equation (2.27) provides good performance in comparison with other routing schemes. However, for Highly-Dynamic scenarios a quasi-permanently metric based solely in static NSI such as the one shown in Equation (2.28) exhibits a better performance. This is because as the inaccuracy related to NSI (in particular the overall wavelengths availability) increases (since the available bandwidth is rapidly changing due to high frequency of lightpaths being set-up and torn-down) it is better to rely on routes that span less hops in order to use less bandwidth, and add less inaccuracy. Indeed, the simulation results presented in the next section validate that in Highly-Dynamic scenarios selecting paths that span several hops might potentially lead to suboptimal performance. Therefore, for Highly-Dynamic scenarios it is better to rely on local NSI such as the number of hops along a path. This has been discussed by authors in [84]. Nevertheless, authors do not consider inaccurate NSI.

In order to select which cost-metric will be used, HPBR gathers NSI such as destination and arrival time of CRs in order to evaluate the offer load in the network. The overall procedure of how this is done is elucidated in Algorithm 2. As it can be observed, if a WR receives a high amount of CRs during a short period of time (determined by variable $Delta - value$ in Algorithm 2) and the destination WRs (variable $Destinations$ in Algorithm 2) of these requests are highly heterogeneous, then HPBR computes a route cost based on Equation (2.28), otherwise it uses Equation (2.27). In Algorithm 2, the parameter th_1 determines the amount of collected information; whereas th_2 specifies when the condition of CRs with highly-heterogeneous destinations is met; finally, th_3 specifies temporal proximity of the CRs. The threshold values of th_1, th_2 were set to 20 and 8 respectively, whereas th_3 value was set to 4 time units.

Algorithm 2 Overview of the metric adaptation mechanism of HPBR.

Input: $(dest, b_{req}, timestamp)$
Output: $(metric)$

{$dest, b_{req}$ and $timestamp$ are the destination, bandwidth, and arrival time of the requested lightpath respectively.}
if $H.size() < th_1$ **then**
  $H.append(dest)$ {H is a set containing different lightpath's destinations, the maximum size of H ($H.size()$) is determined by th_1, $append()$ is a method that inserts the specified content into a given set. }
  $RT.append\big(timestamp\big)$
else
  $Destinations = distinct(H)$ {$distinct$ is a method returning a set of distinct elements of a set (H). }
  for i in range $(0, size(RT) - 1)$ **do**
    $Delta.append\big[abs\big(RT_{[i+1]} - RT_{[i]}\big)\big]$ {rate of change of the connection requests.}
  $Delta\_value$=$Distinct2(Delta)$ {$Distinct2$ is a method returning the amount of elements with values less than th_3, where th_3 is a predefined threshold.}
  if $|Destinations| > th_2$ and $Delta\_value > 0.7 \times th_1$ **then**
    $metric = Equation(5)$
    $Set\ Increment/Decrement\ Values\ of\ counters$
  else
    $metric = Equation(4)$
    $Set\ Increment/Decrement\ Values\ of\ counters$
  RT, $H = \emptyset$ {Reinitialize H and RT}

2.4.3 Simulation Results with regard to Hybrid Prediction based Routing

In this section, we present the simulation results obtained by the proposed scheme namely HPBR versus other similar predictive routing schemes available in the literature, such as PBR, FRA, Balanced Vulnerable Predictive Path (BVP2) –BVP2 was adapted for WRNs–, the well-known Least-Congested Path (LCP) [26], which among the evaluated schemes is the only one that requires periodically dissemination of NSI, and finally these proposals are compared with a routing scheme referred to as HOPS, which uses a routing metric based solely on the number of hops along a path.

In order to ensure realistic findings, we build a simulation model for a WRN based on the NSFNET topology (see Fig. 2.2a) using the well-known network simulator Omnetpp [85]. All the plotted values have a 95% confidence interval not larger than 0.5% of the plotted values. The traffic models for the evaluated scenarios are as follows.

- **Moderate-Dynamic**. In this scenario the Connection Requests Arrivals (CRA) for each source WR are as follows: $CRA_1(t_1)$, $CRA_2(t_1 + t)$,... $CRA_n(t_{n-1} + t)$, where t and t_n are Poisson-distributed. This scenario has a moderate offered load since CRs arrive in an incremental manner.

- **Highly-Dynamic**. In this scenario the CRA for each source WR are as follows: $CRA_1(t_1)$, $CRA_2(t_2)$,... $CRA_n(t_n)$, where t and t_n are Poisson-distributed.

In addition, the following settings are assumed:

- The Holding time per connection is exponential distributed with a mean never exceeding ten times the inter-arrival time.
- The average bandwidth requested per connection is Poisson-distributed with a mean never exceeding the 10% capacity of an optical link respectively.
- WRs with 80 channels on a $50\,GHz$ fixed-grid, which is one of the channel spacing standards defined by the International Telecommunication Union (ITU).
- The time required by a connection setup is neglected.
- Reattempt is not done once a connection is blocked.
- All WRs support full wavelength conversion.
- Single fiber system.
- Once a connection is established it cannot be reconfigured during its lifetime.

On one hand, Fig 2.23. shows the blocking probability versus distinct update times for a Moderate-Dynamic scenario. In this scenario, LCP and HPBR present a good performance. However, the performance of HPBR is not affected by the update time (as well as all the predictive schemes), as it is the case with LCP. This is because for HPBR the dissemination of NSI information is solely restricted to topological changes. Notice that the performance of a prediction routing scheme based on coarse-granularity predictive counters such as PBR is not as optimal as HPBR. Other similar schemes such as BVP2 and FRA yield a better performance, but their performance is also not as good as HPBR.

On the other hand, Fig.2.24 shows the blocking probability versus distinct update times for a Highly-Dynamic scenario. The purpose of this trial is to simulate a network heavily loaded of CRs. For this purpose, a holding time ten times higher than the inter-arrival time is selected, in order to increase the number of active connections in the network at any time.

Based on the results shown in Fig. 2.24 it can be seen that in Highly-Dynamic scenarios the performance of routing schemes relying on Global NSI, such as LCP is significantly degraded. This is not the case for predictive routing schemes. Among all predictive routing schemes, HPBR yields the best performance.

Finally, Fig. 2.25 depicts the simulation results related to a mixture of a Moderate-Dynamic and a Highly-Dynamic scenario, where the total amount of possible destinations is increased –increase the number of active connections at any moment– for a CR. In this scenario, HPBR switches from computing routes as shown Equation (2.27), to computing routes using Equation (2.28), once it detects (according to Algorithm 2) that the network conditions entail a different network scenario –based on the offered load.

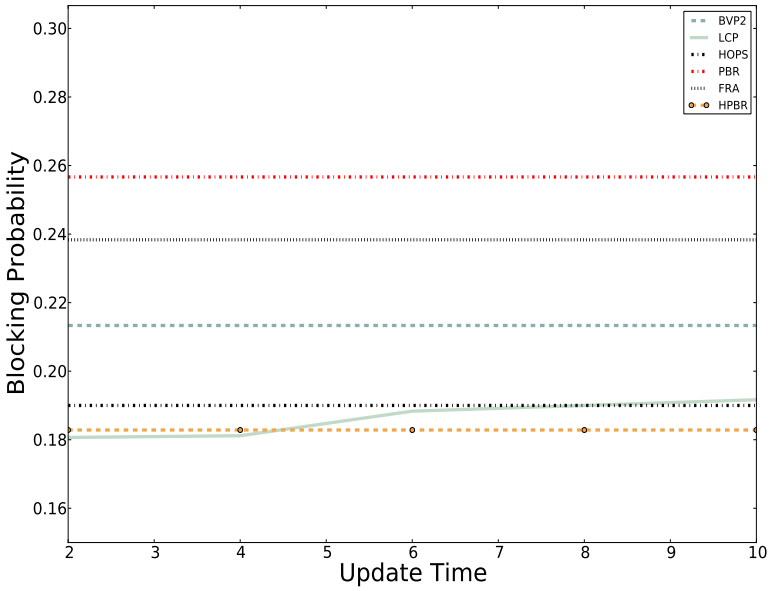


Figure 2.23: Blocking probability for a Moderate-Dynamic scenario with 100 requests per WR, 6 WRs as sources and 14 WR as destinations, average holding time and inter-arrival time of 4 units; average b_{req} =10% of total link capacity; and ϵ=5%.

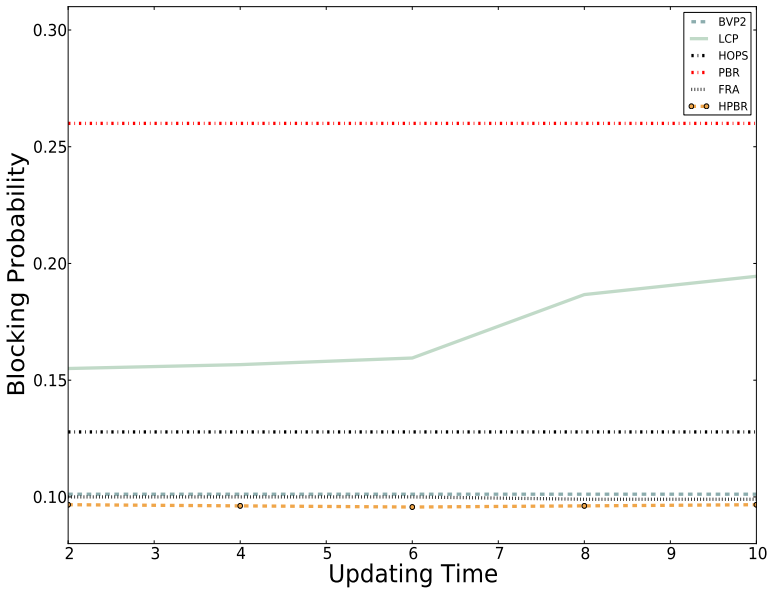


Figure 2.24: Blocking probability for a Highly-Dynamic scenario with 100 requests per WR, 6 WRs as sources, average holding time and inter-arrival time of 100 units and 10 units respectively; average b_{req}=2% of total link capacity; and ϵ=5%.

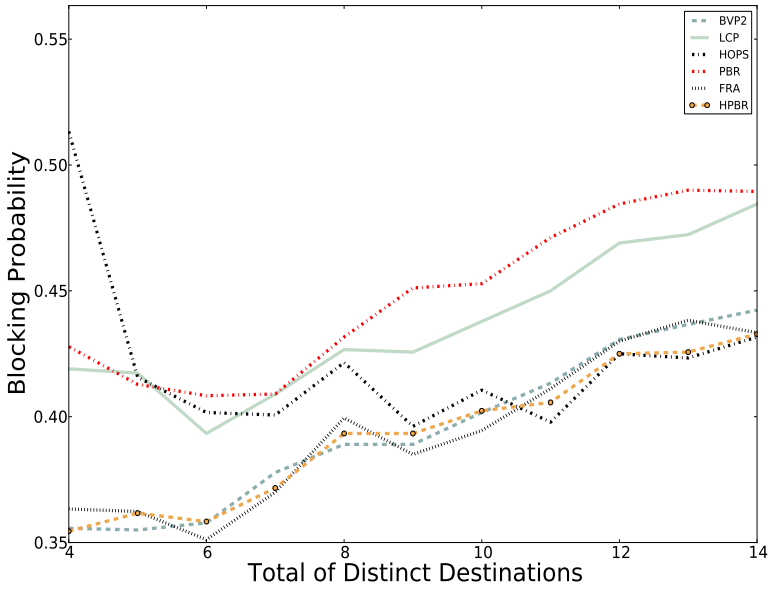


Figure 2.25: Blocking probability for a mixture of a Moderate and a Highly-Dynamic scenario with 200 requests per WR, 3 WRs as sources, average holding time and CRA time of 4 units respectively; average b_{req}=5% of total link capacity; and ϵ=5%.

It is important to notice how the performance of a static routing strategy, such as HOPS, improves as the inaccuracy degree increases, i.e., more active connections at any moment. This behavior was validated by a similar approach followed by authors in [84]. This rationale motivated us to incorporate a hybrid metric system for HPBR.

Notice that for a low amount of distinct destinations FRA shows a low blocking probability, but for higher amounts its performance is reduced. The opposite occurs with the HOPS scheme. However, HPBR shows low blocking probability independently of the network scenario type.

A lesson learned from the obtained simulation results is that fine-granularity predictive counters are the best option for addressing the routing inaccuracy problem in WRNs. Indeed, the obtained results validate that the proposed scheme shows a better performance related to the blocking probability in distinct dynamic scenarios.

2.4.4 Finer Prediction based Routing

As mentioned in this thesis, an important constraint limiting the performance of RWA algorithms –specifically in WRNs without wavelength conversion capabilities referred to as Wavelength-Selective (WS) networks– is the so-called WCC, which states that a lightpath can be solely established if the same wavelength is available on the path selected from the source

to the destination WR pair.

The rationale driving this section focuses on the RI problem caused by both large periodic updating policies and high offered loads, **considering the WCC**. Thus, the envisioned scenario considers WRNs without wavelength conversion capabilities under dynamic traffic, assuming both single and multi-fiber systems. This section proposes a distributed source-based RWA algorithm addressing the RI problem, referred to as Fine Prediction based Routing (FPBR), which is based on prediction techniques (predictive counters) to model lightpaths availability. Contrary to the prediction routing scheme proposed in this thesis called HPBR, FPBR considers the WCC for routing purposes.

The main contribution of this section is twofold. First, to introduce a new algorithm, namely FPBR that, similar to HPBR, makes use of fine-granularity predictive counters, conversely to other Predictive RWA algorithms which use predictive-counters per route/wavelength –coarse-granularity predictive counters; second, to introduce an insightful discussion on how the topology connectivity, the offered load, and the update time, all affect the blocking probability of a RWA algorithm when the WCC is considered. The study presented in this section may help to understand the performance of distinct source RWA algorithms in different dynamic large scenarios. For the sake of understanding Table 2.10 lists the set of symbols and terminology used in this section.

Table 2.10: List of Symbols and Terminology used for Section 2.4.4.

Symbols and Terminology	Meaning
$\lceil\rceil$	The ceiling operator, given a value it returns the smallest integer greater than or equal to the given value.
λ_n	An optical wavelength where $n \in \{1,...\|W\|\}$.
W	The set of wavelengths used by any WR, where $W \in \{\lambda_1, \lambda_2, ...\lambda_{\|W\|}\}$.
F	The set of fibers on any link, all links have the same number of fibers.
$M_{s,d}$	The set of candidates paths for a source-destination pair, where $s, d \in V$ and $s \neq d$.
$C^s_{i,\lambda}$	Fine-granularity predictive counter locally computed by WR s, where $i \in E$, $\lambda \in W$, and $C^s_{i,\lambda} \in \{0,1,2,3\}$.
C^s	Total of fine-granularity predictive counters per WR.
$l^s_{j,\lambda}$	Availability of a lightpath using route j and wavelength λ locally computed by WR s.
N_j	Is the length of route j in terms of hops.
W^s_i	Available wavelengths on link i locally computed by WR s.
$R^s_{i,\lambda}$	Residual capacity of wavelength λ on link i locally computed by WR s, where $R^s_{i,\lambda} \in \{1,...\|F\|\}$.
t_h	Average holding time for all CRs.
t_u	Periodic time interval defining the dissemination of NSI.
$P_{j,\lambda}$	Blocking probability of a lightpath using route j and wavelength λ.
ρ_i	Utilization of a wavelength on link i.

FPBR is proposed as a predictive source RWA algorithm that selects lightpaths based on historic information related to lightpaths availability. To this end, FPBR uses two-bit predictive counters $C^s_{i,\lambda}$ with fine-granularity, whose values from 0 to 1 and values from 2 to 3 model a lightpath availability and unavailability respectively, on link i using wavelength λ. Notice that predictive counters values are locally computed by each WR. Therefore, in case that a lightpath selected by WR s cannot be setup on link i using wavelength λ, the predictive counter $C^s_{i,\lambda}$ is increased (solely when the counter value is less than 3), otherwise it is decreased solely when counter value is greater than 0). As mentioned in this thesis, the use two-bit counters because as is enough to predict routes availability since higher or lower counter values add a high or low hysteresis level both leading to the improper modeling of lightpaths availability.

The availability of a lightpath ($l^s_{j,\lambda}$) using wavelength λ and route j is locally computed by a WR as shown in Equation (2.29). As a result, a WR using FPBR can select lightpaths without requiring global NSI.

$$l^s_{j,\lambda} = \left\lceil \frac{\sum_{i \in j} C^s_{i,\lambda}}{N_j - 1} \right\rceil \tag{2.29}$$

To illustrate the rationale driving the fine-granularity approach for predictive counters when considering the WCC, (conversely to other Predictive RWA algorithms which adopt a coarse-granularity approach), let's observe the scenario depicted in Fig.2.26 . In this scenario, $W^{accurate}$ stands for real (accurate) wavelengths availability on an optical link, whereas WS stands for the wavelengths availability locally computed by a WR; hence, it is potentially inaccurate. Moreover, FF is a source RWA algorithm consisting in the use of First-Fit (FF) for wavelength assignment purposes and shortest-path algorithm for routing.

A CR arrives to WR S_1 demanding a lightpath between WR nodes S_1 and D. For this purpose, in case that a conventional source RWA algorithm such FF or either Predictive RWA algorithms using fine or coarse-granularity predictive counters are used, all three RWA algorithms will select lightpath $S_1, 1, 2, D$ using wavelength λ_1. This will cause the blocking of the selected lightpath, because the NSI computed by WR S_1 related to the available wavelengths (specifically for link $1-2$) is outdated (inaccurate), see Fig. 2.26a.

For subsequent CRs (connections arriving before the next update time) FF and predictive algorithms will work differently. FF will continue selecting lightpath $S_1, 1, 2, D$ using wavelength λ_1, as long as the NSI in WR 1 is kept outdated, see Fig. 2.26b. Conversely, Predictive RWA algorithms can capture the unavailability of lightpath $S_1, 1, 2, D$ using wavelength λ_1, thus selecting wavelength λ_2 for the same path as shown in Fig. 2.26b. It is also important to show the different behavior shown by fine and course granularity predictive counters. Fig. 2.26c shows how coarse-granularity predictive counters cannot model the unavailability of wavelength λ_1 on the path $S_1, 1, 2$.

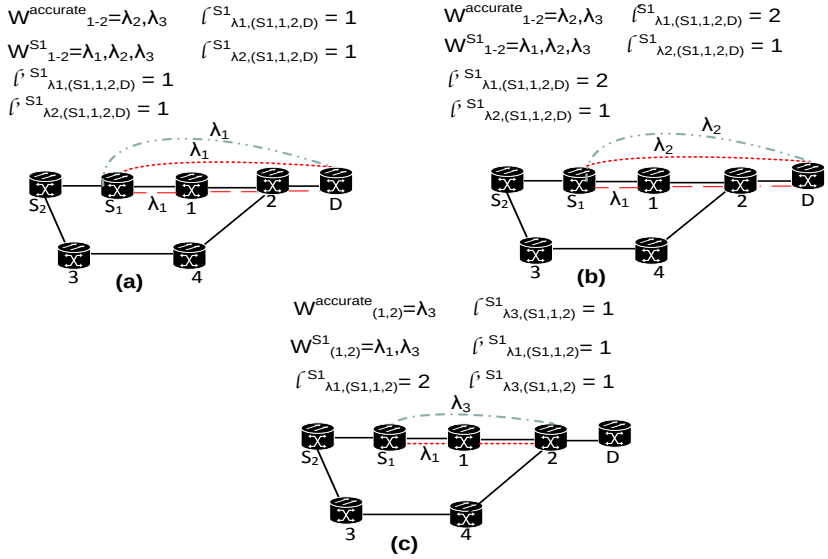


Figure 2.26: An illustrative example of the RI problem in WRNs: a) Handicaps of conventional RWA algorithms; b) Advantages of Predictive source RWA algorithms; c) Advantages of FPBR.

This is a handicap for coarse-granularity predictive counters that can be overcome by using fine-granularity predictive counters such as FPBR. Indeed, FPBR will select lightpath $S_1, 1, 2$ using wavelength λ_3, avoiding the blocking of future CRs.

The overall procedure related to the lightpath selection of FPBR for single-fiber networks is shown in Algorithm 3. The first step (step. 1) of the FPBR algorithm consists in randomly selecting –assuming uniform probability distribution for all wavelengths– a wavelength for each path to a given destination d. In this thesis it is considered that under inaccurate NSI a random selection of wavelengths is more suitable in order to offset the negative impact of inaccurate NSI. Then, by means of fine-granularity predictive counters, FPBR computes the availability of the selected lightpath, cf. Equation (2.29). If a lightpath's availability is lower than 2 and the selected wavelength is available on the output link towards the destination, the lightpath is then selected; otherwise, FPBR will continue to evaluate the availability of the remaining optical wavelengths. FPBR assumes that the NSI on the output link for all source WRs is 100% accurate. It is worth mentioning that in the case that all lightpaths are predicted to be unavailable, a source WR selects a lightpath based solely on its output links wavelengths availability, cf. step. 2 in Algorithm 3. This is also useful to unblock the lightpaths that are in an unavailable state ($l^s_{j,\lambda}$). Finally, if there are no lightpaths available the incoming connection is blocked.

For multi-fiber scenarios, the operation of FPBR is similar to the single-fiber scenario, but for the former scenario FPBR selects the optical wavelength with the highest residual capacity, i.e., the optical wavelength less used along an output link. In case that two or more optical wavelengths have the same residual capacity, then an optical wavelength is randomly selected. Algorithm 4 depicts this wavelength assignment strategy.

Moreover, the computational complexity of both Algorithm 3 and Algorithm 4 is discussed in the following lines.

FPBR algorithm works in two phases: an offline route generation phase, and an online lightpath selection phase (shown in Algorithm 3 and Algorithm 4). In the route generation phase, $|M_{s,d}|$ pre-computed (candidate) routes are (offline) generated for each source-destination pair by means of Dijkstra algorithm with a complexity of: $O\big(|M_{s,d}| \times |E| + |V| \times log\,(|V|)\big)$.

For a worst case scenario, the online phase has a complexity of (assuming a single-fiber network): $O\big(2 \times |M_{s,d}| \times |W|\big)$, whereas for a multi-fiber network the complexity is:

$O\big(2 \times |M_{s,d}| \times |W|| \times |F|\big)$.

The rationale driving the adoption of a route pre-computation strategy is to minimize the complexity of the lightpath selection phase and to ease the deployment of FPBR in comparison with an adaptive routing strategy. Despite the fact, that a pre-computation strategy can impose a storage overhead per source WR of $|M_{s,d}| \times |V| - 1$, assuming all WRs as possible destinations. This storage overhead can be neglected for a low amount of candidate paths. Indeed, a

Algorithm 3 Lightpath selection of FPBR for single fiber networks.

Input: (d)
Output: $(route, \lambda)$

{step. 1}
$W' = W$ {create a copy of wavelengths set}
$state$=0
$route, \lambda = \emptyset$
for j in $M_{s,d}$ **do**
  λ =select(W') {Randomly select a wavelength.}
  $W'.remove(\lambda)$ {Remove selected wavelength from the wavelengths set.}
  if $l^s_{j,\lambda} < 2$ *and* λ is available on the output link of s to route j **then**
    $route$=j
    Provision($route, \lambda$)
    $state$=1
    BREAK{end loop execution}
if $state == 0$ **then**
  $W' = W$
  {step. 2. In case that all lightpaths are considered unavailable}
  for j in $M_{s,d}$ **do**
    λ =select(W') {Randomly select a wavelength.}
    $W'.remove(\lambda)$ {Remove selected wavelength from the wavelengths set.}
    if λ is available on the output link of s to route j **then**
      $route$=j
      $state$=1
      Provision($route, \lambda$)
      BREAK{end loop execution}
for i in $route$ **do**
  if $state == 1$ *and* $C^s_{i,\lambda} > 0$ *and* $i \neq$output link **then**
    decrease $C^s_{i,\lambda}$ {1 means that selected lightpath could not be provisioned.}
  else if $state == 0$ *and* $C^s_{i,\lambda} < 3$ **then**
    increase $C^s_{i,\lambda}$

high amount of candidate paths is an overkill since they do not significantly improve the performance of a route pre-computation strategy [23].

On the other hand, as described in this section, FPBR does not require global knowledge of NSI. This is a feature highly appreciated for the scalability of source RWA algorithms using distributed control planes, mainly motivated by the difficulty to maintain accurate NSI with the ever increasing constraints such as PLI factors, energy consumption, among others.

Finally, the total amount of predictive counters per WR is shown in Equation (2.30).

$$C^s = |E| \times |W| \tag{2.30}$$

However, for a worst-case scenario, i.e., a full-mesh network, the total amount of predictive counters is computed as $C^s = (|V| \times (|V|-1) \times W|)|$. Therefore, in large full-mesh networks C^s

Algorithm 4 Lightpath selection of FPBR for multi-fiber networks.

Input: (d)
Output: $(route, \lambda)$

{step. 1}
$state$=0
$route, \lambda$=$\emptyset$
for j in $M_{s,d}$ **do**
  W' =Select the wavelengths with the highest residual capacity on the first output link of s to route j
  λ =select(W'){Randomly select a wavelength.}
  $W'.remove(\lambda)$ {Remove selected wavelength from the wavelengths set.}
  if $l^s_{j,\lambda} < 2$ **then**
    $route$=j
    $state$=Provision($route, \lambda$)
    $state$=1
{The rest of the algorithm is similar to Algorithm. 3}

is significantly greater than the total amount of coarse-granularity predictive counters per WR which is $\sum_{d \in V} |M_{s,d}| \times W$. However, C^s can be significant reduced if it is assumed that each WR solely maintains state of predictive counters state related to the optical links part of its set of candidate paths. In this case the computation of C^s is shown in Equation (2.31).

$$C^s = \sum_{d \in V} \left| \bigcup_{k \in M_{s,d}} k \times |W| \right| \tag{2.31}$$

2.4.5 Simulation Results with regard to Finer Prediction based Routing

This section presents extensive simulation results related to the blocking probability of distinct source RWA algorithms evaluated on the well-known NSFNET topology (14 nodes, 21 links) with an Average Shortest Path Length (ASPL) of 2 hops, Spanish Backbone topology, ASPL=3.30 hops, and the DCN topology, ASPL=2.14 hops, see Fig. 2.27.

The simulation results presented in this section were obtained using the well-known network simulation framework Omnetpp [85], and all plotted values have a 95% confidence interval not larger than 0.5% of the plotted values.

The network models used namely Single-Fiber and Multi-Fiber to obtain the simulation results are described in the following lines. For these two network models the following settings apply.

- CRs arrive at a node according to a Poisson process with an inter-arrival mean time t in an incremental manner: $CR(t_1 = t)$, $CR(t_1 + t)$,....$CR(t_{n-1} + t)$.
- The connection holding time (t_h) is exponentially distributed with a mean of 50 time

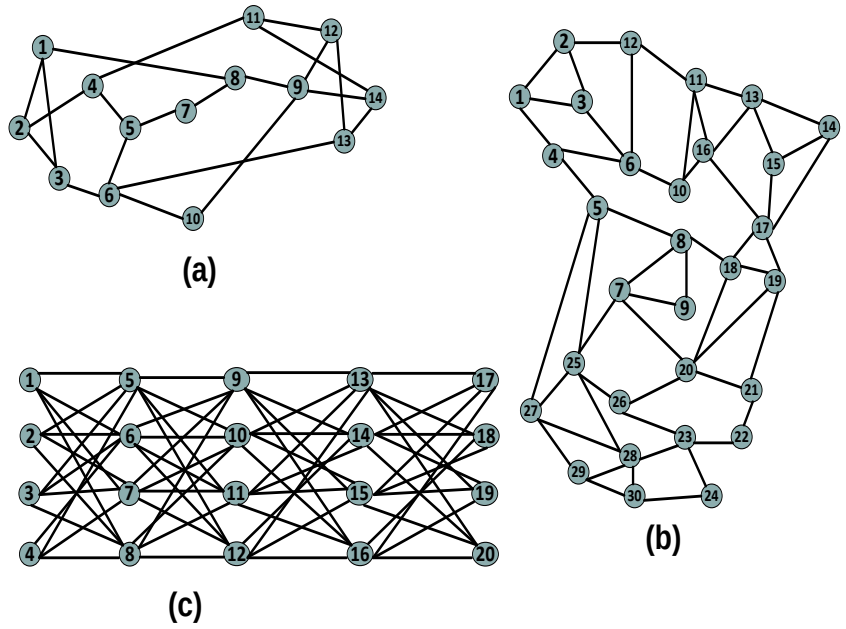


Figure 2.27: Evaluated network topologies: a) NSFNET topology (14 nodes, 21 links); b) Spanish Backbone Topology; c) DCN topology.

units.

- WRs without wavelength conversion capabilities –a WS optical network.
- Each CR requires a full wavelength on each link –grooming is not assumed. This assumption is motivated by the high bandwidth demands within DCNs scenarios –above 100 Gbps.
- NSI is periodically disseminated according to the update time parameter (t_u), after this dissemination it is instantaneously available at all WRs.
- Blocked CRs are not reattempted.
- Once a lightpath is provisioned it cannot be reconfigured during its lifetime.
- Control messages loss as well as both propagation and connection setup delays are neglected. This assumption enables to solely focus on the study of RI affected by t_u intervals, which impact on the blocking probability is dominating.
- For every source-destination pair each WR has 4 candidate paths. The candidate paths were computed off-line by means of Dijkstra's algorithm with the number of hops as the routing metric and are sorted according to their cost, i.e., the first and last candidate path are the shortest and longest path respectively. The candidate paths are recomputed whenever the network topology changes.

- 5 WRs as sources and 6 as destinations.
- 8 wavelengths per fiber.

Moreover, for the single-fiber model it is assumed the following settings: 1) An offered load of 5 erlangs per node; and 2) a single fiber per link; whereas for the multi-fiber model it is assumed the following settings: 1) An offered load of 10 erlangs per node; and 2) 2 fibers per link.

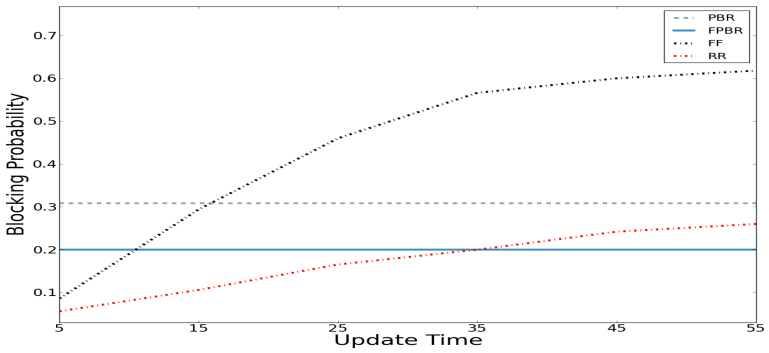
Figures 2.28a, 2.28b and 2.28c depict the blocking probability versus a diverse spectrum of t_u values in a single-fiber network for FF, RR (consisting on a random strategy for wavelength assignment with shortest-path routing), a Predictive RWA algorithm using coarse-granularity counters hereinafter referred to as PBR , which is proposed by authors in [19], and the FPBR RWA algorithm. As shown in Fig. 2.28, both FF and RR algorithms yield the lowest blocking probability only when the update time is very low. However, when the update time increases to higher (realistic) values, the blocking probability of both FF and RR increases significantly and tends to settle when the update time is approximately equal to the holding time. Contrary to FF and RR, Predictive RWA algorithms such as PBR and FPBR are not affected by the update time. Remind that Predictive RWA algorithms do not require the global NSI.

It is worth mentioning that the Average Shortest Path Length (ASPL) is a collateral source of inaccuracy with regard to NSI. According to [86],$P_{j,\lambda} = 1 - \left(\prod_{i \in j}\left(1 - \rho_i^F\right)\right)^{|W_i|}$; Hence, a longer path length implies a higher probability of no meeting the WCC constraint. This is aggravated under inaccurate NSI where the blocking probability is highly sensitive to the update time, as proved by [36]. For convenience let $P_{j,\lambda}^{t_0}$ be the blocking probability of a lightpath under a update time of 0 time units (100 % accurate NSI), hence it can be assumed that $P_{j,\lambda}^{t_0} < P_{j,\lambda}^{t_1} < ...P_{j,\lambda}^{t_h}$. This is assumption is validated by the obtained simulation results. Notice that for single-fiber networks the blocking probability of RWA algorithms dependent of NSI is lower and tends to settle faster with topologies with low ASPL such as NSFNET and DCN; the contrary occurs with topologies with high ASPL such as the Spanish Backbone topology.

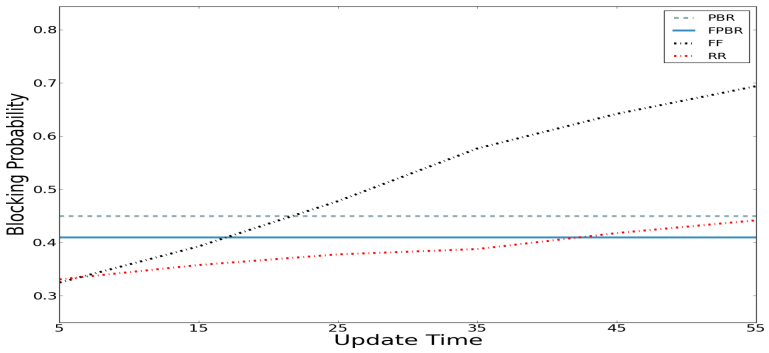
A secondary cause of inaccurate NSI is the dynamicity of CRAs. High frequency of CRAs leads to highly inaccurate NSI, the contrary occurs with low arrival rates. In light of this, in Fig. 2.29, it is shown the blocking probability versus a large spectrum of arrival rates with an update time of 25 time units –the rest of the network parameters assumed for single-fiber networks are the same. As it can be seen Predictive RWA algorithms are not affected by the connection arrival rate, conversely to FF and RR.

Based on the obtained results, it can be stated that the performance of conventional RWA algorithms such as FF and RR are strongly affected by the inaccuracy of NSI. This is not the case for the proposed scheme, which has a low dependency of NSI dissemination. In conclusion, FPBR presents the best performance of the evaluated schemes in single-fiber scenarios, whereas FF (optimal under accurate NSI) yields the worst performance.

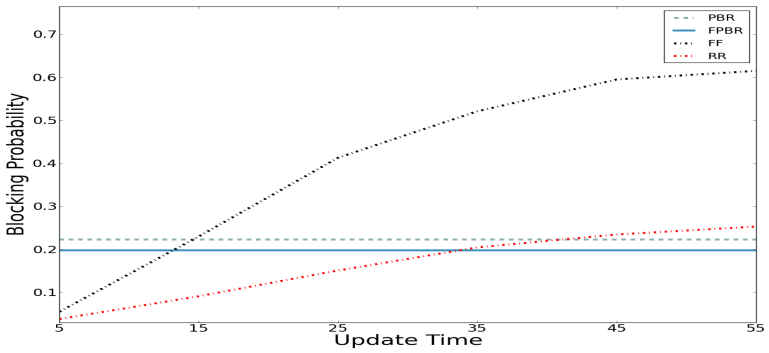
On the other hand, the proposed RWA algorithm was evaluated in multi-fiber scenarios and



(a)

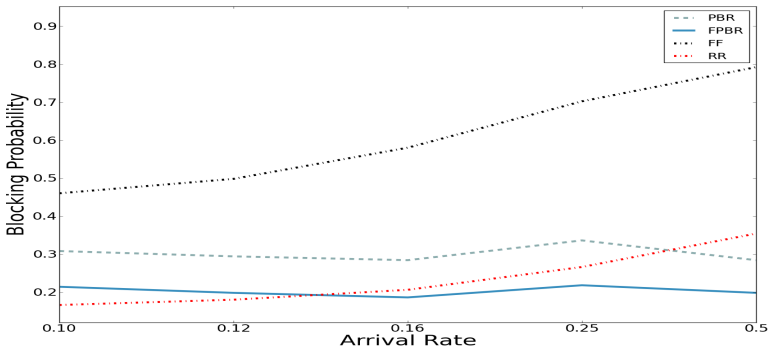


(b)

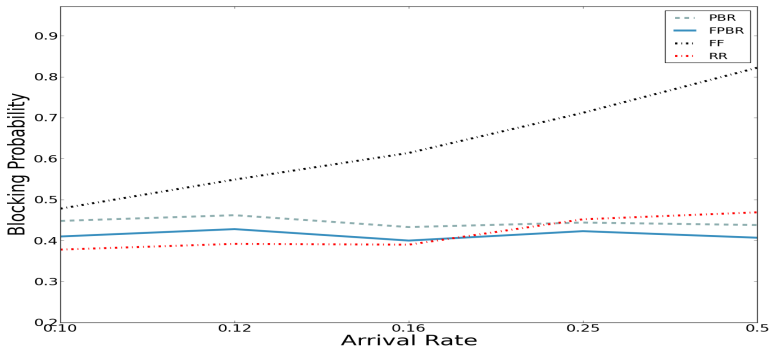


(c)

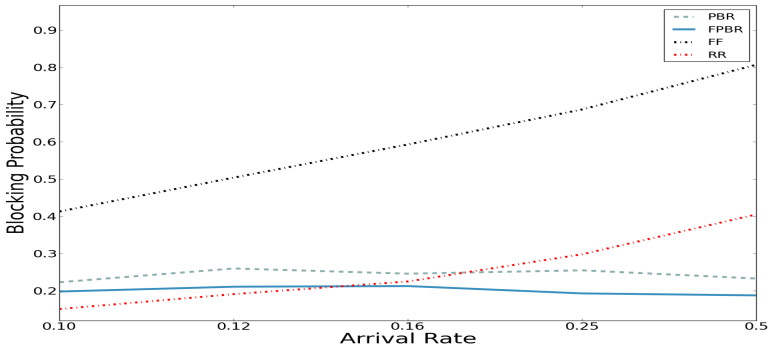
Figure 2.28: Blocking probability versus a large spectrum of update times considering a single-fiber model for: a) NSFNET; b) Spanish Backbone Topology; c) DCN topology.



(a)



(b)



(c)

Figure 2.29: Blocking probability versus a large spectrum of arrival rates considering a single-fiber model for : a) NSFNET topology; b) Spanish Backbone Topology; c) DCN topology.

compared with multi-fiber RWA algorithms such as Least Loaded (LL), PBR, and the Weighted Selective Adaptive Routing (WSAR) [87].

The LL algorithm is explicitly designed for multi-fiber WRNs, this algorithm selects the wavelength that has the largest residual capacity (the wavelength least used among the fibers on a link) on the most congested link. The LL operation is as follows: 1) $i' = min_{i \in j}(|W_i^s|)$; then, 2) selected_wavelength= max$\left(R_{i',\lambda}^s\right)$, where i' is the most congested link on a path j. It is worth mentioning that the LL algorithm needs global NSI, hence is susceptible to inaccurate NSI.

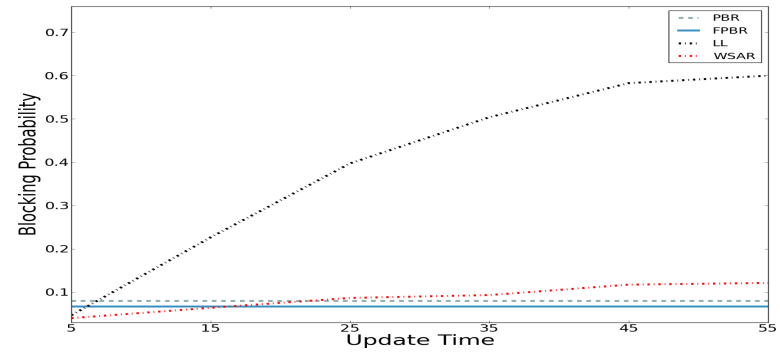
The WSAR algorithm is another algorithm devised for multi-fiber WRNs which incorporates an innovative weight function method that enables to choose lightpaths based on the wavelength availability and multi-fiber segments. Once the shortest path is selected, WSAR chooses a wavelength in a random manner.

Based on the results shown in Fig. 2.30, it can be concluded that indeed, multi-fiber systems can reduce the blocking probability of any RWA algorithm. Moreover, similar to single fiber scenarios, global RWA algorithms such as LL and WSAR outperform the evaluated Predictive RWA algorithms solely when the update time is low; but when the update time increases the opposite occurs. Unfortunately, though expected, the performance of fine-granularity predictive counters tends to approximate to coarse-granularity predictive counters performance. This is because the inaccuracy of NSI is decreased due to the use of multi-fiber systems. Therefore, the advantages of fine-granularity predictive counters are neglected under accurate NSI. Moreover, notice that since the adoption of multi-fiber systems decrease the inaccuracy related to NSI, it is not surprising that the blocking probability of global RWA algorithms increases significantly in topologies with low ASPL such as the DCN topology.

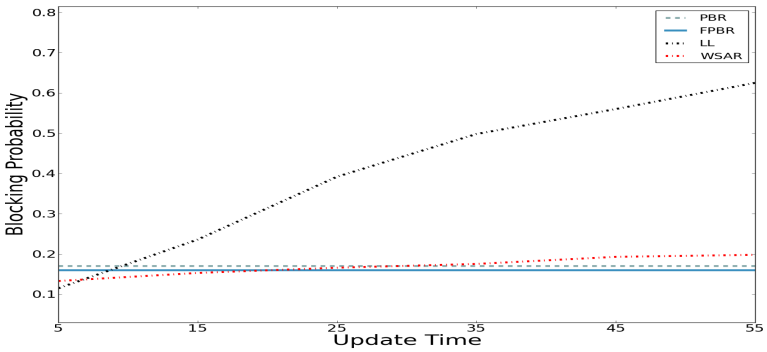
Finally, Fig. 2.31 shows the blocking probability for a large spectrum of arrival rates with an update time= 25 time units –the rest of the network parameters assumed for multi-fiber networks are the same. Similar to the results shown in Fig. 30, Predictive RWA algorithms are not affected by the connection arrival rate in multi-fiber scenarios, conversely to LL and WSAR, even thought the latest is less affected by the inaccuracy caused by high arrival rates.

Based on the obtained simulation results presented in this section, the following lessons are learned related to the study of RWA in dynamic scenarios under inaccurate NSI.

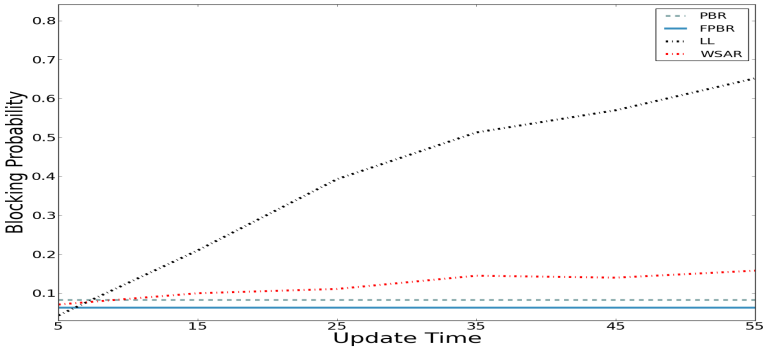
- The update time, connection arrival rates, as well as topology characteristics such as the ASPL impact the blocking probability of RWA algorithms under inaccurate NSI.
- Predictive RWA algorithms can outperform RWA algorithms relying on global NSI under the assumption of realistic (greater than 5 time units) update times.
- Fine-granularity predictive counters are more suitable for predicting lightpaths availability in comparison with a coarse-granularity approach in single-fiber networks. However, in multi-fiber networks the performance of both types of predictive counters is similar.



(a)



(b)



(c)

Figure 2.30: Blocking probability versus a large spectrum of update time values considering a multi-fiber model for: a) NSFNET topology; b) Spanish Backbone Topology; c) DCN topology

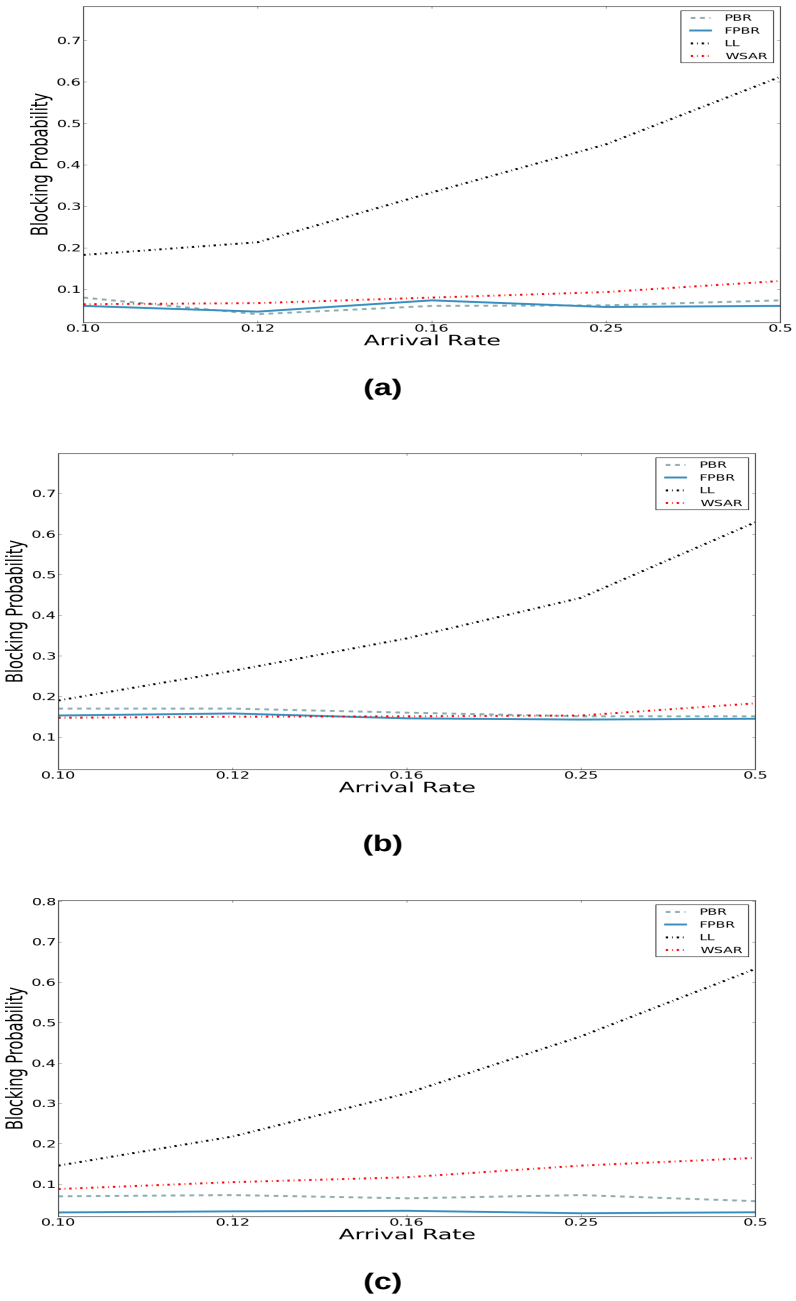


Figure 2.31: Blocking probability versus a large spectrum of arrival rates considering a multi-fiber model for: a) NSFNET topology; b) Spanish Backbone Topology; c) DCN topology.

2.4.6 Routing Inaccuracy Problem in Dynamic Protected Scenarios

In network research, the negative effects of inaccurate NSI on the routing performance have been widely evaluated and analyzed in unprotected network scenarios. However, there is limited information of the RI problem in protected scenarios. It is intuitive observing that the negative effects of inaccurate NSI are more harmful in protected scenarios where two paths must be computed per CR, i.e., a primary and a link-disjoint backup lightpath in order to provide resilient services. This section focuses on the study of RI in protected scenarios considering the WCC. To this end, proactive protection schemes such as NCP and Dedicated Path Protection (DPP) are evaluated.

NCP has been widely studied in network research for offline scenarios, where CR demands are known beforehand. Unfortunately, there is limited information related to the study of NCP in online scenarios under inaccurate NSI. The rationale driving this section is to fill this gap, by extensively studying the behavior of NCP in a source RWA scenario under dynamic traffic and inaccurate NSI caused by a periodic updating policy. This section shall show that NCP yields a lower P_{cost} because of its efficient usage of network resources in comparison with conventional proactive protection schemes, such as DPP. However, NCP might be less resilient to inaccurate NSI due to the routing constraints that must be met in order to obtain the benefits of NC.

In order to mitigate the negative effects of inaccurate NSI on the performance of an NCP scheme, we propose an innovative NCP scheme referred to as Predictive Network Coding Protection (PNCP). PNCP leverages predictive techniques in order to compute both primary and backup lightpaths meeting the follow requirements: 1) low blocking probability, 2) low signaling overhead, 3) efficient utilization of network resources allocated for path protection (hereinafter referred to the P_{cost}), and 4) low recovery time. The main characteristic of PNCP is that it can successfully mitigate the negative effects of inaccurate NSI by means of predictive techniques and that it optimizes the P_{cost}.

Extensive simulation results are presented assessing that the proposed scheme significantly outperforms conventional proactive protection schemes such as a DPP with regard to both blocking probability and the P_{cost}.

Operation of NCP in Dynamic Scenarios

This section first introduces in a comprehensive manner the advantages of NCP with regard to the P_{cost}. Then, it is distilled the negative effects that inaccurate NSI might have on the performance of an NCP scheme. For the sake of understanding Table 2.11 lists the set of symbols and terminology used through this section.

Table 2.11: List of Symbols and Terminology used for Section 2.4.6.

Symbols and Terminology	Meaning
$G(V, E)$	Directed graph where E is the set of optical links and V is set of WRs.
CR	Connection Request.
P_{cost}	Amount of network resources allocated for path protection.
W	The set of optical wavelengths available for any WR.
λ_n	An optical wavelength, where $n \in \{0,..\|W\|-1\}$.
W_i^s	The set of optical wavelengths along a link i locally computed by a WR s, where $i \in E$ and $s \in V$.
Pr	Blocking probability.
$P_i^s{}_{,\lambda}$	Predictive counter for link i and wavelength λ locally computed by a WR s, where $\lambda \in W$, and $P_i^s{}_{,\lambda} \in \{0,1,2,3\}$.
$A_{j,\lambda}^s$	Availability of a lightpath using route j and wavelength λ locally computed by WR s.
N_j	Is the length of route j in terms of hops.
$M_s{}_{,d}$	The set of candidate paths for endpoints s, d, where $d \in V$.
ρ_i	Utilization of a wavelength on link i.
U	The allocation of one optical wavelength on an optical link.
Coding lightpath	A lightpath that conveys coded (protected) traffic.

To clearly illustrate the operation of proactive protection schemes under dynamic traffic we consider the scenario shown in Fig 2.32. In this scenario, a CR (CR_1) arrives to WR S_1

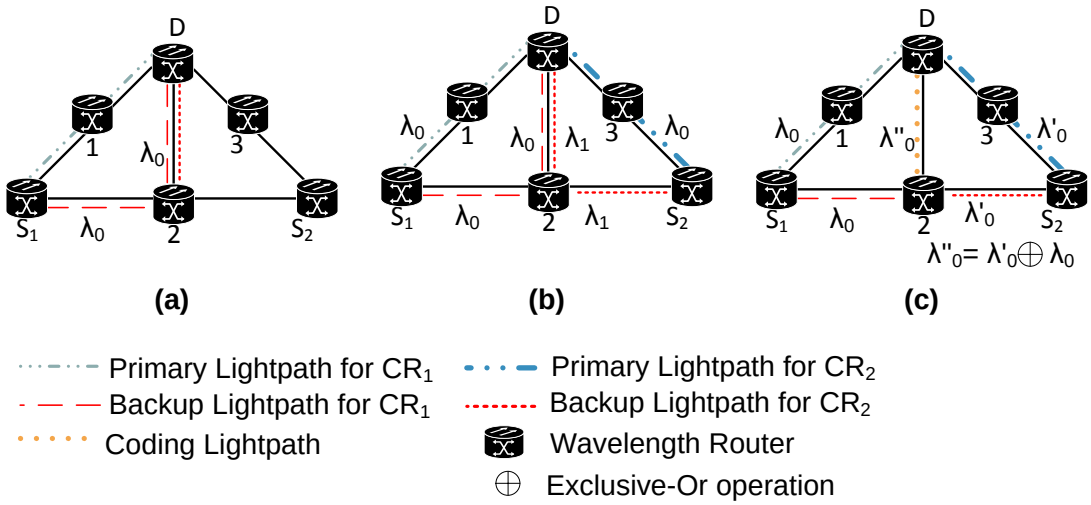


Figure 2.32: Operation of proactive protection schemes: a) and b) Protection with a DPP scheme c) Protection with a DPPNC scheme.

requesting a resilient lightpath for endpoints $S_1 - D$ with a holding time of 50 time units. For this purpose, a proactive DPP scheme based on shortest-path routing jointly with FF for wavelength assignment purposes, hereinafter referred to as LCP-FF (in the rest of this section the terms DPP and LCP-FF are used interchangeably) assuming source routing, will compute two link-disjoint paths: 1) a primary lightpath consisting on the path $S_1 - 1 - D$ using wavelength λ_0, and; 2) a backup lightpath consisting on a path $S_1 - 2 - D$ using wavelength λ_0, see Fig. 2.32a.

Afterwards, a subsequent CR (CR_2) arrives to WR S_2 requesting a resilient lightpath between endpoints $S_2 - D$ with a holding time of 70 time units. As a result, a DPP scheme computes paths $S_2 - 3 - D$ using λ_0 and path $S_2 - 2 - D$ using λ_1 for the primary and backup lightpaths respectively, see Fig. 2.32b. The total P_{cost} allocated for the backup lightpaths of both CR_1 and CR_2 using DPP is $4U$.

Nevertheless, the P_{cost} can be reduced if a DPPNC* scheme is used for the scenario depicted in Fig.2.32. A DPPNC scheme will select lightpath $S_1 - 2$ using λ_0 for CR_1, lightpath $S_2 - 2$ using λ_0' for the backup path of CR_2, and along optical link $2 - D$ on optical wavelength λ_0'' will be coded (perform the all-optical XOR operation) the backup traffic received along optical links $S_1 - 2$ and $S_2 - 2$ corresponding to CR_1 and CR_2 respectively. Under this setting, the total P_{cost} would be $3\,U$. This is because by doing NC is possible to convey information of more than one data stream allocating the same amount network resources to do so. Indeed, the advantage of an NCP scheme relies on the coding of traffic. For the sake of convenience we use the notation λ, λ' and λ'' to differentiate traffic sent along different optical links allocated on the same optical wavelength.

As a result, in case of a failure affecting either the primary lightpath of CR_1 or CR_2, WR D

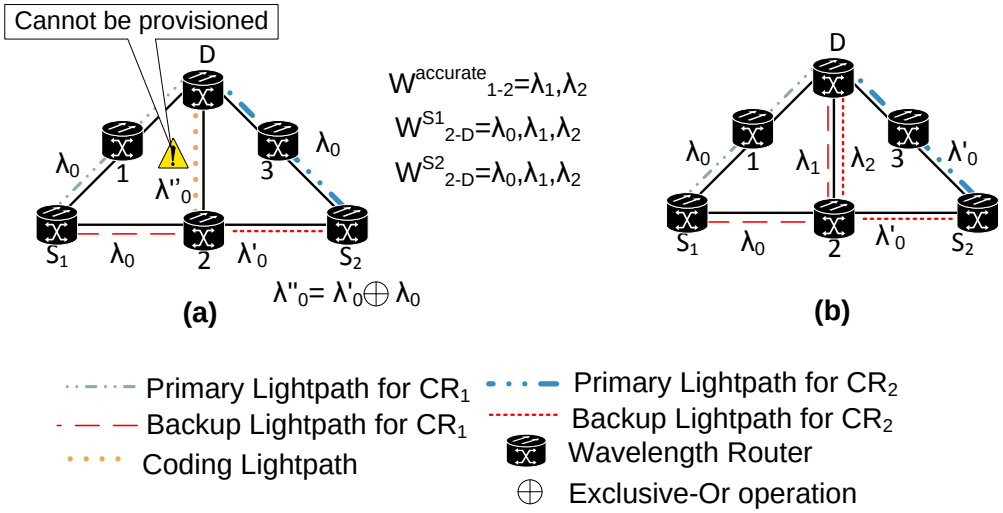


Figure 2.33: Operation of proactive protection schemes under inaccurate NSI: a) protection using DPPNC scheme; b) protection using a DPP scheme.

can recover the affected traffic by doing the all-optical XOR operation of λ_0'' and the traffic sent along the unaffected primary lightpaths. Notice that the lightpath $2-D$ using λ_0'' is referred to as a *coding lightpath*. A coding lightpath is a lightpath that conveys coded traffic.

It is worth mentioning that the holding time corresponding to a coding lightpath must be equalized (extended) to the holding time corresponding to the CR with the longest holding time allocated on this coding lightpath. For instance, once the backup traffic of CR_2 is coded along lightpath $2-D$, the holding time of this lightpath —-set previously to the holding time of CR_1– must be equalized to the holding time of CR_2, since this connection will remain longer on the network. Otherwise, once the holding time of CR_1 expires, the coded lightpath $2-D$ will be torn down; hence; impacting on the traffic sent along the backup lightpath assigned to CR_2. For more information related to holding time connection awareness the reader is referred to as [88].

The coding strategy used in Fig.2.32c, hereinafter referred to as *Preference Coding* consists in considering the following rule. When NC may be enabled along a selected backup route the selected optical wavelength is changed to the optical wavelength that enables NC, e.g., optical wavelength λ_1 was changed to optical wavelength λ_0' in Fig. 2.32c.

The scenario shown in Fig. 2.32 shows the operation of an NCP scheme under accurate NSI. However, there are several issues that must be considered under inaccurate NSI. In light of this, consider the scenario shown in Fig. 2.33. In this scenario, both WR S_1 and WR S_2 have inaccurate (outdated) NSI regarding the wavelengths availability on link $2-D$. This issue might lead to the blocking of the backup lightpath for both CR_1 and CR_2, since wavelength λ_0'' might not be really available along link $2-D$ ($W_{2-D}^{accurate} = \lambda_2, \lambda_3$), despite that appears as available in the routing tables of both WRs, see Fig. 2.33a.

Authors in [89] provide a simple but effective analytical model to evaluate the blocking probability of a lightpath for single-fiber networks as shown in Equation (2.32). This equation is valid for DPP schemes. However when NCP is enabled, the blocking probability of a coding lightpath is computed as shown in Equation (2.33). Notice that a backup lightpath suitable for NC is shorter, i.e., spans fewer optical links, since some of the network resources used for a backup lightpath are already reserved for a different backup lightpath, e.g., lightpath $2-D$ using λ_0'' was previously reserved, see Fig. 2.33a, hence the lightpath to be provisioned –hereinafter referred to as the uncoded lightpath– is path S_2-2 using λ_0. Recall that the advantage of NCP relies on sharing backup network resources. However, under this constraint the wavelength allocated along the coding lightpath (λ_0) must be available along the uncoded lightpath in order to enable NC. Therefore, in the absence of wavelength conversion, the use of NCP (DPPNC) over no NCP (DPP) might be counterproductive in the presence of inaccurate NSI. This is validated by extensive simulation results shown in the next section.

$$Pr = \left(1 - \prod_{i \in j} \left(1 - \rho_i\right)\right)^{|W_i|} \tag{2.32}$$

$$Pr = 1 - \prod_{i \in j'} \left(1 - \rho_i\right) \tag{2.33}$$

In order to reduce the blocking probability of the DPPNC scheme while still exploiting the advantages of NCP, a possible strategy to follow (hereinafter referred to as *Opportunistic Coding strategy*) is to enable NC as long as it does not lead to blocking, as shown in Fig. 2.33b. To this end, the wavelength selected by a DPPNC scheme will be changed for a wavelength that enables NC as long as this wavelength is available along the uncoded path. Nevertheless, the advantages of an Opportunistic Coding strategy can be reduced by the negative effects driven by having inaccurate NSI.

Prediction based Network Coding Protection

By carefully observing the negative effects of inaccurate NSI on protected scenarios, an intuitive reasoning is to devise a protection scheme that does not require global NSI. This protection scheme should be able to compute lightpaths with low blocking probability by solely using local NSI. This kind of protection scheme is referred to as predictive RWA algorithms. Predictive RWA algorithms are not affected by the inaccuracy added by updating policies; hence, avoiding periodic dissemination of NSI. As a matter of fact, this is a scalability advantage looking forward to network scenarios such as DCNs where large signaling overhead substantially impacts the performance of protection schemes.

Motivated by the good performance of Predictive RWA algorithms under the presence of inaccurate NSI, in this thesis it is devised a protection scheme so-called Predictive Network Coding Protection (PNCP). PNCP is a proactive protection scheme that combines the advantages related to network throughput improvement of NCP and the benefits of Predictive RWA algorithms related to the blocking probability in order to mitigate the negative effects of inaccurate NSI.

In this section, it is presented a thorough description with regard to the PNCP mechanism, supported by two aligned conceptual arguments. Firstly, it is justified a potential utilization of predictive protection RWA algorithms instead of conventional protection schemes by highlighting its strengths and benefits. Then, it is provided a scalability analysis of the proposed NCP scheme that clearly validates PNCP utilization in large scale scenarios.

PNCP is proposed as a predictive source RWA algorithm that computes resilient lightpaths by means of both predictive counters and the wavelength availability of output links –it is assumed that a WR have accurate NSI of its output links. To this end, PNCP extends the predictive concepts used by [19],[77], and used in protected scenarios. However, PNCP adopts a fine-granularity approach for predictive counters (predictive counters per link-wavelength), contrary to authors in [19],[77], which used a coarse-granularity approach (predictive counters per lightpath).

A predictive counter ($P^s_{i,\lambda}$) measures the availability of a lightpath along a link. Specifically, PNCP adopts two-bit predictive counters. Values from 0 up to 1 predict that a lightpath is available along link i using wavelength λ, whereas values from 2 up to 3 predict the contrary. The reasoning driving us to adopt two-bit predictive counters is to control the degree of hysteresis of predictive counters. It is proven by authors in [19] that two-bit counters are sufficient for predicting lightpaths availability. It is worth noting that predictive counters are locally computed by each source WR. The availability of a lightpath using route j and optical wavelength λ ($A^s_{j,\lambda}$) is computed as it shown in Equation (2.34). Notice that the predictive counter values are squared in order to minimize the selection of lightpaths with predictive counters greater than 2.

$$A^s_{j,\lambda} = \frac{\sum_{i \epsilon j} \left(P^s_{i,\lambda}\right)^2}{N_j - 1} \tag{2.34}$$

In Algorithm 5 it is presented the overall procedure for PNCP. As it can be observed, PNCP selects a primary lightpath based on is availability computed by the predictive counters and the available bandwidth on its output link to a route j.

On one hand, PNCP selects optical wavelengths in a random manner, where the probability of selection of each wavelength is uniformly distributed. On the other hand, for the routing purposes, each source WR has at least 2 link-disjoint candidate paths. These candidate paths

Algorithm 5 Overall Procedure of PNCP.

Input: (d)
Output: ($ResilientLightPath$)

$primaryLightPath, backupLightPath$=∅{Initialize primary and backup lightpaths}
{A lightpath is a tuple formed by a route (lightpath[0]) and an optical wavelength (lightpath[1])}
$primaryLightPath$ =ResilientLightPath (d,∅,0)
if $primaryLightPath$! = ∅ **then**
  $backup, \lambda'$ =ResilientLightPath (d,$primaryLightPath$[0],1)
$ResilientLightPath$=$\left(primaryLightPath, backupLightPath\right)$

FUNCTION ResilientLightPath (d,$primary$,opt)
{ opt Indicates if is a primary (0) or a backup lightpath (1) }
$W' = W${create a copy of wavelengths set}
$route, \lambda$=∅
$step$=0
for j in $M_{s,d}$ **do**
  if $primary \cap j$! = ∅ and $primary$! = ∅ **then**
    Continue
  if opt == 0 **then**
    λ =random_select(W'){Randomly select a wavelength fir primary lightpaths}
  else
    λ =FF_select(W'){select a wavelength in a First-Fit Fashion for backup lightpaths}
  $W'.remove(\lambda)${Remove selected wavelength from the wavelengths set.}
  if $A^s_{j,\lambda} < 2$ and λ is available on the output link of s to route j **then**
    $route$=j
    $state$ =Provision($route, \lambda$)
    $step$=1
    BREAK{end loop execution}
if $step$ == 0 **then**
  $W' = W$
  {$step$. 0. In case that all lightpaths are considered unavailable}
  for j in $M_{s,d}$ **do**
    if opt == 0 **then**
      λ =random_select(W'){Randomly select a wavelength fir primary lightpaths}
    else
      λ =FF_select(W'){select a wavelength in a First-Fit Fashion for backup lightpaths}
    $W'.remove(\lambda)${Remove selected wavelength from the wavelengths set.}
    if λ is available on the output link of s to route j **then**
      $route$=j
      $step$=1
      $state$=Provision($route, \lambda$)
      BREAK{end loop execution}
if opt! = 0 **then**
  λ =DoOpportunisticCoding($route, \lambda$)
for i in $route$ **do**
  if $state$ == 1 and $P^s_{i,\lambda} > 0$ and $i \neq$output link **then**
    decrease $P^s_{i,\lambda}$ {1 means that selected lightpath could not be provisioned.}
  else if $state$ == 0 and $P^s_{i,\lambda} < 3$ **then**
    increase $P^s_{i,\lambda}$
RETURN (route,λ)
END FUNCTION

are sorted from the shortest to the longest path taking into account the number of hops as a routing metric. In case that all lightpaths are predicted to be unavailable, then, PNCP selects a primary lightpaths solely based on the output links availability (which it is assumed to be 100% accurate). However, when a lightpath cannot be selected because all optical wavelengths are unavailable along the output links, the primary lightpath is blocked.

When a primary lightpath is successfully provisioned, PNCP proceeds to compute a protection lightpath that must be link-disjoint from the primary lightpath recently provisioned. For this purpose the operation of PNCP is similar as the computation of primary lightpaths, but instead of using a random strategy for wavelength selection, PNCP selects protection wavelengths in a First-Fit fashion, where wavelengths are sorted in a low frequency manner. This is done in order to efficiently pack the optical spectrum, hence avoiding disperse optical spectrum allocation. In this way, there are more chances to do NC in the absence of wavelength conversion capabilities.

Once a protection lightpath is selected, a source WR with PNCP will proceed to provision the selected lightpath. Nevertheless, in case that NC may be potentially enabled along the selected route and another wavelength must be selected, PNCP will proceed to do an Opportunistic Coding strategy. In this case, the previously selected wavelength might be changed for an optical wavelength suitable for NC solely based on local NSI, i.e., lightpaths availability and output links bandwidth. Otherwise, NC would not be enabled.

In order to illustrate the operation of PNCP we consider the scenario shown in Fig. 2.34. In this scenario a resilient lightpath must be provisioned between endpoints S and D_1. In the case that either a conventional RWA algorithm such as LCP-FF or PNCP are used, both will select path $S-D$ using λ_0 for the primary lightpath and path $S-2-D_1$ using λ_0 for the backup lightpath. Unfortunately, using any of the two schemes, the backup lightpath will not be provisioned because λ_0 is not available on link $2-D_1$, see Fig. 2.34a. This occurs because the NSI related to link $2-D_1$ locally computed by WR S is inaccurate.

Now suppose that a subsequent CR arrives (before the next update time) to WR S requesting a resilient lightpath between endpoints $S-D_2$. In order to compute this resilient lightpath, LCP-FF and PNCP will work differently. LCP-FF will continue selecting lightpath $S-2-D_1-D_2$ using λ_0 as a backup lightpath, this will lead to the blocking of this lightpath because WR S haven't updated its NSI related to optical link $2-D$. Conversely, PNCP will be able to capture the unavailability of wavelength λ_0 along link $2-D$, i.e., ${P^s}_{i,\lambda_0} = 2$. Thus, it will select λ_1 instead, see Fig. 2.34b.

On the other hand, the operation of the PNCP algorithm comprises two phases: 1) an offline route generation phase –assuming fixed-alternate path routing, and 2) an online lightpath selection phase (as shown in Algorithm 5). In the route generation phase, $|M_{s,d}|$ pre-computed (candidate) link-disjoint routes using a two-step approach are generated offline for each source-destination WR pair by means of Dijkstra's algorithm.

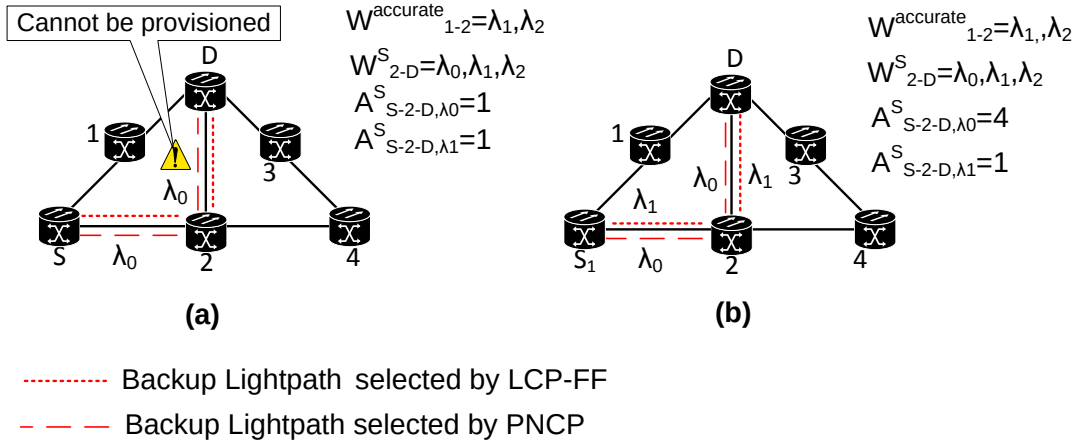


Figure 2.34: Operation of conventional and predictive proactive protection schemes under inaccurate NSI: a) protection using DPP; b) protection using a PNCP.

An offline route generation strategy is adopted due to both, minimizing the complexity of the lightpath selection phase and the scalability of fixed-alternate path routing strategies (in comparison with adaptive routing strategies).

2.4.7 Simulation Results with regard to Dynamic Protection schemes Considering the RI Problem

In this section we introduce extensive simulation results with regard to the performance of DPP and NCP evaluated using the NSFNET topology. The simulation results presented in this section were obtained using the widely used network simulation framework Omnetpp. Moreover, all plotted values have a 95% confidence interval not larger than 0.5% of the plotted values. The evaluated protection schemes are the following. DPP implemented using LCP-FF, DPPNC and DPPNC+, also based on LCP-FF, but along with a Preference and Opportunistic Coding strategy respectively, and finally PNCP.

The following settings apply for the simulation results presented in this section:

- CRs arrive at a source WR according to a Poisson process with a inter-arrival mean time t following an incremental manner: $CR(t_1 = t)$, $CR(t_1 + t)$,....$CR(t_{n-1} + t)$
- The holding time of each connection is exponentially distributed with a mean of 50 time units. All CRs demand the provisioning of both a primary and backup lightpath (resilient CRs).
- A backup lightpath is computed solely when its primary lightpaths was successfully provisioned.
- Each connection request requires a full wavelength on each link –grooming is not assumed. Therefore, the cost to send traffic along an optical link is $1U$.

- A periodical updating policy where NSI is disseminated it is instantaneously available at all WRs.
- Blocked CRs are not reattempted.
- Once a (primary or backup) lightpath is provisioned it cannot be reconfigured.
- WRs do not have wavelength conversion capabilities.
- Control messages losses as well as both propagation and connection setup delays are neglected. This assumption enables us to focus on the study of inaccurate NSI affected solely by high update times, which impact with regard to the blocking probability is dominating.
- For any source-destination pair each WR has at least 2 link-disjoint candidate paths.
- The candidate paths were computed off-line by means of Dijkstra's algorithm considering the number of hops as the routing metric and are sorted according to their cost, i.e., the first and last candidate path are the shortest and longest path respectively. The candidate paths will be recomputed if the network topology changes.
- 5 WRs as sources.
- 16 wavelengths per WR.
- Single-fiber per optical link.
- An offered load of 5 erlangs per source WR.
- NSI regarding channel occupancy is not affected by updating policies. Hence, only the inaccuracy NSI related to wavelengths availability is considered.
- For the sake of simplicity, it is assumed that NC operations are based on the Exclusive-Or operation (XOR) and are done over $GF(2)$, i.e., the Galois field of two or more data streams. In addition, it is also assumed all optical XOR gates.

Finally, only lightpaths with the same destination are suitable for NC in order to minimize the P_{cost} [18], [90]. Recall that even though the protection of lightpaths with different terminal vertices using a DPPNC scheme is possible, in this thesis we consider that this strategy is more scalable in order to minimize the complexity of the control plane operations required on the decoding process.

Figure 2.35 shows the blocking probability for all the evaluated protection schemes for large spectrum of update time values. As it can be observed, the performance of DPP, DPPNC and DPPNC+ is highly sensitive to the update time and it is optimal only for low update time values. This is not the case for PNCP because it computes lightpaths based solely on local NSI; hence, it does not require periodically dissemination of NSI.

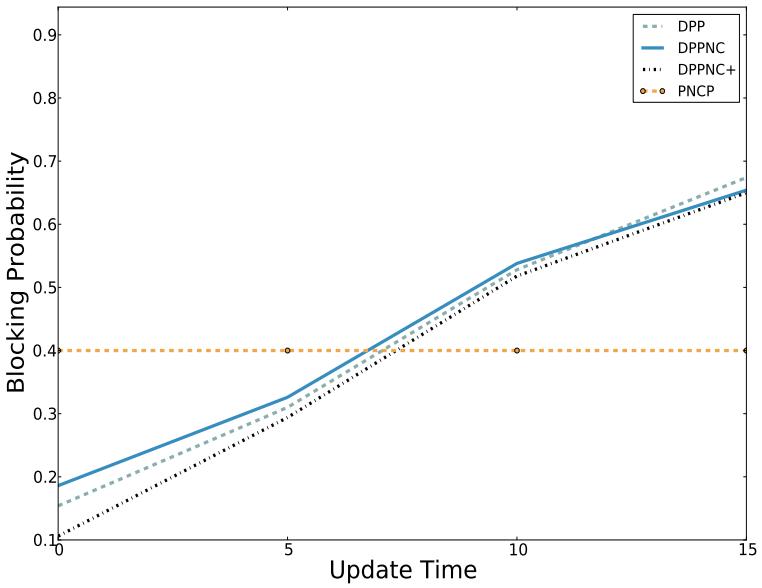


Figure 2.35: Blocking Probability vs update time.

On the other hand, Fig. 2.36 depicts the Average Protection Cost (APC) versus the Update time. The APC is computed as the total P_{cost} divided by the number of backup paths successfully provisioned. It is not surprising that among the evaluated schemes DPP yields the highest APC, an average of $3.50U$ per backup path. This is because of DPP inability to code traffic. PNCP yields the lowest APC, an average of $3U$, because of its preference for selecting shortest-routes as long as it successfully enables either NC. Notice that the APC of DPPNC+ is not as low as DPPNC since the former does not give preference to NC. Therefore, it can be stated that there is a tradeoff between the blocking probability and APC achieved by an NCP scheme. In addition, notice that the APC is not as sensitive to the update time as it is the case for the blocking probability.

Finally, Figure 2.37 depicts the blocking probability versus the interarrival mean time. In this test, it is attempted to evaluate the inaccuracy added by the dynamic of CR arrivals. Low inter-arrival mean times leads to highly inaccurate NSI. The contrary occurs with high inter-arrival mean times.

Based on the simulation results shown in this section, the following lessons were learned related to the study of dynamic proactive protection schemes under inaccurate NSI.

- The frequency of NSI dissemination as well as the inter-arrival mean time substantially impact on the blocking probability of protection schemes in dynamic scenarios.
- Predictive NCP schemes can outperform conventional protection schemes as well as NCP schemes which rely on global NSI under the assumption of realistic (greater than 5

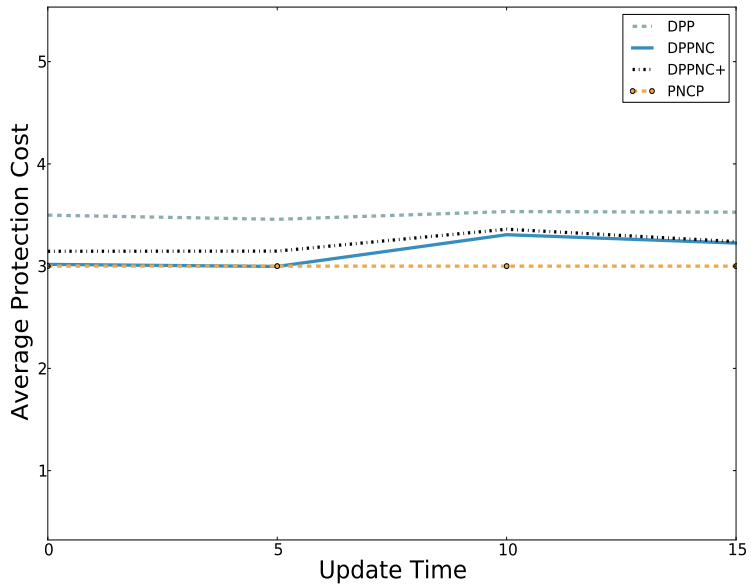


Figure 2.36: APC vs update time.

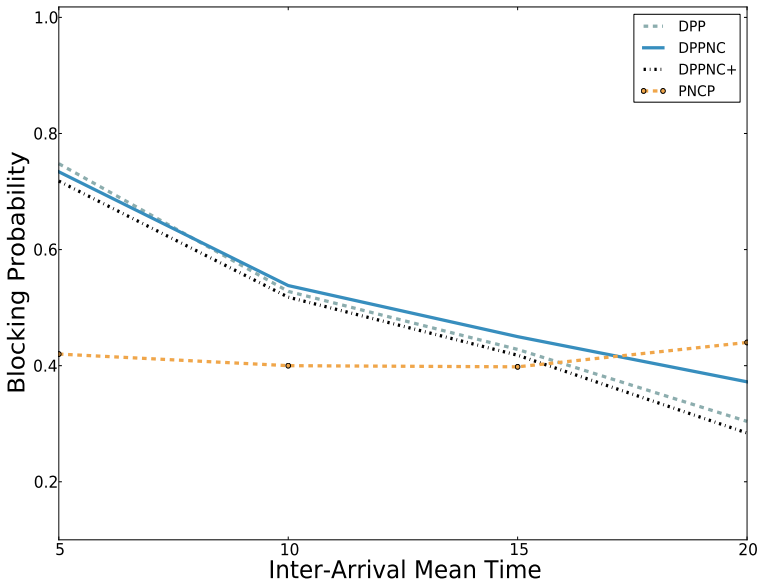


Figure 2.37: Blocking Probability vs inter-arrival mean time.

time units) update time values.

- The blocking Probability of an NCP scheme with a Preference Coding strategy is slightly higher than conventional DPP under inaccurate NSI.

3 Routing and Resilience in Multi-Layer Carrier-Grade Networks

This section is devoted to the study of resilience in multi-layer CGNs. To this end, this section first surveys distinct resilience schemes. Secondly, it distills the challenges for managing resilience. Then, it presents a novel study with regard to the deployment of NCP schemes in multi-layer CGNs. Finally, it introduces an innovative vendor-agnostic algorithm for dynamic discovery of cross-layer connections.

3.1 Resilience Schemes for Managing Resilience in Multi-Layer CGNs

The introduction of optical technologies in telecommunication networks enables high speed transmissions, necessary for the provisioning of services requiring a high amount of bandwidth, such as, IPTV or Video on Demand (VoD). This set of services is usually associated to a Service Level Agreement (SLA) clearly specifying and demanding the end-to-end connectivity expected characteristics and features that must be maintained. When these features refer to resilient communications, the IP/MPLS over Optical networks must be endowed with resilience mechanisms in order to provide fault tolerance services. This section surveys several protection and restoration schemes commonly deployed to meet such objective.

The schemes devoted to enhance the resilience level in multi-layer CGNs can be categorized into three main approaches according to which network layer (IP/MPLS or Optical) is executing the recovery actions require to restore affected traffic: ***1) Single-Layer Resilience (SLR); 2) Multi-Layer Resilience Bottom-UP (MLRBU)***; and ***3) Multi-Layer Resilience Top-Down (MLRTD)***.

SLR schemes do not require cross-layer coordination, i.e., interaction between network layers. Thus, the recovery actions are executed solely within one network layer. Many studies available in the literature argue that SLR schemes are inefficient and in some cases ineffective, because there are failure scenarios where SLR schemes might not be able to recover the affected traffic [3], [91]. In table 3.1 it is shown a collection of SLR schemes for both IP/MPLS and Optical network layers.

Table 3.1: Protection and Restoration schemes for IP and Optical networks.

Network Layer	Recovery Schemes	
	Protection	**Restoration**
IP/MPLS	IP Fast Reroute [92], Resilient Routing Layer (RRL) [93], Redundant Trees [94], NCP [55],	Equal Cost Multipath Forwarding (ECMF) [95].
Optical	Dedicated Path Switched WDM self-healing ring (DP-WSHR), Dedicated Path Protection (DPP), Generalized Loop-back (GL) [96], Backward Restoration (BRS) [97], Shared Line Switched WDM self-healing ring (SL-WSHR), Dedicated Line Protection (DLP), Redundant Trees [98], RPR (Resilient Packet Ring), Shared Link Protection (SLP), p-cycles [99], Hamiltonian cycles [100], NCP [42] [90].	Maximal Mutual Links (MML)[101], PCE based restoration [102], Threshold based selective restoration [103], Active restoration (AR) [104], Other related works: [105],[106],[107].

MLRBU is a resilience scheme which upon a failure event triggers the execution of recovery actions at the bottom network layer, e.g., the Optical layer; in case the affected traffic cannot be restored, then recovery actions are triggered at an upper network layer, e.g., the IP/MPLS layer. On the other hand, the operation of a MLRTD scheme is similar to a MLRBU scheme, but the former triggers recovery actions from the upper down to the bottom network layer.

It must be noticed that in this thesis we consider that the IP/MPLS layer is on top of the Optical layer, because the latest is used as the transport medium for the former. Other authors refer to the IP/MPLS domain as the Client layer and the Optical domain as the Service layer [61].

Authors [108] [109] are devoted to the study of MLRBU schemes. They claim that MLRBU schemes are more agile (lower recovery times) compared to MLRTD schemes because of their coarse-granularity recovery actions. Other studies such as [110] propose MLRBU schemes, because they argue that this type of resilience schemes can execute recovery actions with a finer granularity level. This is an advantage in order to select distinct protection paths for traffic flows with distinct characteristics, e.g., data traffic, video traffic (sensitive to network delay).

Both MLRBU and MLRTD schemes require cross-layer coordination in order to trigger the recovery actions at different network layers. In case that no cross-layer coordination is used, this type of resilience scheme is referred to as Uncoordinated Multi-Layer Resilience (UMLR). A UMLR scheme triggers recovery actions in a parallel manner at all network layers. This can lead to both suboptimal recovery actions (high P_{cost}) and inconsistent network states, i.e., a network setting leading to traffic loss.

A Multi-Layer Resilience (MLR) scheme can achieve cross-layer coordination by means of

two main approaches: ***1) Sequential Strategy***; and ***2) An Integrated Strategy***.

For a Sequential Strategy a pre-defined mechanism defines when a network layer is not able to restore affected traffic. When this occurs, recovery actions at an upper or inferior network layer are triggered. A sequential strategy uses two mechanisms to delegate recovery actions to a network layer.

1. ***Hold-off timers:*** are based on the use of predefined timeouts. Upon a failure event, the recovery mechanism of each network layer has a predefined timeout defining the time limit to restore affected traffic. Once this timeout expires, the recovery mechanism being executed stops, and the recovery mechanism of the upper or bottom network layer is then triggered. It is worth mentioning that there may be several timeouts assigned to each network layer, i.e., timeout per failure event. Each timeout could be customized according to the expected behavior to certain failure scenarios.
2. ***Signaling Messages***: refers to the use of notification messages among the network layers to either stop or start the execution of recovery actions.

On the other hand, an Integrated Strategy gathers NSI related to all network layers upon a failure affecting the network. Then, based on the collected NSI, a recovery action is chosen. A MLR scheme that uses an integrated strategy can follow three approaches.

For the first approach referred to as Fully Integrated MLR scheme, the Network Elements (NEs) within the IP/MPLS and Optical network layers can coordinate and trigger recovery actions. As it is shown in Fig. 3.1a, each NE embeds a set of recovery capabilities that enables the control and management of its own and other NEs features. Therefore, substantial modifications are needed for conventional routing protocols such as OSPF as well as for distributed control planes technologies such as GMPLS or ASON in order to enable the signaling required to support multi-layer recovery capabilities such as Multi-Layer Traffic Engineering (MTE). MTE provides enhanced recovery capabilities with low recovery times. In [111] it can be found extensions for conventional routing protocols such as OSPF to achieve MTE. For more information related to signaling technologies supporting MTE the reader is referred to [112].

On the other hand, for Relay MLR schemes, all the recovery capabilities are embedded into a centralized NE referred to as Relay Coordinator, which coordinates recovery actions in all network layers. In a Relay MLR scheme, the NEs (IP/MPLS routers and ROADMs from different vendors) are restricted solely to send NSI –they do not trigger any recovery action. As it is shown in Fig. 3.1b, a Relay MLR scheme uses the features of the management plane of all network layers (IP/MPLS and Optical) to coordinate and trigger the required recovery actions. Therefore, a Relay MLR scheme must be able to orchestrate multi-layer interactions. This implies the modification of the data models of IP/MPLS and Optical NEs, as well as the communication with Third-Party systems, such as a PCE, or an OpenFlow Controller.

A proposal of a Relay MLR scheme referred to as ONE Adapter can be found in [113]. The

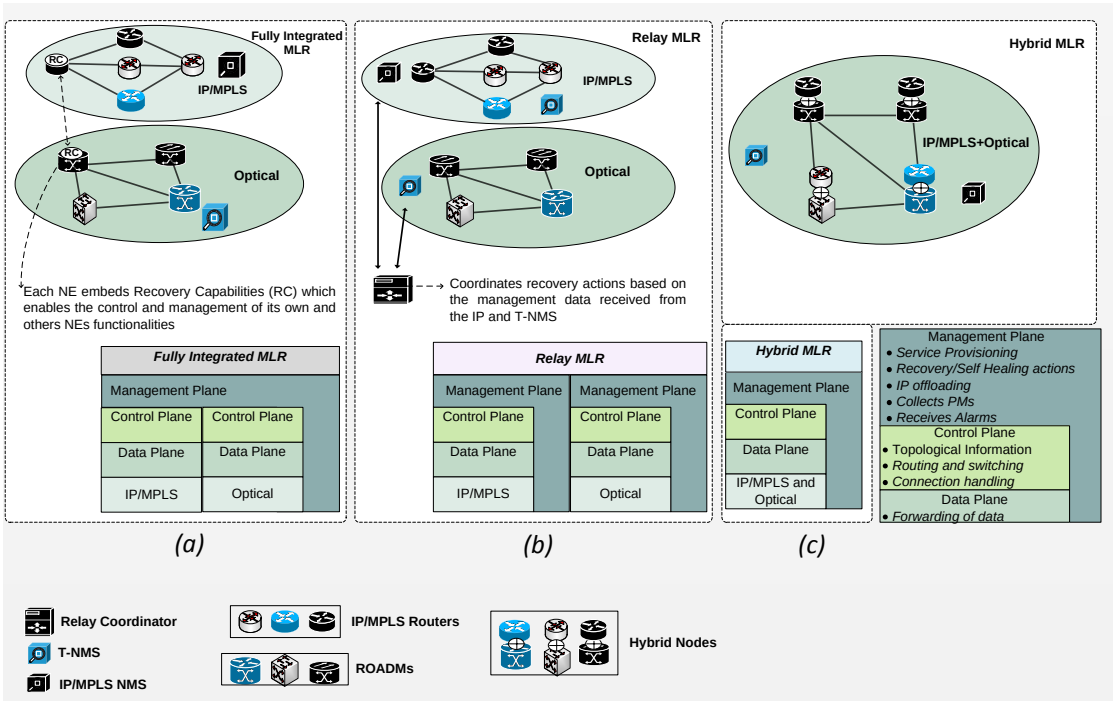


Figure 3.1: Integrated Strategies for MLR schemes: a) Fully Integrated MLR; b) Relay MLR; c) Hybrid MLR.

ONE adapter is a middle box which may communicate with the IP/MPLS and Optical layer and coordinate actions between them. The ONE Adapter enables the dynamic provision of a wide and diverse set of services, such as IP service Provisioning, IP offloading actions, and recovery actions.

Finally, another strategy used for Integrated MLR scheme is referred to as Hybrid MLR. As it is described in this thesis, in the recent years, there is a trend in network research referred to as hybrid optical network architectures, which consists in combining the functionalities of optical circuit and packet switching technologies. In these hybrid networks, hybrid nodes also known as PHRs can be programmed in order to enable packet switching, optical switching, low-level electronic packet routing or even all network features at the same time.

Moreover, in hybrid network scenarios both management and control planes of the Optical and IP/MPLS network layers are merged, as it can be observed in Fig. 3.1c. Thus, a Hybrid MLR scheme must be able to orchestrate a variety of services, as well as to reconfigure the routing and optical features of the PHRs. To this end, protocols such as ForCES have been proposed [114].

In Fig. 3.2 is depicted a taxonomy of the MLR schemes discussed in this section.

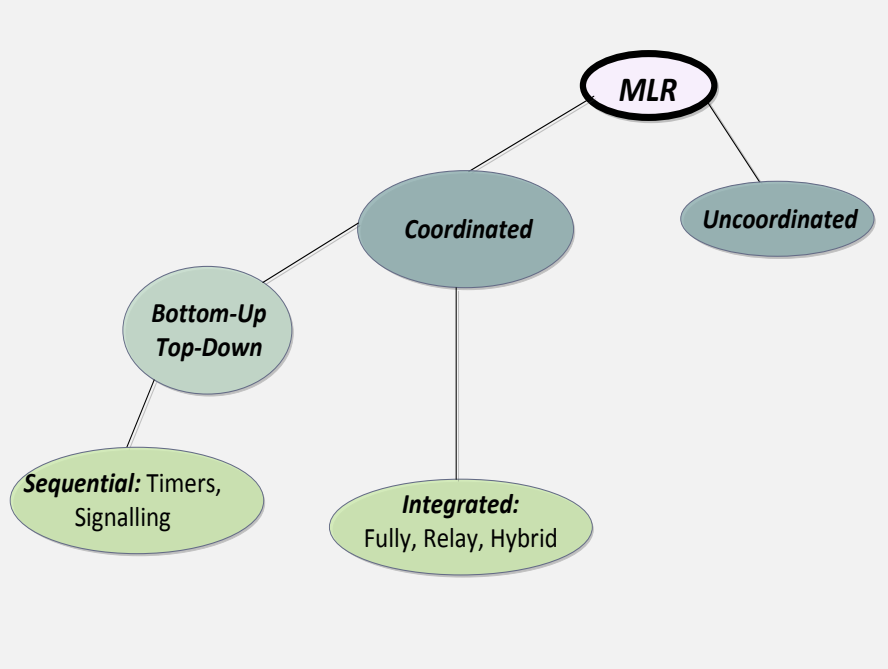


Figure 3.2: Taxonomy of MLR schemes.

3.2 Challenges for Managing Resilience in Multi-Layer CGNs

In this section it is differentiated three issues that hinder the deployment of resilience schemes in multi-layer CGNs: (1) Coordination of actions, (2) Correlation of NSI, and (3) Integration with Third-Party Systems. These three issues are the major reason preventing the deployment of MLR schemes proposed in current multi-layer CGNs. A comprehensive knowledge of these issues is mandatory to understand why the deployment of MLR schemes is not fully achieved at present. In the following lines, these problems are described in detail.

3.2.1 Coordination of Actions

The coordination of actions in multi-layer CGNs can be grouped into two sets. (1) The cross-layer coordination (communication among network layers); and (2) the intra-layer coordination (communication between NEs belonging to the same network layer technology). In the following lines it is illustrated in a compressive manner the need for both cross and intra-layer coordination in multi-layer CGNs by means of illustrative network scenarios. Even though, it may be debated that the network scenarios shown in this section came up as a result of a bad network planning. It must be noticed that these scenarios may show a network state caused by multiple failures, i.e., the scenarios shown represent a topology previously affected by one or more link failures. This section is intended to show that even if a careful network planning is made, a recovery scheme might not be able to recover the affected traffic in certain network scenarios. Therefore, both protection and restoration schemes are required to fully guarantee

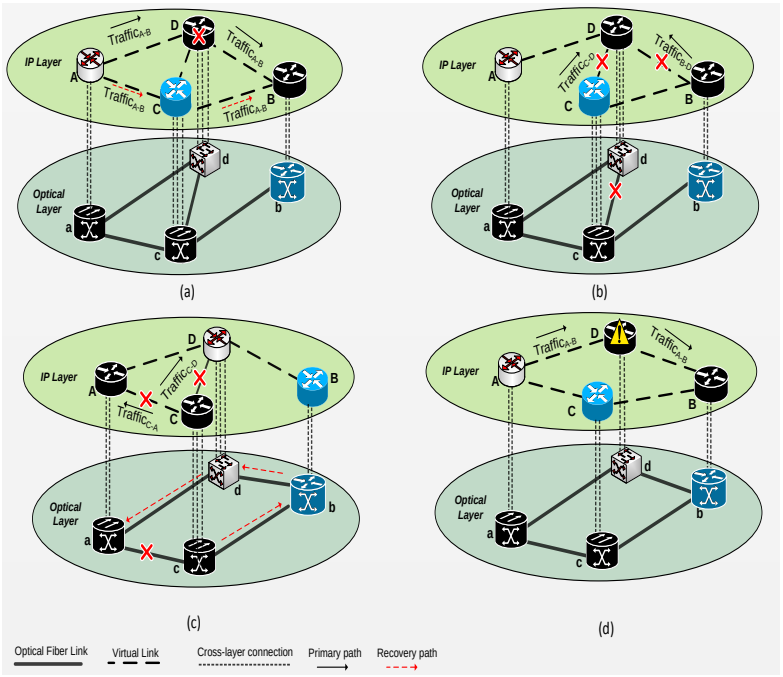


Figure 3.3: Operation of SLR schemes in multi-layer CGNs.

network resilience against failures.

As described in the last section, network resilience can be achieved by means of SLR schemes. Unfortunately, SLR schemes are inefficient in multi-layer CGNs. To illustrate the inefficiency of SLR schemes consider the network scenarios shown in Fig. 3.3. It must be noticed that the network scenarios shown in this section are modeled by the following network layers: the IP or the IP/MPLS layer, and the Optical layer. All IP or IP/MPLS inks are virtual links. This means that the optical links serve as a transport medium (server links) for the virtual links.

In Fig. 3.3a, the traffic sent by $routerA$ destined to $routerB(traffic_{A-B})$ is sent along the path $A-D-B$. If there is a failure affecting $routerD$, a SLR deployed at the IP layer can successfully restore the affected traffic, by rerouting this traffic along the path $A-C-B$. However, consider the scenario shown in Fig. 3.3b. If there is a failure on the optical link $c-d$, this failure affects both virtual links $B-D$, and $C-D$, because the optical link $c-d$ is the server link of these virtual links (the traffic on the links $B-D$ and $C-D$ is sent along the optical link $c-d$). In a failure scenario such as the one shown in Fig. 3.3b is impossible to restore the affected traffic by solely using a SLR scheme deployed at the IP layer, because at present the IP layer cannot trigger the provision of an optical circuit (even though there are recent advances for addressing this issue [115]). Consider that a possible recovery action could be to set up a new optical circuit between WRs c and d along the path $c-a-d$ for reestablish virtual link $C-D$, and set up a new optical circuit between WRs b and d along the path $b-c-a-d$ to reestablish virtual link $B-D$.

There are also network scenarios where a SLR scheme deployed at the Optical layer can be ineffective to restore affected traffic. In this regard, Fig. 3.3c depicts a scenario where there is a failure on the optical link $c-a$, affecting the virtual links $C-A$, and $C-D$ (assuming that the traffic sent along these links is forwarded along the optical link $c-a$). As a consequence, an optical circuit is setup between WRs c and d, along the optical path $c-b-d$, in order to restore the traffic sent on the virtual link $C-D$. In addition, an optical circuit is setup between WRs a and c, along the optical path $c-b-d-a$, to restore the traffic sent on the virtual link $C-A$. In this failure scenario a SLR scheme deployed a IP/MPLS network layer is ineffective to restore the affected traffic since a IP/MPLS recovery scheme cannot trigger the setup of an optical circuit.

However, consider the scenario shown in Fig. 3.3d. In this scenario there is a malfunction in $routerD$ (a software failure) affecting its data plane features, such as the traffic sent by $routerA$ destined to $routerB$ ($traffic_{A-B}$). A possible recovery action will be to reroute to $routerC$ the $traffic_{A-B}$. This recovery action can be triggered solely at the IP layer because NEs belonging to the Optical layer are not aware of any failures affecting an IP NE, i.e., the alarms generated by a malfunctioning router will be received only by the Network Management System (NMS) of the IP layer (IP-NMS).

Moreover, there are also scenarios where a SLR scheme deployed either at the IP or Optical layer cannot restore affected traffic. To illustrate this, we consider the topology shown in Fig. 3.4. This topology shows a network scenario where there are multiple simultaneous failure events. A failure on optical link $a-c$ that affects virtual link $A-C$, and there is a malfunction affecting $routerD$. A SLR scheme deployed at either the IP or the Optical layers cannot restore the traffic sent by $routerA$ destined to $routerB$ sent along the path $A-D-B$ because the Optical layer is not aware of the failure on $routerD$, and the IP layer cannot trigger the provision of a new circuit between WRs a and c. However, if the recovery mechanisms on the IP and the Optical layers would coordinate their respective recovery actions, it may be possible to set up a new optical circuit between WRs a and c along the optical path $a-d-c$, to reestablish the virtual link $A-C$, and then reroute the traffic from A destined to router B along the path $A-C-B$, restoring in this way the affected traffic.

An intuitive reasoning is to deploy recovery schemes at all network layers in order to cope with all possible failure scenarios. To this end, the coordination of recovery actions is highly required. Otherwise, an uncoordinated MLR scheme may lead to both inconsistent network states and high P_{cost}. With the aim of illustrating the negative effects of the lack of coordination in multi-layer CGNs, consider the network scenario shown in Fig. 3.5, which depicts a multi-layer topology based on the convergence of IP/MPLS and Optical technologies.

During normal operation, the traffic sent by $routerA$ destined to $routerE$ is sent along the virtual link $A-C$ (along optical link $a-c$) using the MPLS label $L6$. Suppose that there is a failure affecting the optical link $a-c$. As a consequence of this failure, the traffic sent from router A destined to E is lost. Thus, the recovery mechanism of the IP/MPLS layer reroutes

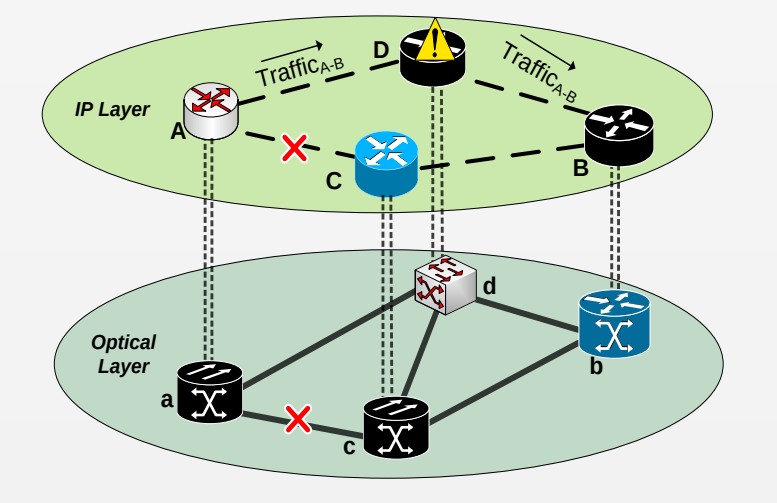


Figure 3.4: Operation of a SLR scheme in multi-layer CGNs with multi-failure events.

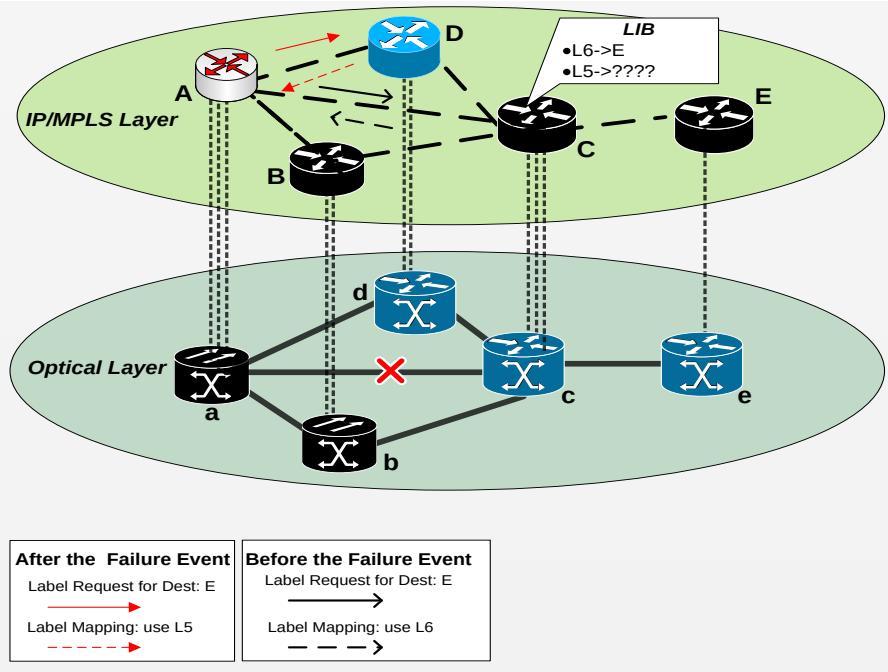


Figure 3.5: Suboptimal operation of Uncoordinated MLR schemes.

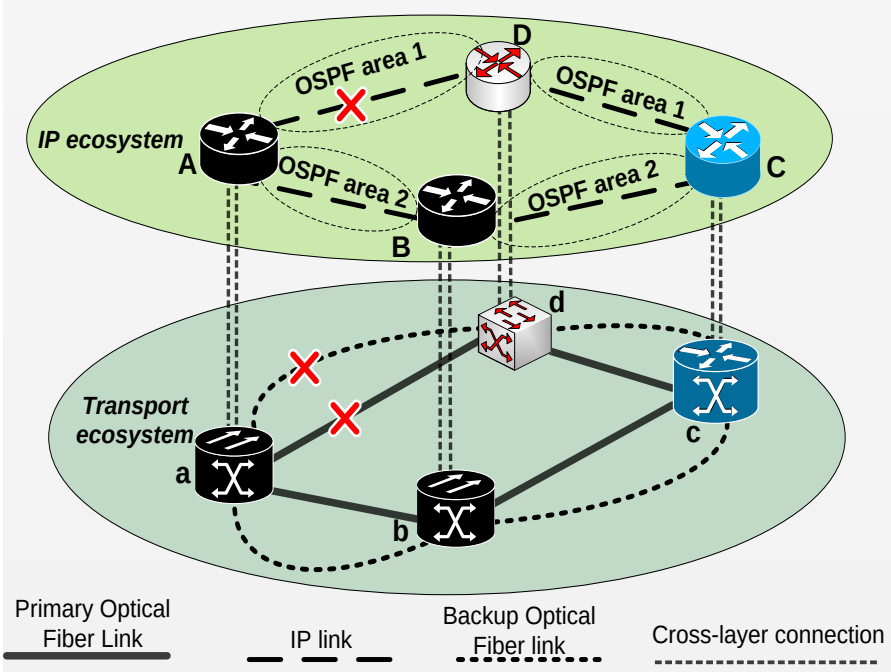


Figure 3.6: Single layer protection under the presence of multi-failure events.

the traffic sent by $routerA$ destined to $routerE$ by assigning a new MPLS label ($L5$), and forwarding this traffic along the IP path $A-D-C-E$. In a parallel manner, the recovery mechanism of the optical layer triggers the provision of a new optical circuit between WRs a and c, along the optical path $a-b-c$, which reestablishes the virtual link $A-C$. Once this virtual link is reestablished, the IP routing mechanism reroutes the affected traffic.

A recovery policy of the IP/MPLS layer states that the label $L5$ (the MPLS label used for all traffic sent from A destined to E) must be maintained, i.e., no to swap label $L5$ with another MPLS label such as $L6$. However, $routerA$ acquired the MPLS label $L5$ from $routerD$. Thus, $routerC$ does not have label $L5$ on its Label Information Base (LIBs). This will cause that $routerC$ drop all traffic with a $L5$ label.

The scenarios described so far in this section are valid for SLR and UMLR schemes where the recovery action is computed after of the failure event occurs, i.e., restoration schemes. However, both UMLR and SLR schemes based on protection actions, i.e., protection schemes can be also inefficient with the aim of recovering affected traffic in multi-layer network scenarios. In this regard, consider the network scenario depicted in Fig. 3.6, where a SLR scheme based on a DPP strategy is deployed at the Optical layer. In addition, notice that within the IP layer broadcast segmentation is enabled by means of OSPF areas at the IP layer [116].

In case that both primary and the backup optical links between WRs a and d fail. As a result, the virtual link between routers A and D also fail. In order to restore the affected traffic it is necessary to set-up on the fly a new virtual link between routers A and D – restoration

scheme is required. This is so because the recovery actions of a protection scheme are defined at the network planning phase.

It is worth mentioning that is not possible to restore the affected traffic, by rerouting it along the path $A-B-C-D$, since the virtual link $A-B$ belongs to a different OSPF area from link $A-D$.

By carefully observing the network scenarios described in this section it can be concluded that if recovery actions are only executed at a certain layer it will not be possible to restore affected traffic. For this reason, a SLR scheme might be inefficient in multi-layer networks. Thus, coordinated MLR schemes are required. Moreover, the coordination of recovery actions is a must in multi-layer CGNs in order to recover affected traffic in an efficient and agile manner.

The MLR schemes described in the last section are proposed to coordinate actions between the IP/MPLS and the Optical network layers aiming at achieving multi-layer resilience. Main of the issues related to both cross-layer and intra-layer coordination are rooted on the fact that each network layer has its own NMS System (IP-NMS and T-NMS). As a matter of fact, to manage NEs belonging to the same network technology—intra-layer coordination– each network vendor imposes its own NMS. This leads to both an unnecessary duplication of actions, and difficult configuration tasks for network administrators, because each vendor has its proprietary configuration data model. This situation is magnified at the IP/MPLS layer because a network administrator often uses a CLI interface for the configuration of NEs.

As a consequence, the use of control plane proposals such as GMPLS or ASON is becoming a widespread practice among network operators. However, neither of these two technologies provides the required granularity level for the provisioning tasks required by recovery actions. This is the reason why mainly in the IP/MPLS layer, network operators rely on the use of management protocols such as Simple Network Management Protocol (SNMP).

Nevertheless, SNMP has been mainly adopted for monitoring purposes, rather than for configuration tasks. This mainly occurs for two reasons. (i) The coarse granularity level offered by SNMP, and; (ii) the lack of a vendor-agnostic data model. These two issues motivated the development of more robust and flexible management protocols such as NETCONF. NETCONF is an XML based protocol that provides configuration actions with a high degree of granularity compared to SNMP, and also offers a vendor-agnostic data model termed as YANG [117]. NETCONF jointly with YANG can provide support to deal with the intra-layer interoperability issues. However, the development of a common data model that embraces the configuration syntax for all different network vendors is a difficult task, because each network vendor adds its own “ingredient” to the protocols running on their NEs. It is also worth mentioning that another issue requiring attention is the impact of a common data model might have on the current business models, because it may lead to reluctance among network vendors.

Despite of the several management features provided by NETCONF but there is an issue

that deserves further attention which is the coordination of recovery actions in multi-layer CGNs. As described in the previous section, two major approaches for multi-layer coordination are Sequential and Integrated strategies.

On one hand, for the case of MLR scheme based on an integrated strategy, current control plane protocols are not capable enough to orchestrate complex recovery actions, i.e., actions that involve the simultaneous configuration of several features of a NE. Moreover, the necessary signaling mechanisms must be defined to achieve cross-layer coordination [118]. Even though there are proprietary solutions, there is not a vendor-agnostic signaling mechanism available.

On the other hand, for MLR schemes based on a Sequential strategy there are several issues related to the escalation mechanisms that must be addressed. For instance, even thought the use of Hold-off timers is very pragmatic and easy to implement, adjusting it could be troublesome. Consider that a high Hold-off timer may prolong the restoration time, causing the loss of packets. On the contrary, a Hold-off timer too low may not be enough for restoring affected traffic in some network scenarios.

Nonetheless, the use of signaling messages are more efficient than timers [3], but are harder to implement in multi-vendor environments due to standardization issues.

As a matter of fact, since the required signaling messages for cross-layer communication needs to be defined, the deployment of MLR schemes is limited. Moreover, the deployment of MLR schemes is also limited by intra-layer interoperability issues caused by multi-vendor settings.

On the other hand, the lack of coordination in multi-layer CGNs also affects other types of MLR schemes such as Hybrid schemes. Indeed, several issues must be considered to the proper achievement resilience in hybrid networks. Figure 3.7 illustrates as a didactic example a failure scenario in a hybrid network where node b is an hybrid node programmed as optical switch. The traffic from $routerA$ destined to routers B and C is groomed in optical lambda λ_1 and then is sent to hybrid node b. The hybrid node b does the dropping of λ_1 and forwards the traffic to $routerB$; $routerB$ receives the traffic destined to itself; then checks its forwarding table and forwards the traffic destined to $routerC$ by allocating it into optical lambda λ_2 so it can be able to reach its destination.

In the case that the cross-layer connection between $routerB$ and hybrid node b (marked as 1 in the Fig. 3.7) fails, the traffic destined to $routerC$ will be lost, because it is necessary to have a cross-layer connection on $routerB$ to establish virtual link between this one and $routerA$. However, the only cross-layer connection available in $routerB$ is already being use for the virtual link $B-C$. A possible recovery action could be to reconfigure the hybrid node b with both IP and optical routing features. Thus, a new virtual link can be established between router A and hybrid node b, then hybrid node b can be able to route IP $traffic_{A-C}$ without dropping it to $routerB$ –notice that an optical switch cannot route IP traffic. Also note that a virtual link between routers A and C cannot be established because there are not cross-layer

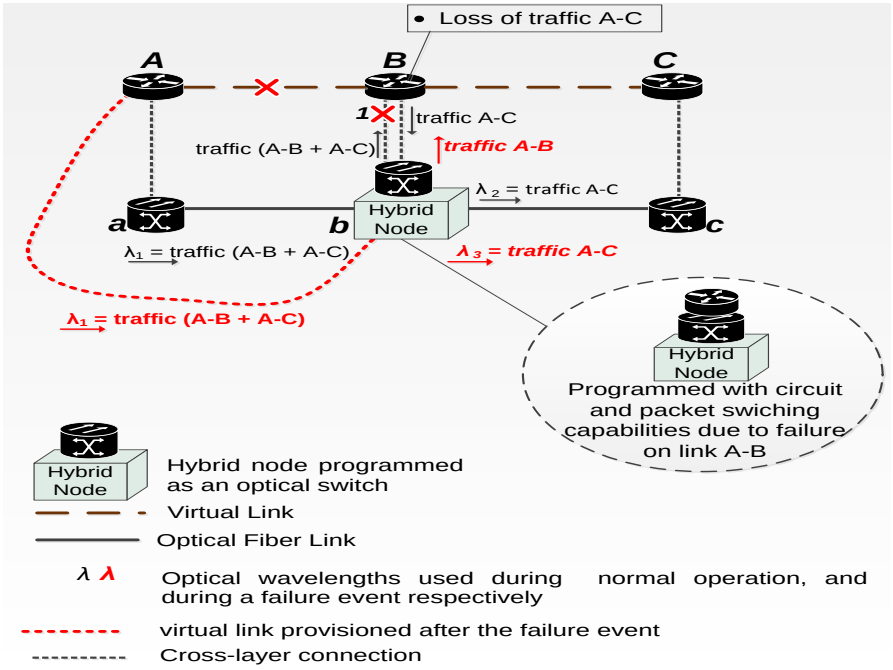


Figure 3.7: Negative effects caused by the lack of coordination in Hybrid MLR schemes.

connections available in $routerC$.

Therefore, in case of a failure event the functionality to be activated on a hybrid node (IP or optical routing) is a decision that a MLR scheme must take. To this end, a high level of coordination is demanded.

Based on the described in this section, it can be stated that the coordination of actions is a major drawback hindering the management of resilience in multi-layer CGNs. Even though the issues related to cross-layer coordination may be minimized by using a simple escalation strategy such as Hold-off timers, intra-layer coordination can be arduous to achieve in multi-vendor environments. Although major advances have been done on this aspect regarding the control plane, the management plane needs further development.

3.2.2 Correlation of NSI

In current multi-layer CGNs there is a lack of mechanisms that enable multi-layer topology discovery in multi-vendor settings in a dynamic manner. The multi-layer topology should reflect the set of physical (optical) paths followed by a virtual link, as well as the set of cross-layer connections (connection between an IP/MPLS node and a WR). In order to build the multi-layer topology it is required NSI from both IP/MPLS and Optical network layers. The dynamic discovery of the multi-layer network topology could be troublesome in multi-vendor settings because of two reasons; 1) The IP/MPLS NEs are not aware of the optical network topology and vice-versa, and; 2) the protocols available for topology discovery only operate

among NEs belonging to the same vendor.

To the best of our knowledge, the multi-layer topology can be only obtained statically (this is traditionally done by observing the status of the virtual links when disconnecting an optical link in the optical layer), but there is not an algorithm for dynamically obtaining the multi-layer topology.

The lack of a multi-layer topology discovery algorithm limits operations such as the alarm correlation and the severity assessment of a failure (Fig. 3.8). These two features are very useful for a network operator to know how failures in a server layer could affect a client layer [119], e.g., computation of Shared Risk Link Groups (SRLGs), which it could be very handy for an MLR scheme to have accurate NSI regarding the scale of a failure, as well as how the network performance will be affected by a failure, e.g., assessing the Expected Traffic Loss (ETL) [120].

At present the most realistic solution for alarm correlation in multi-layer networks is LMP [121]. Nevertheless, LMP is a notification protocol; hence, it does not perform advanced computational tasks such as the assessment of the ETL.

To clearly illustrate the negative effects that the lack of correlation of NSI we consider the scenario depicted at the upper side of Fig. 3.8, where there is a failure affecting optical link $c-b$. In case that the multi-layer topology is known, it can be easily estimated that the failure affecting link $c-b$ impacts on the traffic sent along virtual links $A-C$ and $C-B$. As a consequence, an MLR scheme might compute the P_{cost} required with the aim of avoiding possible network congestion. Furthermore, an MLR scheme (specifically, a stateful MLR scheme) could store recovery state information, so in case optical $c-b$ fails again, the recovery paths are foreknown.

The features of alarm correlation, the severity assessment of a failure, and the multi-layer topology discovery, are all adding computational complexity in the NEs. As a result, it would be very difficult to embed these features in the NEs. Therefore, it may be more feasible (from a deployment perspective) to embed these functionalities into a centralized-server, such as a Relay MLR scheme.

Indeed, a Relay MLR scheme outperforms a Fully-Integrated scheme in scenarios where a significant amount of NSI needs to be processed to perform the online computation of a recovery path [113]. As a matter of fact, the implementation of a Fully-Integrated MLR scheme is difficult because the integration of the control and management planes is an arduous task due to several issues, such as the computational burden and the high complexity added to each NE. This limits correlation-actions such as the multi-layer topology discovery. Notice that the mechanisms provided by routing mechanisms such as OSPF for IP topology discovery may cause both routers congestion, and high convergence time in large network scenarios [122]. This can be more severe if this is extrapolated to the multi-layer topology scenario. Thus, regarding the execution of correlation-actions, it can be stated that the scalability of

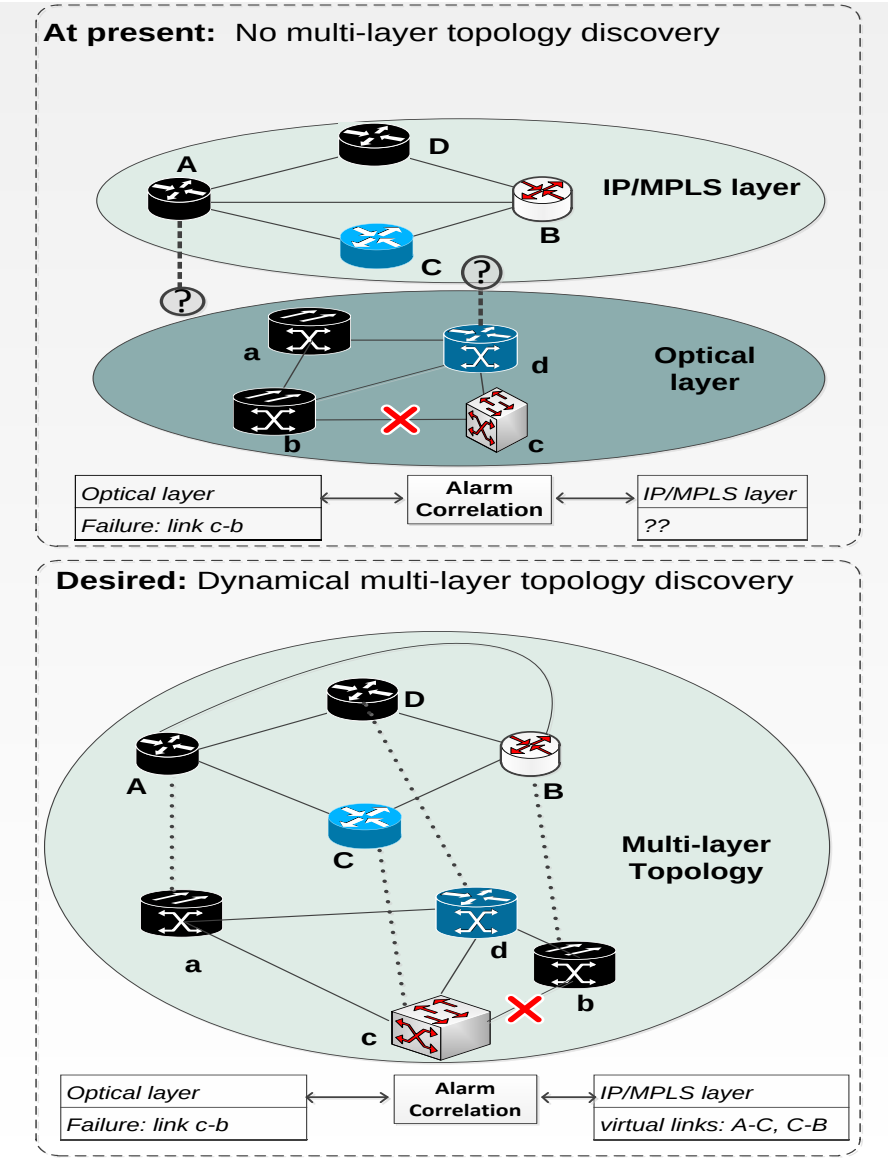


Figure 3.8: Correlation of NSI in multi-layer CGNs.

Fully-Integrated MLR schemes is low.

3.2.3 Integration of third-party systems and new network architectures

The integration of third-party systems and new network architectures such as a Path Computation Element (PCE) or an OpenFlow controller may enhance the resilience level of a MLR scheme. Therefore, MLR schemes should be able to leverage the features provided by other network architectures in order to improve their performance.

Notice that by means of a PCE it is possible to compute recovery paths in an agile manner [123]. In light of this, there are significant efforts related to standardize the integration of PCE schemes in current CGNs [124], [125], [126]. Moreover, there are several studies available in the literature dealing with path computation in multi-layer CGNs by means of a PCE , [127], [128].

In order to illustrate the features provided by a PCE scheme consider the scenario shown in Fig. 3.9. In this scenario, a PCE interact with a NE called as the Virtual Network Topology Manager (VNTM). The VNTM is in charge of provisioning tasks. It also has NSI related the multi-layer topology.

As it can be seen in Fig.3.9, an IP NMS informs a MLR scheme of the failure affecting virtual link $A-C$ (step 1). As a consequence, a MLR scheme sends to a PCE a path computation request for a new lightpath between nodes A and C using the Path Computation Element Protocol (PCEP), (see step 2). Then, the PCE checks with the VNTM the set of candidate optical paths available between WRs A and C (step 3). Hereafter, the VNTM initiates the provisioning of a new lightpath between nodes a and d (step 4). After the optical path is provisioned, the virtual link $A-C$ is reestablished. By means of a PCE an MLR scheme can coordinate path computation while reducing signaling overhead and delay as well avoiding the additional components and extended capabilities required at the NEs [107].

Despite of the advantages provided by PCE schemes, enabling interaction between the NEs and a PCE might be very sophisticated, e.g., consider a Fully-integrated MLR scheme, since this leads to an extra feature which is not native in current NEs, i.e., deploy the PCEP in the NEs.

In addition, consider that in a multiple failure scenario it is required to coordinate recovery actions in order to avoid sending duplicated path computation requests to a PCE. Otherwise, some contention process must be done by a PCE. Therefore, the coordination of actions must be done either by the MLR or the PCE scheme (“choose your poison”).

On the other hand, SDN is gaining momentum in recent years. The rationale behind SDNs is to enable the programmability of the forwarding table of NEs. SDNs are conceptually based on decoupling the control and data planes. In this way the mechanisms related to traffic forwarding (data plane) are placed within the NEs, whereas the control planes features

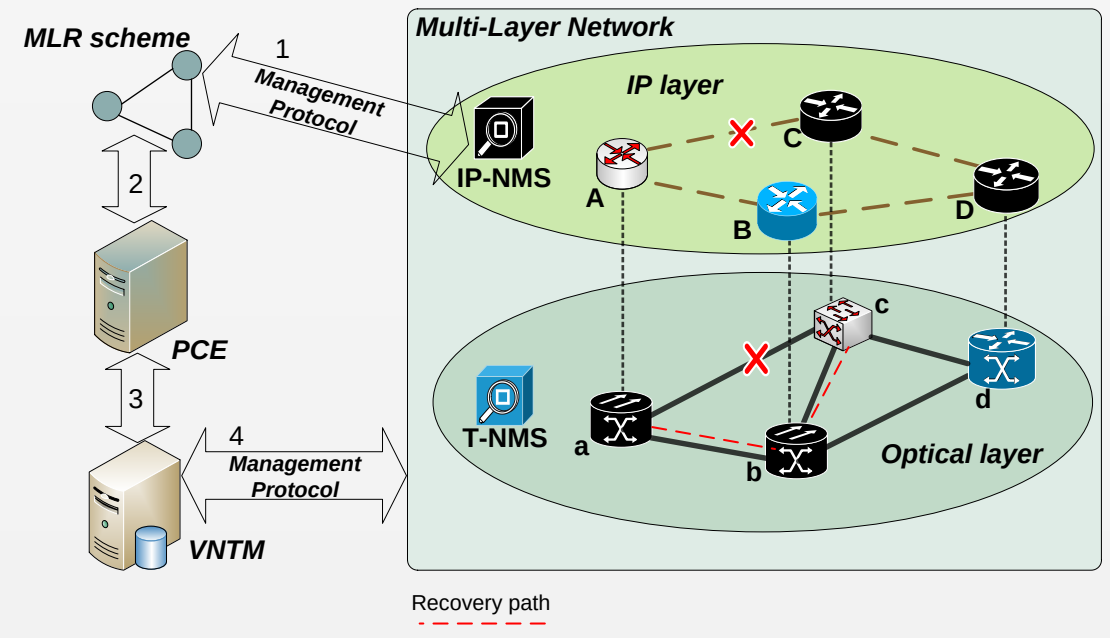


Figure 3.9: Integration of a PCE in multi-layer CGNs.

are embedded in a separate hardware entity called Remote Controller, which is typically a stand-alone server.

At present, network vendors are unwilling to expose the internal operation of their network products because of their business policies, which hinders the design and evaluation of new protocols, such as a TE features or an MLR scheme. As a consequence, the flexibility provided by SDNs for real-time programming of the traffic flow has been well received in network research.

One of the possible solutions facilitating a real deployment of SDNs is OpenFlow. OpenFlow is a protocol used to program the forwarding table of a NE by means of a Remote Controller. With OpenFlow the traffic flow can be controlled using several parameters such as VLAN ID, source/destination IP/MAC address, or a TCP port. It is worth mentioning that "flow" in the jargon of OpenFlow refers to packets or circuits (optical circuit); hence, optical circuit parameters can be also employed for controlling the direction of a flow, e.g., an optical wavelength.

Indeed, OpenFlow can be considered as a useful tool with the aim of designing more advanced MLR schemes. This is because OpenFlow offers a more flexible interface to configure the forwarding plane of NEs from different vendors in comparison with conventional management protocols, such as SNMP or NETCONF. One of the advantages provided by OpenFlow is that it reduces complexity of an MLR, since fewer protocols are required for configuration actions. This may result in agile MLR schemes.

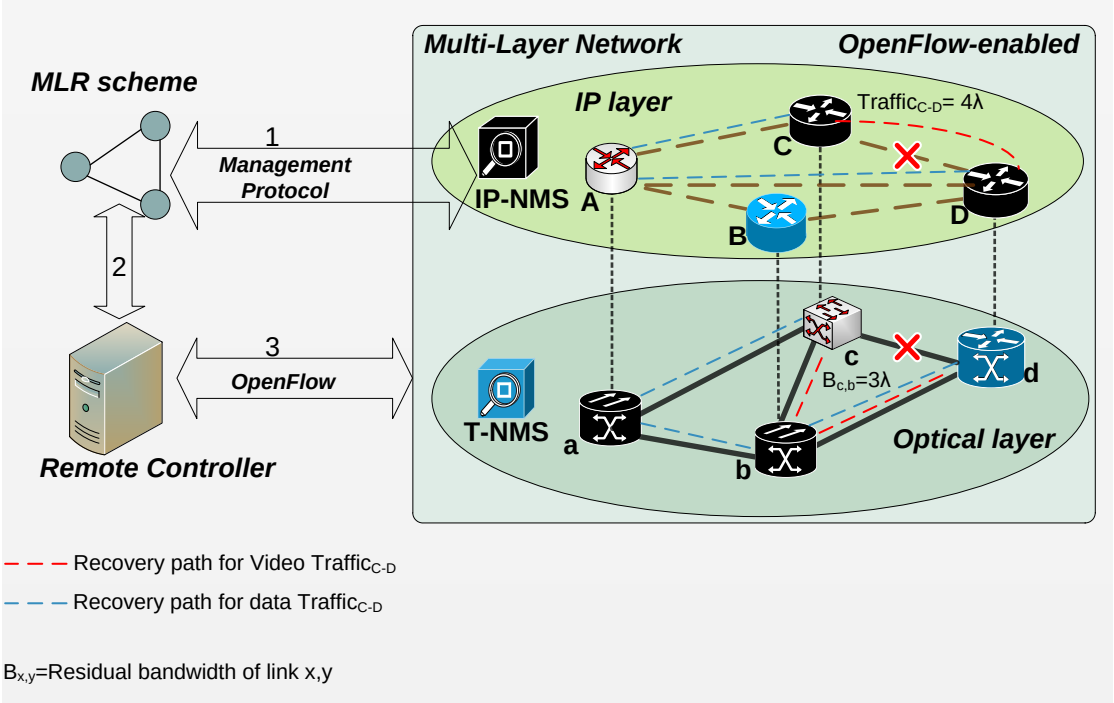


Figure 3.10: Integration of SDN in CGNs.

Figure 3.10 depicts a scenario that illustrates how a MLR scheme can leverage the features provided by OpenFlow. The failure in link $c-d$ affects the traffic sent by $routerC$ destined to $routerD$. As a consequence, an IP-NMS informs the MLR scheme of the failure event (step 1). Then, the MLR scheme sends a request to modify the forwarding table of WR c to the Remote Controller (step 2). The Remote Controller sends an OpenFlow message for modifying the forwarding table of WR c in order to route the delay sensible traffic (Video traffic) destined to $routerD$ along optical path $c-b-d$ and non-sensible delay traffic (data traffic) destined to $routerD$ along the optical path $c-a-b-d$, (step 3). This is so, because of the lack of network resources on the optical link $c-d$ to convey all traffic from $routerC$ destined to $routerD$. A MLR scheme can achieve this fine-granularity related to the selection of recovery paths by means of OpenFlow.

3.3 NCP in Multi-Layer CGNs

As mentioned in the previous sections, multi-layer CGNs formed by the convergence of IP/MPLS and Optical technologies are nowadays a widespread practice among network providers because of the vast transmission capacity offered by optical technologies. This section presents a promising NCP scheme referred to as DPNC+. The main objective of DPNC+ is to improve network reliability at the network planning phase, specifically for link protection in single failure scenarios in multi-layer CGNs –even though it can be easily extended for path protection. The novelty of the proposed scheme is that exploits cross-layer NSI for computing backup

paths. In particular, the main goals of DPNC+ are: 1) maximizing the amount of coded traffic; and 2) minimizing the P_{cost}.

To the best of our knowledge, there are no studies addressing the deployment of NCP schemes that leverage cross-layer NSI in multi-layer CGNs despite the fact that the current network backbone is mainly a multi-layer network. Indeed, this challenge is the rationale driving the design of DPNC+.

Cross-layer NSI is required to guarantee that both primary and backup paths are link-disjoint at all network layers. This is very relevant in multi-layer CGNs in order to avoid SRLGs that can lead to multiple failure scenarios. For instance, a failure affecting an optical link may affect a primary virtual link, i.e., a virtual link and its backup path at the IP/MPLS layer. Moreover, cross-layer NSI is useful to compute the P_{cost} at different network layers to enable link protection. Notice that even though the IP/MPLS (Packet) P_{cost} required to protect a certain group of virtual links using two different backup paths may be the same, the Optical P_{cost} may be different. Thus, computing the P_{cost} solely for a single network layer may lead to an improper deployment of an NCP scheme in multi-layer scenarios.

To illustrate the basic operation and limitations of an NCP scheme we consider the directed graph topology shown in Fig. 3.11a. In this scenario, as well as those shown in Fig.3.11b, c, d and e, we assume the following.

- The cost to send a data stream along any link is $1U$.
- The network resources required to send traffic along both ways of a link are the same.
- We consider a systematic coding strategy similar to the found presented by authors in [42] [45].
- All links are bidirectional.
- Traffic data units are fixed and equal in size.
- The proposed protection strategy is deployed at the IP/MPLS network layer.
- The backup paths associated to a certain set of primary connections that are jointly coded (protected) are link-disjoint.
- All primary virtual links follow a transparent model, i.e., transparent lightpaths are assigned to all primary virtual links.
- For simplicity, coding operations are done electrically, based on the exclusive-or operation (XOR) and are done over GF(2).

We also assume the following network model. A directed graph $G(E,V)$ representing a Virtual network topology, where V is the set of nodes, specifically IP/MPLS nodes, and E is

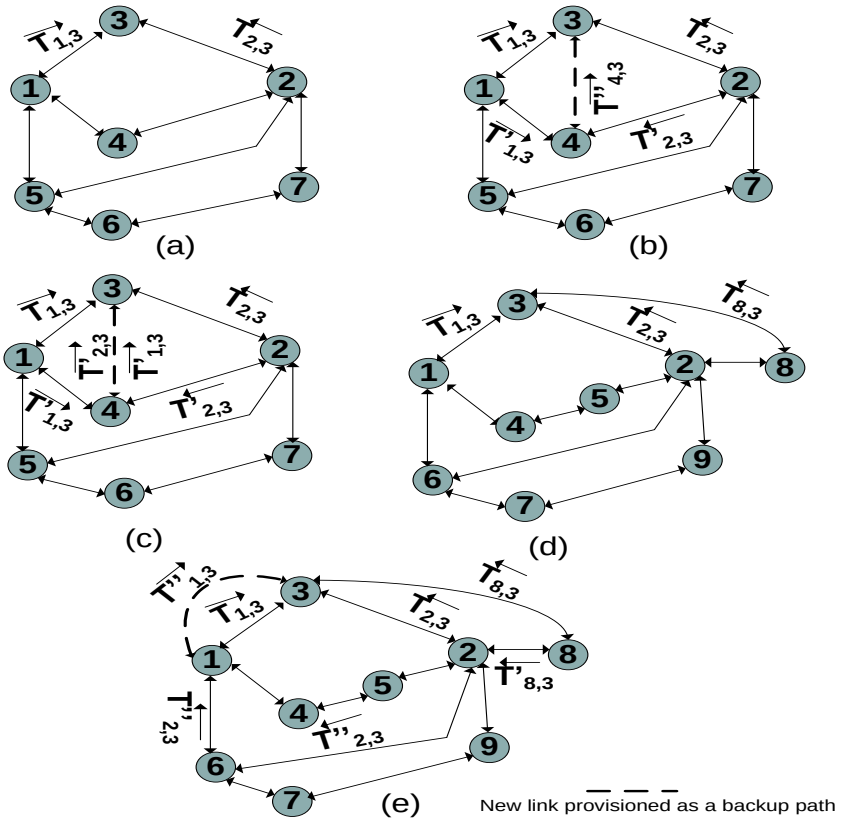


Figure 3.11: a) and d) Scenarios where it is not possible to code traffic; b) and e) Path provisioning to enable NC; c) DP operation.

the set of edges, specifically packet connections, i.e., virtual links. Our objective is to obtain a new graph $G'(E', V)$, which is a directed multigraph, where G is an edge-induced subgraph of G' with $E \subseteq E'$, such that the amount of coded traffic can be maximized. Our proposal can be useful to any NCP scheme (such as the ones using a systematic coding strategy) highly impacting on protecting those topologies where the network connectivity hinders the coding of traffic. In addition, the symbols and terminologies used in the rest of this section are listed in Table 3.2.

In the topology shown in Fig. 3.11a the traffic sent along links $e_{1,3}$ and $e_{2,3}$ ($T_{1,3}$, $T_{2,3}$) cannot be coded (protected) because there is not a link-disjoint backup path from $e_{1,3}$ and $e_{2,3}$ with node 3 as its terminal vertex. Recall one of the main goals of an NCP scheme is to code traffic aiming at reducing the bandwidth used for protection. In the case that the traffic $T_{1,3}$ and $T_{2,3}$ are jointly coded ($T_{1,3} \oplus T_{2,3}$) and sent along either link $e_{1,3}$, $e_{2,3}$, or sent on both links, this would be inefficient compared to the use of conventional proactive protection schemes such as DP. It is important to notice that coded traffic must be sent along a path link-disjoint from the primary links to be protected, i.e., $\rho_{1,3} \cap e_{1,3} \cap e_{2,3} = \emptyset$, and $\rho_{2,3} \cap e_{1,3} \cap e_{2,3} = \emptyset$, where $\rho_{x,y}$ is a backup path of link $e_{x,y}$.

As a consequence, two possible solutions can be followed. One is to use a DP scheme for those links that could not be coded. Contrary to an NCP scheme, a DP scheme does not code traffic. Thus, the path $(e_{1,4}\, e_{4,2}\, e_{2,3})$ and $(e_{2,4}\, e_{4,1}\, e_{1,3})$ can be the backup path for links $e_{1,3}$ and $e_{2,3}$ respectively.

The other possible solution includes the provisioning of a new link that serves as backup path. For instance, if a new link is provisioned between nodes 4 and 3 ($e_{4,3}$), it would be possible to code the traffic $T_{1,3}$ and $T_{2,3}$ and obtain $T''_{4,3}$ see Fig. 3.11b. This can be achieved by setting up node 4 as a *coding node* and link $e_{4,3}$ as a *coding path*. As a result node 3 receives the data stream $T''_{4,3}$, that codes $T_{1,3}$ and $T_{2,3}$ ($T''_{4,3} = T_{1,3} \oplus T_{2,3}$). Thus, in the case of a failure of links $e_{1,3}$ or $e_{2,3}$, node 3 can decode $T''_{4,3}$ and obtain $T_{1,3}$ or $T_{2,3}$ by executing $T''_{4,3} \oplus T_{2,3}$ or $T''_{4,3} \oplus T_{1,3}$ respectively.

Indeed, when traffic $T_{1,3}$ and $T_{2,3}$ are coded at node 4 (see Fig. 3.11b) the P_{cost} is $3U$ of bandwidth. But when conventional DP is used the P_{cost} is $4U$ of bandwidth (count the number of $T'x, y$ and $T''x, y$ on Fig. 3.11c).

Note that the terminal vertex of the protected links and the terminal vertex of their coding path is the same (node 3 in the topology shown in Fig. 3.11b), that is termed as node d. This holds true if it is assumed that only links with common terminal vertices are protected [1].

The other endpoint of the coding path is the *coding node* (node 4 in the topology shown in Fig. 3.11b). However, there can be more than one single *coding node*. Indeed, in a connected graph, all nodes i are potential *coding nodes*, where $i \epsilon \{1, ..., |V|\}$ and $i \neq d$. Aligned to this, we

[1] It is worth mentioning that there are studies available in the literature that deal with NCP with different destinations [44].

Table 3.2: List of Symbols and terminologies used thought section 3.2

Symbols and Terminology	Meaning
$G(V, E)$	Directed graph where V is the set of nodes and E is the set of edges.
$T_{x,y}$	Traffic sent by node x destined to node y, where $x, y \in V$.
$T'_{x,y}$	Replica of traffic $T_{x,y}$.
$T''_{x,y}$	Coded traffic sent by node x destined to node y.
Coding Node	Node that codes protected traffic.
$\phi()$	Function that returns the shortest-path between two nodes (we consider the number of hops as the routing metric).
$h()$	Function that given a path returns the set of nodes belonging to this one.
Ω	Set of potential coding nodes.
L	Set of links suitable for NC.
χ	Set of provisioned links to be used as backup paths.
$\mathscr{L}_m$	Set of lightpaths assigned to each virtual link, where $m \in \{1, ...\|E\|\}$.
β	Set containing all combinations of shortest-paths among the source vertices of the links to be coded. β=$\phi(n_k, n_{k+1})$,$\phi(n_k, n_{k+2})$,..,. where $k \in \{1, ...\|L\|\}$, and n_k is a source vertex of link k.

propose the following procedures to obtain the set of *coding nodes* offering minimum P_{cost} according to the links to be protected.

1. Only two links (with common terminal vertex) are to be protected:

- Remove links to be protected from $G(V,E)$, then compute $\Omega = h\{\phi(n_k, n_{k+1})\}$.

2. More than two links (with common terminal vertex) will be protected by enabling NC:

- First, obtain the set β, where $|\beta| = \binom{|L|}{2}$, and L is the number of links suitable for NC.
- Second, remove links to be protected from $G(V,E)$, then obtain $\Omega = \cap_{s=1}^{|\beta|} h(b_s)$, where $b_s \epsilon \beta$.

In the following lines we illustrate the procedure to obtain the set of *coding nodes* with a simple example. Consider the topology depicted in Fig. 1d where the links suitable for NC are: $L = \left(e_{1,3}, e_{2,3}, e_{8,3}\right)$. For this case $\beta = \phi(1,2), \phi(2,8), \phi(1,8)$, and $\Omega = 1, 6, 2, 8$. Therefore, a backup link may be provisioned between node 3 and any of the nodes belonging to the set Ω (such as $e^2_{1,3}$ see Fig. 3.11e) to be used as the backup path for the protected primary links.

The cost of provisioning new links may be expensive when there is no infrastructure currently in place, such as dark fiber. However, if the links to be provisioned are virtual links, e.g., IP/MPLS label switched paths (LSPs) in a multi-layer network scenario, the backup link provisioning process is related to: 1) the availability of physical resources (transponders, optical wavelengths); and 2) the graph properties of the optical topology, e.g., graph connectivity.

Moreover, all coding nodes belonging to the set Ω offer the same P_{cost}. This holds true assuming that the cost to send a data stream along a given link is the same independently of the path length.

In the scenario depicted in Fig. 3.11e, the P_{cost} required to protect links $e_{1,3}$, $e_{2,3}$ and $e_{8,3}$, i.e., $P\left(e_{1,3}, e_{2,3}, e_{8,3}\right)$, is $4U$ (U is a network resource unit) if link $e^2_{1,3}$ is provisioned to be used as a backup path. Notice that node 2 codes the traffic $T'_{2,3}$ (not shown in Fig. 3.11e) and $T'_{8,3}$, producing $T''_{2,3}$. In a similar manner, node 1 codes $T'_{1,3}$ (not shown in Fig. 3.11e) and $T''_{2,3}$ producing $T''_{1,3}$. This traffic is then sent along the recently provisioned backup path. Therefore, the path traversed by the coded traffic is $(e_{8,2}, e_{2,6}, e_{6,1}, e^2_{1,3})$. Moreover, if a new link $e_{6,3}$ is provisioned as a backup path the P_{cost} is also $4U$, since $1U$ is needed for paths $e_{1,6}$ and $e_{6,3}$ respectively, and $2U$ for path $e_{8,2}$, $e_{2,6}$. Nevertheless, we must consider that in a multi-layer scenario, equal protection costs computed at the virtual topology when using two different coding paths –such as the ones obtained when using links $e^2_{1,3}$ or $e_{6,3}$– may be different when the lightpaths assigned to each coding path are considered. For instance, even though Packet P_{cost_1}=Packet P_{cost_2}, it can be possible that the Optical $P_{cost_1} \neq$Optical P_{cost_2}, where P_{cost_1} and P_{cost2} are protection costs obtained when using two different coding paths.

The scenario described in Fig. 3.11 illustrates how to provision backup links to be used as backup paths in such a way that the amount of coded traffic is maximized. As a result, P_{cost} is minimized when an NCP scheme is used in single layer networks. In the following section, we plunge into several issues that need to be addressed to provision backup links in multi-layer scenarios.

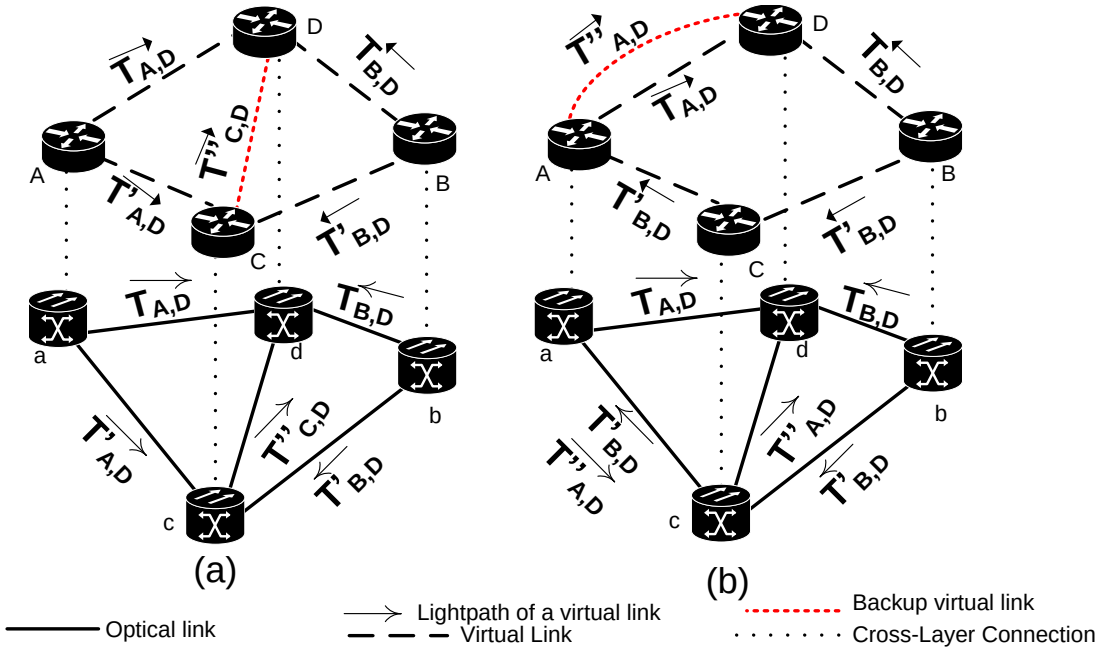


Figure 3.12: a) Multi-layer protection with router C as a coding node; b) Multi-layer protection with router A as a coding node.

3.3.1 Operation of DPNC+

This section introduces a novel NCP scheme for multi-layer networks namely DPNC+. The main purpose of DPNC+ is to improve network reliability by provisioning backup virtual links by means of cross-layer NSI. In particular we intend to: 1) maximize coded traffic; and 2) reduce the P_{cost} on a multi-layer network scenario.

In order to illustrate the operation of the proposed multi-layer protection scheme we consider the multi-layer network scenario shown in Fig. 3.12. The main objective pursued with this example is to elucidate the need of using cross-layer NSI when provisioning links to be used as backup paths.

In order to protect the traffic sent along the virtual links $e_{A,D}$ and $e_{B,D}$ using DPNC+ two approaches can be followed, represented in Fig. 3.12a and Fig. 3.12b respectively. The configuration shown in Fig. 3.12a consists of the following: 1) Virtual link $e_{C,D}$ is provisioned as a backup path; 2) router C is configured as a coding node. The Packet and the Optical P_{cost} are $3U$ each (count the number of $T'_{x,y}$and $T''_{x,y}$).

On the other hand, the configuration shown in Fig. 3.12b consists of the following: 1) a new virtual link $e^2_{A,D}$ is provisioned to be used as a backup path; 2) router A is configured as a coding node. With this configuration the Packet P_{cost} is $3U$, but the Optical P_{cost} is $4U$ (count the number of $T'_{x,y}$ and $T''_{x,y}$), because the primary and its respective protected traffic need to be sent along different paths (Packet and Optical paths) to avoid SRLGs. Thus, the

configuration shown in Fig. 3.12a should be the option chosen to protect the traffic sent along virtual links $e_{A,D}$ and $e_{B,D}$.

To compute the Optical P_{cost} the set of lightpaths ($\mathscr{L}$) associated to each virtual link is required, i.e., cross-layer NSI must be known beforehand. However, cross-layer NSI might be also obtained on demand by a multi-layer coordinator.

After carefully observing the example described in Fig. 3.12 it can be concluded that the backup link provisioning process must consider cross-layer NSI in order to address two issues. First, the backup path (including the provisioned backup link) and the primary links protected by this path must be link-disjoint at both Packet and Optical layers in order to avoid SRLGs. Moreover, primary virtual links suitable for coding must be link-disjoint at the Optical layer as well, in order to properly decode protected traffic, i.e., enable protection against double link failures. Second, both Packet and Optical P_{cost} must be computed to provision the most suitable backup path, i.e., obtain the smallest P_{cost}.

As described in the previous section, two solutions may be applied when NCP does not show enough resources to react to a link failure: 1) use DP; or 2) use backup link provisioning. The protection scheme presented in this section provisions backup links with the aim of enabling the coding of traffic, but also introduces a function to decide when this backup link must be used instead of DP. This is also useful because the avoidance of SRLGs strongly depends on the connectivity of the packet and optical topologies. Thus, when traffic cannot be coded or coding is expensive (a high P_{cost}), conventional DP is used.

Finally, Algorithm 6 shows the overall procedure for DPNC+. Notice that a backup link is only provisioned as long as a P_{cost} reduction is achieved in comparison with conventional DP. This is the reason why in an NCP scenario, the provisioning of a backup link must be done solely when it enables the coding of traffic.

3.3.2 Numerical Results with regard to protection schemes in Multi-Layer CGNs

This section provides numerical results related to the proposed scheme and other similar proactive protection solutions. The proposed protection scheme (DPNC+) is evaluated in terms of IP/MPLS and Optical P_{cost} (using the well known python graph library NetworkX, in comparison with DP (conventional proactive protection), and DPNC (NCP without cross-layer NSI) schemes. To ensure realistic findings the evaluated schemes were modeled over the multi-layer Spanish backbone topology see Fig. 3.13a.

The Virtual topology configuration of the multi-layer Spanish backbone topology is based on realistic network topologies extracted from [129], which is a vast online repository of real telecommunication networks. On this basis, we configured the Virtual Topologies, shown in Fig. 3.13b and Fig. 3.13c. We considered it more reasonable to evaluate more than one Virtual topology, while using only one physical topology design, because on a real multi-layer network scenario the Virtual topology design changes faster than the physical topology (fueled by the

Algorithm 6 Overview of DPNC+

Input: $(G(E,V), G_2(E_2,V_2), \mathscr{L})$
Output: (P_{cost})

{G and G_2 are the IP/MPLS (Packet) and Optical topology respectively.}
$P_{cost} = 0$ {Initialize the total Packet Protection Cost}
S =Group virtual links (E) by common destination node
for i in S **do**
  L =Create Sub-groups of minimum length equal to 2. {since at least 2 working links with common destination are required to enable NC.}
for L in S **do**
  for j in L **do**
    $DPNCP_{cost}$ =Run DPNC for each j (links suitable to NC or link subgroup), then compute the protection cost for each link subgroup {Protect each link subgroup with NCP strategy described in [45]}
    DPP_{cost} =Run DP for each link that could not be protected by DPNC, then compute the protection cost
    $\Omega = ObtainCodingNodes(j)$ {Obtain the set of coding nodes by using the procedure described in Section III.B.}
    $\chi_j = ProvisionBackupLink(\mathscr{L}, G, G_2, j)$ {Provision a backup link which endpoints are one of the coding nodes obtained and the terminal vertex of protected links, consider the both Virtual (packet) and Optical topologies in order to select the optimal backup link}

    $DPNC^+P_{cost}$ =Run DPNC+ for each j then compute the protection cost {protect each link subgroup using the logical backup links}
    if $DPNC\,P_{cost}$+DP_{cost}> $DPNC^+P_{cost}$ **then**
      $\mathscr{F}_j = DPNC\,P_{cost} + DP\,P_{cost}$ {$\mathscr{F}_j$ is the protection cost of sub-group j.}
      {protection group j is protected with DPNC combined with DP.}
      Tear-Down backup link χ_j
    else
      $\mathscr{F}_j = DPNC^+P_{cost}$
      protection group j is protected DPNC+.
  P^L_{cost}=$min(\mathscr{F})$ {Select the sub-group with the minimum P_{cost}.}

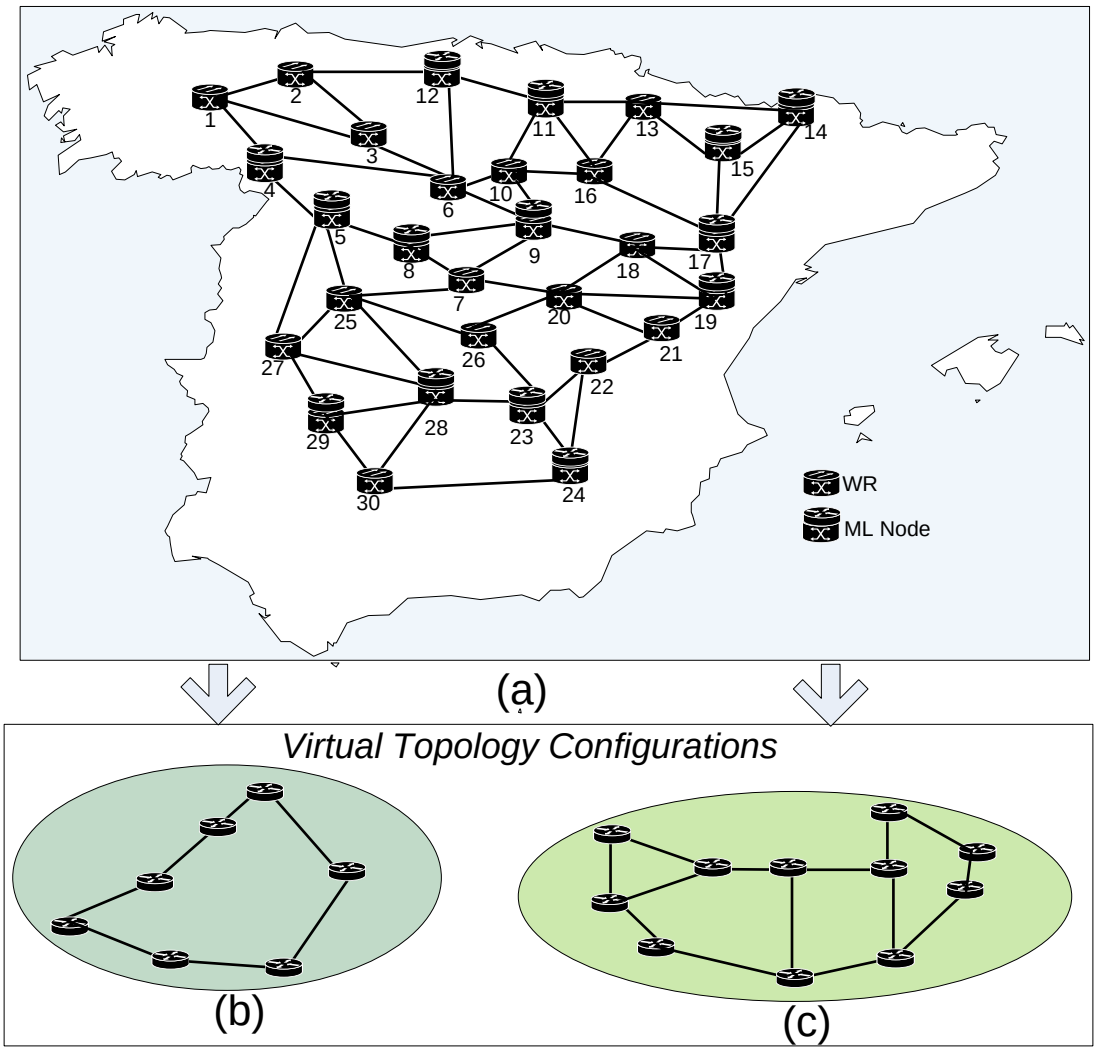


Figure 3.13: Multi-layer Spanish backbone topology; b) Virtual topology based on Sanren topology; c) Virtual topology based on Abilene topology

low economic cost, and ease of provisioning tasks).

Several trials have been carried out assuming the following settings: 1) the shortest-path routing algorithm used for route computations is based on the hop metric; 2) IP/MPLS router line cards of 100 Gbps capacity; 3) homogenous traffic demands of 20 Gbps along each virtual link; and 4) cross-layer NSI is known beforehand.

The IP/MPLS P_{cost} for the three evaluated protection schemes is depicted in Fig. 3.14. It can be seen that with DPNC+ a considerable reduction of the IP/MPLS P_{cost} is achieved, up to 50% reduction. Note that for the Sanren topology the IP/MPLS P_{cost} for DP and DPNC schemes is the same. This is expected since all nodes in this topology have an indegree equal to two. As a consequence, an NCP scheme (based on a systematic coding strategy) cannot

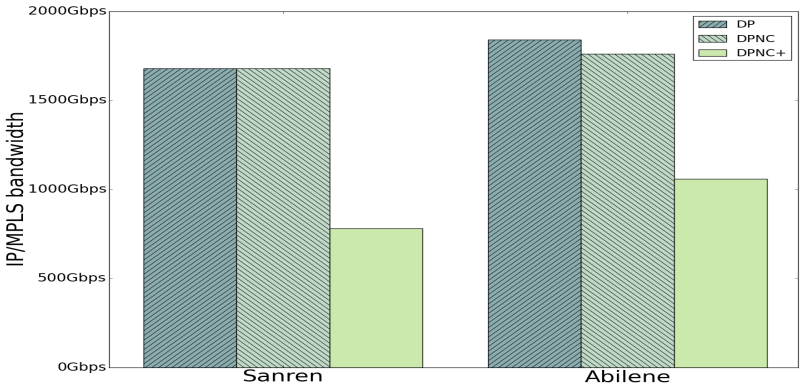


Figure 3.14: Comparison of IP/MPLS P_{cost}.

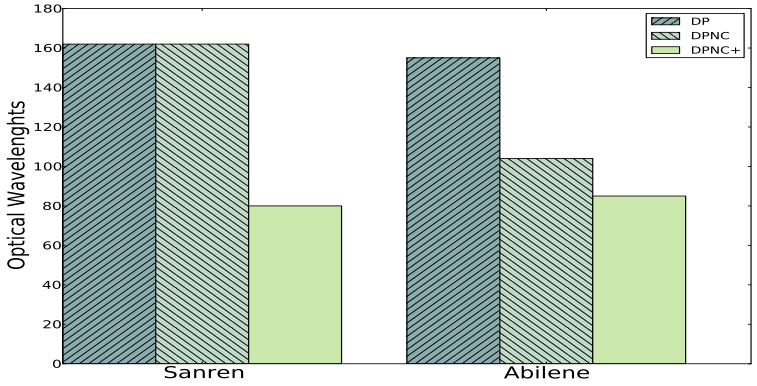


Figure 3.15: Comparison of the Optical P_{cost}.

code traffic, i.e., DPNC does not perform better than DP in this type of topology. Conversely, DPNC+ is able to code traffic due to its capability for backup link provisioning.

Regarding the Abilene topology, DPNC offers a smaller IP/MPLS P_{cost} in comparison with a DP scheme. However, using DPNC+ is possible to obtain a 40% and 35% P_{cost} reduction compared to DP and DPNC schemes respectively.

Furthermore, Fig. 3.15 quantifies the Optical P_{cost}. Similar to the results shown in Fig. 3.15, it can be observed that the DPNC+ scheme requires less Optical resources in comparison with the other schemes evaluated.

Finally, Table 3.3 shows the percentage of non-coded connections by DPNC and DPNC+ respectively. Based on the obtained results it can be stated that the proposed scheme maximizes coding in an effective manner, i.e., enable coding solely when the P_{cost} is reduced. Moreover, the evaluation results substantiate that DPNC+ significantly reduces both the Packet and

Table 3.3: Percentage of Non-Coded Connections.

Protection schemes	Evaluated network topologies	
	Abilene	Sanren
DPNC	100%	35.7%
DPNC+	14.2%	10.7%

Optical P_{cost} compared to other proactive protection schemes.

3.4 Interface Correlation in Multi-Layer CGNs

As mention in previous sections, cross-layer NSI is useful for MLR schemes in order to improve their performance. This was validated by the P_{cost} reduction achieved by NCP schemes that leverage cross-layer NSI such as DPNC+. However, NSI such as cross-layer connections is difficult to obtain in a dynamic and agile manner because of vendor interoperability issues. Indeed, to the best of our knowledge, the NEs belonging to the IP/MPLS are unaware of their directly connected neighbor at the Optical Layer, and vice-versa.

Therefore, driven by the high performance achieved when cross-layer NSI is available, this section presents a topology discovery algorithm referred to as Multi-layer Topology Discovery (MTD). MTD is able to discover cross-layer connections between an IP/MPLS node and a transport (Optical, Ethernet) node in a precisely and dynamic manner. The main advantage of MTD in comparison with similar topology discovery algorithms such as Cisco Discovery Protocol (CDP) [130] is that is vendor agnostic.

The first version of MTD algorithm is based on python and it is depicted in Fig. 3.16. As it can be observed, the operation of MTD is based on the correlation of statistics counters available on NEs. Statistics counters such as packets sent or received are available in NEs of different vendors and technologies. Moreover, they can be easily accessed by means of a management protocol such as SNMP, NETCONF or OpenFlow. By means of simple correlation algorithms and approximation methods, MTD is able to correlate the endpoints of a cross-layer connection.

In order to prove the efficiency of MTD, we build a real simulation testbed shown in Fig. 3.17, whereas the software model of this testbed is shown in Fig. 3.18. In Fig. 3.18, the correlation module is the MTD algorithm. MTD is split into two functional blocks. 1) The Correlation Engine which is in charge of the discovering the cross-layer connections; and 2) the Correlation Presenter, which is in charge of organizing the cross-layer NSI provided by the Correlation Engine in a legible way –xml was used for modeling cross-layer connection information.

Finally, NSI information is gathered by the Data Gathering Module by means of manage-

```
def compare(PR,PO):#compare polynomials checking their magnitudes
    match=[]#router with switches interface matching, router out switch in
    match2=[]#router with switches interface matching, router in swtich out
    for k in range(0,len(PR)):
        match.append([])
        match2.append([])
    #check Router out Optical In packets
    for k in range(0,len(PR)):
        for i in range(0,len(PO)):
            for j in range(0,len(PO[i])):
                if i==0 and j==0:
                    val=abs(PO[i][j][0][1]-PR[k][1][1])
                    val2=abs(PO[i][j][1][1]-PR[k][0][1])
                    match[k]=("router"+str(k)+"/1","switch"+str(i)+"/"+str(j),val)
                    match2[k]=("router"+str(k)+"/1","switch"+str(i)+"/"+str(j),val)
                else:
                    #print abs(PO[i][j][0][1]-PR[k][1][1])
                    #if i==1:return
                    if abs(PO[i][j][0][1]-PR[k][1][1])<val:
                        val=abs(PO[i][j][0][1]-PR[k][1][1])
                        match[k]=("router"+str(k)+"/1","switch"+str(i)+"/"+str(j),val)
                    if abs(PO[i][j][1][1]-PR[k][0][1])<val2:
                        val2=abs(PO[i][j][1][1]-PR[k][0][1])
                        match2[k]=("router"+str(k)+"/1","switch"+str(i)+"/"+str(j),val)
    return match,match2
def getx(x):
    x1=range(0,len(x))
    return x1
def getpol(T,Y):
 #Create a  vandermonde matrix
    V = np.vander(np.array(T), 2)### matrix of 3 columns.
                                 ##therefore the polinomial is of type ax**2+bx+c
                                 # size for 4 implies ax**3+bx**2+cx+d
 #Get Transpose
    VT = V.transpose()
  #multiply the transpose obtained before
    multVTV = np.dot(VT,V)
    multVTb = np.dot(VT,Y)
 # For the matrixe 3x3 apply cholesky factorization,  obtain L
    cholesky = np.linalg.cholesky(multVTV)
 #Solve linear system = Vtb,
    sol1= np.linalg.solve(cholesky, multVTb)
 #Con el resultado anterior resolvemos Lt x = y, put  cholesky  transpose
    sol2=np.linalg.solve( cholesky.transpose(), sol1)
#Create polynomial
    polinomio = np.poly1d(sol2)
    #return polinomial of type ax*2+bx+c
    return polinomio
```

Figure 3.16: MTD algorithm.

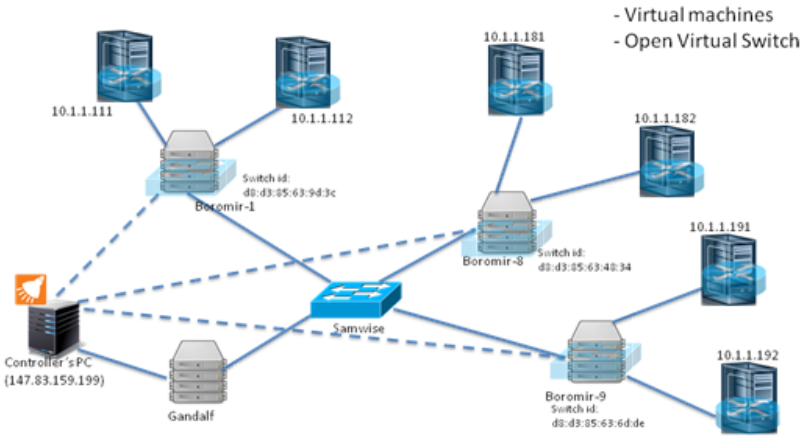


Figure 3.17: Testbed scenario for the evaluation of MTD.

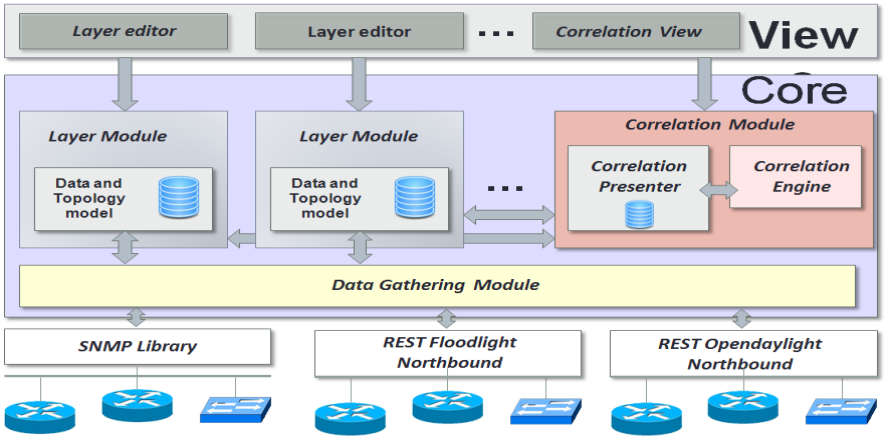


Figure 3.18: Testbed Software Modular View.

ment protocols such as SNMP and OpenFlow.

As shown in Fig. 3.17, we build a multi-layer topology formed by 6 IP routers running Juniper and Quagga [131] software; and 4 virtualized transport switches. The traffic model was created using the well-known traffic generation tool called MGEN [132]. For the network topology shown in Fig. 3.17, MTD has a 100% hit related to discovery of cross-layer connections.

4 Evaluation of New Trends for Routing and Resilience

This section is devoted to the study of new network architectures that can improve the performance of routing and resilience in CGNs. To this end, we propose a new PCE scheme referred to as Context-Aware PCE. The proposed scheme leverages ILSA schemes in order to enhance conventional PCE schemes to fully exploit the advantages provided by new communication models. In the following lines we describe the future challenges and trends for routing and resilience, as well as the proposed PCE scheme.

4.1 Future Challenges for Routing and Resilience

"Networking" as a single word is undergoing a noticeable evolution. On one hand, the advent of novel network paradigms such as SDN, Cloud Networking or Network Virtualization, all as a whole requiring significant changes in the currently deployed network architecture. On the other hand, network users are offered with new services and applications, all accessible from anywhere and at anytime. Fueled by the continuous evolution of networking, the research community started to seek new solutions aiming at optimizing network resources utilization, while facilitating the birth of new markets and business models. It is a must to have a comprehensive knowledge on where the network is and where is the network going in order to have the opportunity to propose new solutions. This section is devoted to describe the challenges faced by current networking architectures.

For many years, Internet has been constantly evolving in a wide set of areas e.g., technical, social, etc, what has been demanding a continuous effort from the scientific community to face the technological challenges linked to this unstoppable evolution. The socialization of Internet as well as the rapid dissemination of new user-friendly/appealing services and applications are both fueling network connectivity to become a basic need for users. Thus, it is widely shared among the scientific community that the near future for Internet will draw a network scenario enriched by network features such as End to End security, Resilient Communications, Mobility, Traffic Engineering and Multi-Homing), with a huge volume of heterogeneous devices all demanding Internet connectivity anywhere, anyhow and at anytime.

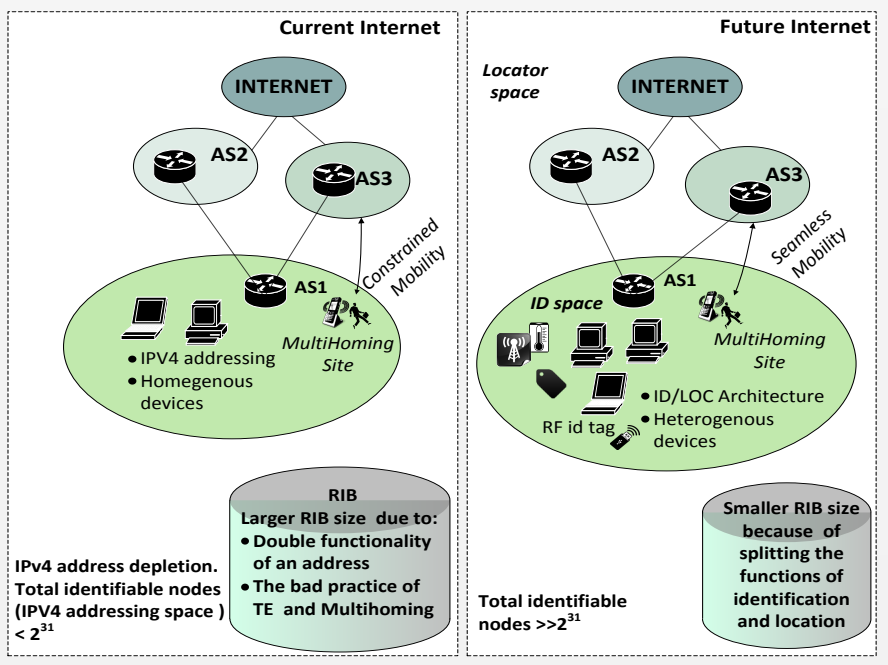


Figure 4.1: Comparison between current and Future Internet.

It is evident that the network protocols supporting the current Internet were not designed to provide such new features. As a result, network research community is pushing for the demise of the conventional location/host-oriented communication model deployed in current Internet and it is starting to migrate to the Future Internet, also referred to as the Internet of Things (IoT) see Fig. 4.1. An IoT architecture must undoubtedly overcome the limitations inherent to the currently deployed network protocols.

The IoT comprises a large and heterogeneous amount of devices demanding ubiquitous and seamless connectivity round the clock [133]. Unfortunately, though expected, the highly demanding constraints required by a Future Internet scenario cannot be appropriately supported by the current location/host oriented communication model, see table 4.1.

As a consequence, particular research efforts must be devoted to study the limitations caused by the existing IP-based addressing scheme, specifically with regard to two main issues: the depletion of addresses, i.e., the availability of the addressing space, and the semantic overload of addresses, i.e., double functionality of an address. The first refers to the fact that the overall size of the IPv4 address space is definitely not enough to support the current and expected increase in the density of identifiable NEs in Internet (worth noticing that the IPv4 address space has almost reached the end of its lifetime [134],[135].

As a matter of fact, since the early days of the Internet, IP is being deployed as the main underlying technology supporting routing and addressing strategies on the Internet. Despite some of the well-known weaknesses and limitations inherent to an IP-based addressing

Table 4.1: Requirements the Future Internet.

Requirements/Features
Devices demanding internet connection $\gg 2^{32}$
Smart devices with enhanced capabilities
Network features: social networking, green networking
Mobility without communication disruption (Full Mobility)
Proactive network reconfiguration
Set up/tear down connections in short-term basis
New network scenarios: Virtualized Data Centers, Smart Cities
New users roles: consumers+producers = prosumers

scheme, traditionally the scientific community has invested much more efforts in routing, in particular in the inter-domain area than in addressing [136], [137]. It was Geoff Houston in [138], who warned the scientific community about the addressing space reality, when he showed that the IPv4 depletion time would be shorter than the one previously foreseen by many organizations (some of them reaching the year 2030).

The second issue of the current addressing scheme is the so-called semantic overload of addresses, refers to the fact that current (IP-based) Internet addresses act as both locator and identifier. Thus, adopting a double functionality clearly imposes a burden on the current routing system, hence affecting several network features (e.g., roaming users or operator portability could be accomplished smoothly if this double functionality is removed).

Mobile communications are indeed affected by the double functionality problem. The following real scenario can better illustrate this statement. Nowadays, users are not statically connected to Internet but rather users are demanding connectivity on the move.

This novel mobility context imposes some effects on the connectivity. In particular, in the current routing architecture, when a user moves from one network to another one or changes his/her ISP, (because he/she moved to a new geographic location or he/she subscribed to a new ISP), his/her assigned IP address also changes. This IP address modification significantly degrades the communications quality or, even worse, causes a disruption of all established connections that are bound to this IP address. Notice however that, in the first case, the user only changes its location but its identity is certainly the same. Hence, while changes on the user location should be only reported to the routing layer, nowadays represent a change in the overall IP address. It is the routing (IP/MPLS) layer that should be aware of any change in a user's location. Even though there are protocols, such as Mobile IP that enable users mobility in a network, these are only a work-around that do not solve the root-cause of the addressing problems, i.e., the double functionality problem.

Resilient communications are also affected by the double functionality of the IP addresses.

Let's assume the case of a data center in which a 1:1 protection is used, therefore having a set of primary and backup servers in different geographic locations.

In the case a failure pops up in a primary server, the routing process at the network layer will shift all traffic routed to the failed server towards the backup server. This shifting action has a significant impact on all established connections with the failed server, potentially causing connections disruption.

Moreover, in today's routing architecture it is very hard to use an address to identify multiple hosts. For instance, if address x.x.x.x is assigned to node A, in the case there is a failure in node A, the process to reassign the address x.x.x.x to a node B can be troublesome. Protocols such as Hot Standby Router Protocol (HSRP) can provide support for this, but unfortunately these protocols are vendor dependent, i.e., they operate only among nodes from the same vendor.

On the other hand, multihoming features also affected by the double functionality of IPv4 addresses. Multihoming is a common practice nowadays that significantly fuels the geometrical growth of the routing tables. It basically consists in setting up different alternatives to connect a client to the network. In fact, multihoming comes up as an essential feature for network administrators mainly due to the two following characteristics: 1) it endows a network with fault tolerant capabilities, and; 2) it enables load balancing. These two patent benefits together with the fall of the cost of Internet connections have highly encouraged network administrators to offer and support multihoming.

But, how does a network manage multidomain?. To achieve multihoming, a site (Autonomous System) acquires a Provider Independent (PI) or a Provider Aggregatable (PA) prefix from its providers. It then announces them through all of its providers. PA and PI prefixes are blocks of IP addresses assigned by a Regional Internet Registry (an organization that manages the assignment and registration of IP addresses and Autonomous System (AS) numbers within a particular region of the world) to a site. The difference between them is that unlike PI prefixes, the PA-prefixes assigned to a site cannot be reused if a site changes its Internet provider.

A multihoming site using PI address space allocates its prefixes in the forwarding and routing tables of each of its providers. Therefore, PI prefixes are not aggregated. For PA prefixes, the Internet provider of a site could aggregate the customer (site) advertisement into a shorter prefix when advertising the prefix to other customers or peers. In the practice of multihoming an Internet Service Provider ISP has to advertise more specific (less aggregated) IP routing prefix to the Internet and rely on the traditional and problematic longest-prefix match route selection algorithm of Border Gateway Protocol (BGP).

In addition, to multi-homing features, the double functionality of addresses hinders the deployment of multi-interface applications. For instance, consider a scenario where a NE can have multiple addresses assigned according to the transport technology used for communication to this NE, i.e., legacy application use the conventional IPv4 wired networks, while new wireless technologies use IPv6 addresses.

On the other hand, the majority of routing and resilience schemes deployed in current CGNs are distributed schemes. It is well known that distributed schemes have several issues such as they are highly susceptible to NSI inaccuracy, high signaling overhead, interoperability issues, among others. As a result, the design of new routing architectures is gaining momentum in network research.

4.2 Trends for Routing and Resilience

As it has been mentioned in this thesis the main technologies commonly adopted in CGNs are IP/MPLS and WDM technologies. In a CGN, features such as fast connection-provisioning, recovery actions, and TE, are commonly achieved by source routing strategies using a distributed control plane scheme (e.g., GMPLS or ASON) handling connections setup and teardown in a short-term basis. However, sourced routing strategies have significant weaknesses when facing path computation actions, specifically in large network scenarios, where it is difficult to have precise knowledge of NSI.

As a result, a centralized entity referred to as PCE has started to gain momentum among both network researches and carriers. There are many studies already available in network research introducing contributions on PCE architectures, all devised for the conventional location/host-oriented network scenario. Despite the fact that the host-oriented model embeds several well-known short-comings, the main one referred to as the semantic overload of addresses as it was described in the previous section. As a consequence, the network community is focusing on the study of new network paradigms, such ILSAs and information-centric or context-aware communication models.

In order to address the limitations of the host-oriented model, a diverse set of ad-hoc solutions have been proposed. Most of these solutions are focusing solely on specific issues. Thus, these solutions introduce two harmful consequences. First, several different solutions must be deployed in order to provide the overall set of requirements for an IoT scenario. This may lead to chaos in multi-technology/vendor scenarios due to the costly actions required to deploy a large and diverse set of solutions. Second, the use of an isolated strategy to propose solutions for each individual issue may raise negative collateral effects on the others that are making the overall solution ever complex, such as difficulties in the deployment of communications protocols such as Session Initiation Protocol (SIP) and IPsec, i.e., NAT sensitive protocols.

There are other attempts with the aim of proposing solutions to face the short-comings related to the current location oriented communication model and offering an alternative to the traditional “OSIfied IP networks”. These research attempts are centered in two research lines: Non-disruptive approaches and clean slate architectures.

Among the clean-slate architectures it is worth to mention IPv6. IPv6 was proposed as an (evolutionary) alternative to cope with the exhaustion of addresses, conceptually supported

by enlarging the addressing space. However, as of today, network providers are reluctant to widely deploy IPv6 mainly because of two reasons [139], [140]: 1) the expenditure of resources, referring to the fact that the required tasks to migrate from an IPv4 to an IPv6 core require a considerable amount of time, and represent an operational expenditure not only in terms of firmware updates of NEs but also on IPv6 training for the operational personnel; 2) the migration process may cause an undesired disruption of the offered network services.

Moreover, while enlarging the addressing space may contribute to solve the depletion of addresses problem, it is clear that this does not have any effect on the semantic overload problem.

On the other hand, ILSA schemes falls into the set of Non-disruptive approaches. ILSA schemes are proposed as a way to address the issues related to the current addressing space. ILSA schemes deal with both the double functionality problem and the exhaustion of addresses by assigning an independent set of addresses for identification and location functions. For instance, the network layer supporting end-to-end connectivity operates with an address scheme commonly referred to as an Identifier (ID); whereas, the network layer responsible for location functions operates with an addressing scheme commonly referred to as Locators (LOCs). ILSA schemes have received a great acceptance in network research. As a matter of fact, conversely to the majority of clean-slate architectures, ILSA schemes are already available as commercial solutions and it is also a IETF standard protocol (with some active working groups)[141],[142].

It is worth mentioning that ILSA schemes and IPv6 can work jointly. As an example, the use of ILSAs along with an addressing scheme such as IPv6 could reduce the IP-based addressing limitations, especially in network scenarios requiring multihoming, traffic engineering and Full mobility [143]. Furthermore, ILSAs can smooth the migration of IPv4 to IPv6, what strongly lowers the barrier operators keep to deploy IPv6 on their IPv4 networks.

Moreover, ILSAs may slow down the address exhaustion issue, even though the latest seems to be solved by the huge address space provided by IPv6. However, migrating from IPv6 to IPv4 is a task not pleasant for network providers. An ILSA scheme can provide support to the migration process between addressing schemes, see Fig. 4.2. For instance, in a network of an ISP, the border routers can have IPv6 addresses assigned whereas the core routers remains untouched with the conventional IPv4 addresses. An ILSA solution may be able to map the IPv6 to IPv4 addresses and vice-versa. This is a reduction of time and tasks that reflects in OPEX and CAPEX. At present, the interoperation between IPv4 and IPv6 using ILSAs is a solution that is already being offered by network-vendors.

In addition, ILSAs schemes may boosts up other network features such as mobility. In this regard, consider the scenario shown in Fig. 4.3a. It can be observed that the user device keeps its same ID even in the case it changes its location site (ISP provider), without impacting on any already established. Furthermore, providers do not have to reassign new IDs to new users, keeping in some cases the hosts' configuration untouched.

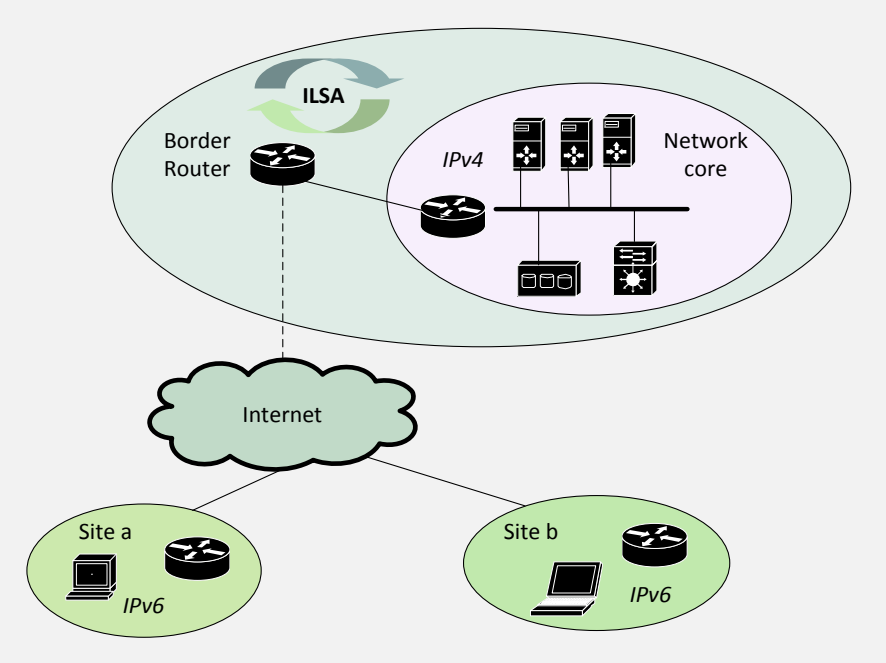


Figure 4.2: ILSAs aiming migration from IPv4 to IPv6.

A highly well positioned real use case for mobility, completely detached from user mobility is drawn nowadays in Virtualized Data Centers. Indeed, Virtual Machines) can be deployed anywhere (supported by real network infrastructure) regardless the address assigned to the IP/MPLS layer, freely moving (migrating) resources across different geographic locations or different racks within a data center. But Virtual Machines migration is not the only added value feature getting benefit from a potential ILSA scheme deployment. Nowadays, a failure in the infrastructure of a data center or a cloud model, will severely impact on the live services offered to the users.

It is possible to enable resilient communications by means of an ILSA scheme. Let's consider the "resilience scenario" shown in Fig. 4.3b, in this scenario a 1:1 protection scheme is employed, i.e., there are two Data centers, the main and the backup, for the purpose of offering fault tolerant services. In the case there is a failure in the main data center, a protection action is triggered for relocating the affected services to the backup Data Center., which consists in mapping the ID xxx to a different locator, the locator z.z.z.z.

The application layer is not aware of any failures in the network layer. The application layer as mentioned before is bound to IDs. In the case of the example shown in Fig. 4.3b, the application layer is not aware that the network element with the ID "xxx", is now a different node in a different geographic location. It is important to remark that even though the connections are not disrupted, the quality of delay sensitive communications, e.g., VoIP or video streaming, could be degraded [144].

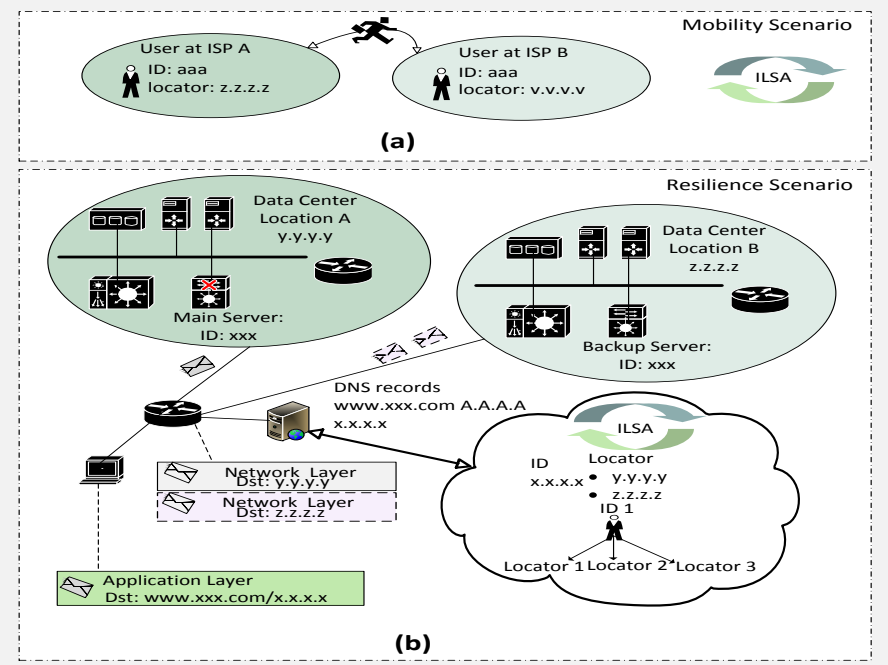


Figure 4.3: An ILSA scheme boosting up mobility and resilience features.

4.3 Dealing with availability and reachability of ILSA schemes

As mentioned in the last section the current internet and routing architecture embeds several issues. Indeed, recent studies including the Internet Architecture Board (AB) report [145], reveal that current Internet routing architecture is facing several scalability problems related to both the size and dynamics of the global routing table in the Internet's Default Free Zone (DFZ). For instance, the global routing table size in the DFZ has been growing at an alarming rate in recent years [146], till reaching now a total of 36.717 ASes that originate 355,262 IPv4 prefixes (see Fig. 4.4) despite several limitations such as lack of IPv4 addresses, strict address allocation and routing announcement policies. Although IPv6 deployment would remove the problem of lack of IPv4 addresses, there is a strong concern that the deployment of IPv6 on a large scale could result in a significant growth of the routing table.

The AB report identified the following sources behind the rapid growth of the global routing table in the DFZ:

- Multihoming.
- Traffic engineering.
- Non-aggregable address allocations.

In [146] authors conclude that address fragmentation, caused by multi-homing and load balancing is the major reason of BGP table growth.

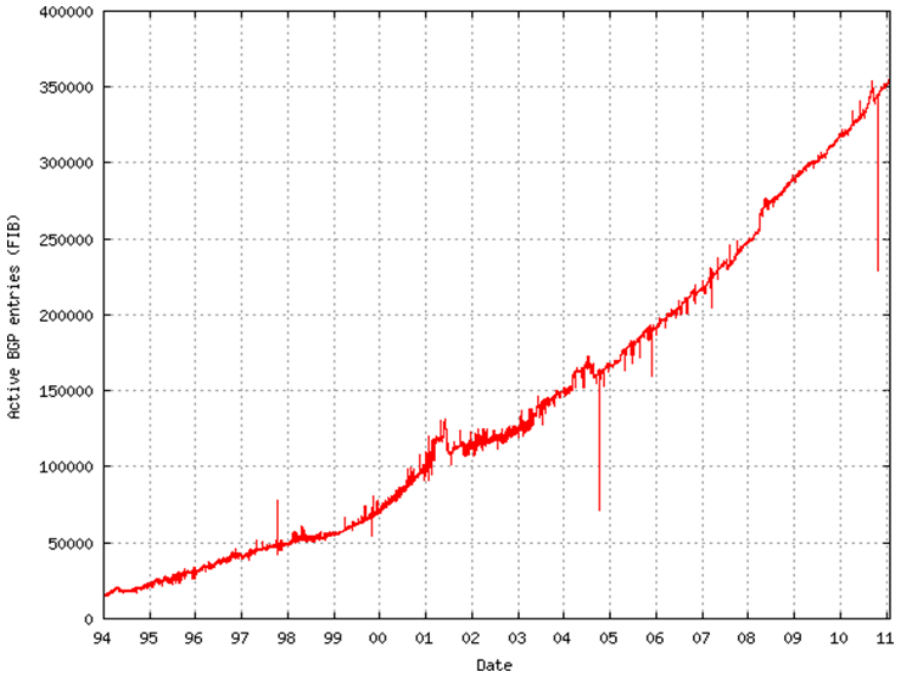


Figure 4.4: Growth of the BGP Tables at DFZ Routers.

4.3.1 ILSAs Overview

Two high level research challenges, the ID/LOC generation and the Mapping System (the entity in charge of the ID/LOC mapping and vice-avers) must be faced when designing an ILSA architecture. As for the first challenge, nowadays, several alternatives may already be found in the recent literature differently handling the ID/LOC generation depending on the network segment they operate at. Thus, as shown in Fig. 4.5, regarding the ID/LOC generation challenge a preliminary ILSA classification turns into two sets of ILSAs schemes, namely Network based and Host based schemes.

Network based schemes: Operating at the network level, usually on the border routers at the network backbone; hence, no modifications are required on the end-nodes (host level). One of the most relevant network based ILSA schemes is LISP [142].

Network based ILSA schemes can be further categorized into: 1) Map-Encap schemes, and 2) Address Rewriting schemes. In Map-Encap schemes (such as LISP), a network packet destined to a certain object (packet with an ID as a destination), is encapsulated into a new packet, whose destination will be a locator. This strategy is widely used in many networking aspects and is usually referred to as tunneling in network jargon. Unlike this tunneling approach, Address Rewriting architectures operate similarly to NAT (Network Address Translation), replacing a packet ID by a locator.

Host based ILSA schemes: Operating at the host level, specifically at the end-nodes, no modifications are required at the network level. A Host based ILSA scheme is a more appealing

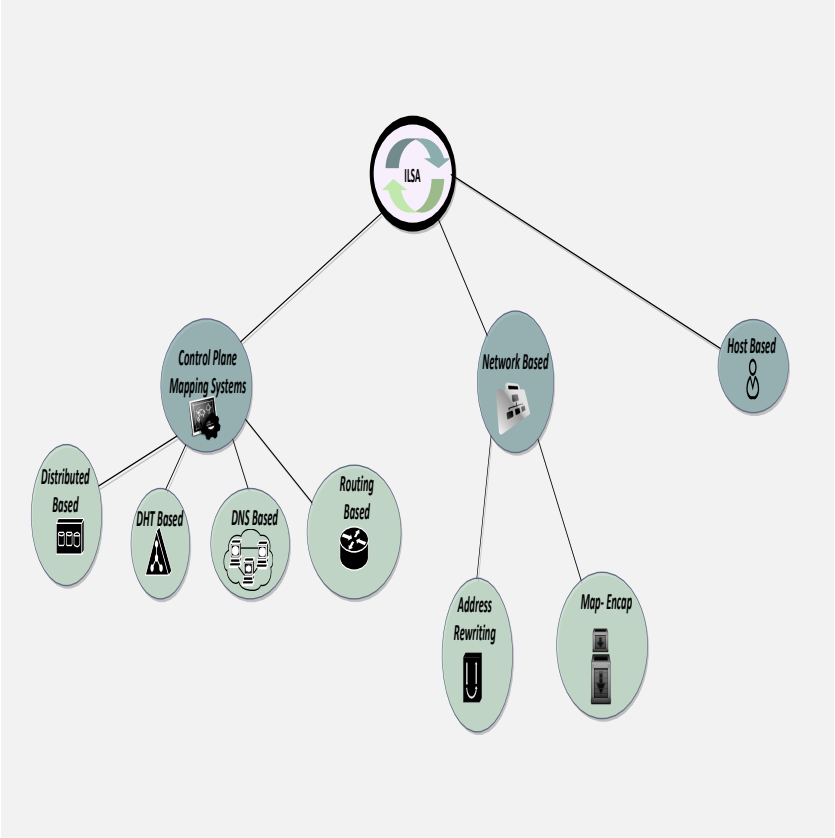


Figure 4.5: Taxonomy of ILSAs.

solution than a Network base scheme for network operators since cost investment is not demanded on the network. However, this solution drives software providers to update their products to meet specific requirements of a Host based ILSA scenario, what of course does not sound that attractive for them. One of the most relevant Host based ILSA schemes is HIP [147].

There is a conceptual difference between both approaches that deserves to be mentioned. Unlike a Network based ILSA scheme, where the ID/LOC space (an ID/LOC space is a collection of all valid ID/LOCs) is fixed, Host based ILSA deployments are not restricted to use a unique LOC or ID space. This feature could be helpful in some scenarios, for example, a two locators space scenario, may assign one locator space for global routing and the other one for local routing, or a two IDs space scenario, may assign one for the identification of virtual objects (e.g., network groups), and the other one for the identification of physical objects (e.g., computers or mobiles nodes). This characteristic increases the addressing granularity.

The second high level challenge refers to the bidirectional mapping between an ID and a Locator ($ID \Leftrightarrow LOC$). Notice that a different level of mapping is also needed in Host based ILSA schemes between ID spaces (ID_{s1}, ID_{s2}), i.e., an ID could be mapped to another ID which may belong to the same or to a different ID space.

While initial ILSA proposals, such as LISP and Six/One, handled the ID/LOC mapping process over the data plane (Data plane architectures), the current trend on ILSAs design is pushing for Control plane architectures. In these architectures the mapping process is run by a Mapping System completely decoupled from the data plane. The Mapping System is a crucial component on any ILSAs scheme, since it is responsible for the mapping between IDs and locators. Mapping systems are conceptually supported by different technological approaches: Domain Name System (DNS), Distributed Hast Table (DHT), Distributed Mapping Systems and Routing protocols generating different Mapping System types. It can be stated that the chosen type of Mapping System is an important design decision because this one will adopt most weaknesses and flaws of its parent technology. For example, a Mapping System based on a routing technology such as BGP will inherit most of the yet unsolved problems of this routing protocol.

4.3.2 LISP Operation

As mentioned in the last sub-section LISP is a Network based ILSA scheme that use Map-Encap processes, e.g., IP-over-IP tunnels deployed between border routers located at different domains. To this end, the IP addresses allocated to the external interfaces of the border routers act as Routing Locator (RLOC) addresses for the end systems in the local domain. Since an AS usually groups several border routers, the local Endpoint Identifier (EID) addresses can be reached through multiple RLOC addresses. Hence, LISP separates the overall address space into two parts, where only addresses from the RLOC address space are assigned to the transit Internet. Therefore, only RLOC addresses are routable through the Internet, that is, EID addresses are considered routable only within their local domain.

The basic idea is that an EID represents an end-host IP address, while a RLOC represent the IP addresses where end hosts are located. At border routers EID are mapped into RLOC, according to a Map-and-Encap scheme. The scaling benefits arise when EID addresses are not routable through the Internet — only RLOC addresses are globally routable. This allows efficient aggregation of the RLOC address space.

Moreover, recent studies show that LISP offers some key advantages. For instance, authors in [148] show that the size of the global routing table can be reduced by roughly two orders of magnitude with LISP.

To illustrate the basic operation of LISP see Fig. 4.6. Notice that when the host *S* with EID 190.1.1.1 wants to communicate with host *D* with EID 200.1.1.2 in a different domain, the following sequence of events occur in LISP: 1) The usual lookup of the destination address EID in the DNS is performed; 2) Once the EID is obtained, data is forwarded to at least one of the local border routers referred to as Ingress Tunnel Routers (ITRs); 3) since only RLOC addresses are globally routable, when an ITR receives packets destined to a host outside its domain, it queries a mapping system to retrieve the EID-to-RLOC mapping; 4) After the EID-to-RLOC mapping resolution, the ITR encapsulates and tunnels packets between the local RLOC address (ITR address 3.3.3.2 in the Fig. 4.6) and the RLOC address retrieved from the mapping system, which is a Egress Tunnel Router (ETR) address in LISP terminology (either 4.4.4.2 or 10.0.0.2); 5) At the destination domain, the ETR decapsulates the packets received through the tunnel and forwards them to their final destination (host *D*). It is worth mentioning that from the first packet received, a ETR is able to cache a new entry, solving in this way the reverse mapping for the packets to be tunneled back from destination to source.

The adoption of an ILSA scheme such as LISP provides several benefits such as 1) reduction of the routing tables size; 2) cost-effective multihoming; 3) easy address renumbering; 4) TE capabilities; 5) Full Mobility, among others. Nevertheless, there several issues related to LISP resilience capabilities that must be addressed in order to consider a future deployment of LISP in current CGNs.

4.3.3 Making the way to a Fault Tolerance LISP

One of the main weaknesses of LISP is related to resilience. In order to increase the resilience degree of LISP, authors in [149] propose control plane based on the novel concept of retrieving EID-to-RLOC mappings within the DNS Resolution time. To this end, the mappings between EIDs and RLOCs are replicated in all of the edge routers within the same AS. Despite the fact that this approach ensures improved reachability, it may lead to scalability problems since each border router must store mapping information that rarely needs to be used, increasing in this way the mapping table size. In order to minimize the amount mapping information managed by a border router, while ensuring the highest possible reachability, the author of this thesis contribute to the design of the so-called LISP Redundancy Protocol (LRP). LRP is inspired by the Cisco's Hot Standby Routing Protocol (HSRP) [150]. HSRP permits to configure

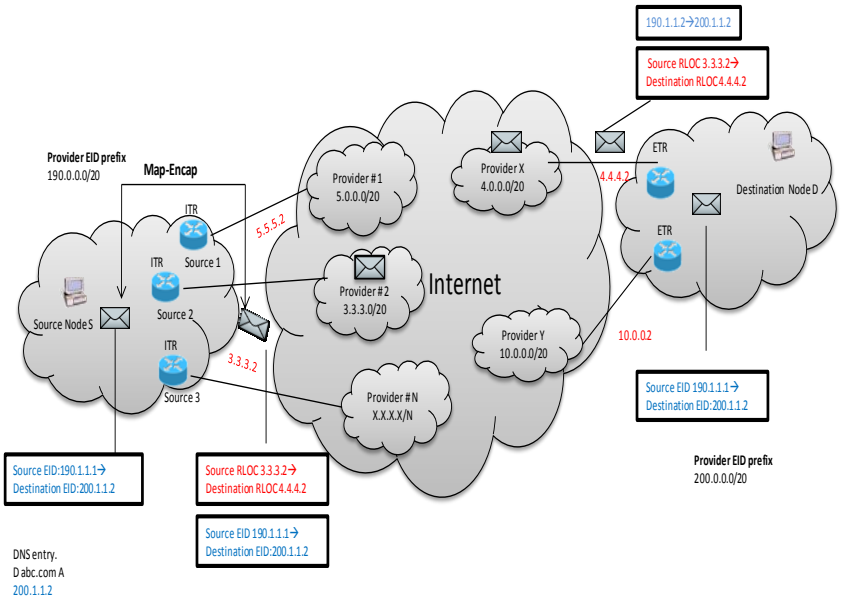


Figure 4.6: Operation of LISP.

two routers one as a main border router and the other as a backup –Master-Slave model. LRP extends this functionality by creating different logical groups. By means of logical groups, border routers are not restricted solely to Slave or Master modes (see Fig. 4.7). This leads to 1) all border routers can forward data –since only primary routers do so, and; 2) there is not need to replicate the entire mapping information on all the border routers, it can be obtained on the fly.

In summary, the main features offered by LRP are:

- The xTRs (ETR or ITRs) can be clustered into different LRP groups or pairs.
- The Mappings are pushed onto the LRP groups or pairs.
- All the xTRs in the group can carry traffic (active rather than standby).
- No need for data-probes (message used to obtain mapping entries) when the xTR does not have a mapping.

Handling Inter-domain link failures with LRP

In the following we will discuss and describe the actions that are executed in order to address a failure affecting an inter-domain link. In step 1 of Fig. 4.8, a high volume of traffic is sent to a border router (ITR1), which is responsible for encapsulating the traffic and send it through its international links to the destination. When a inter-domain link fails (step 2), ITR1

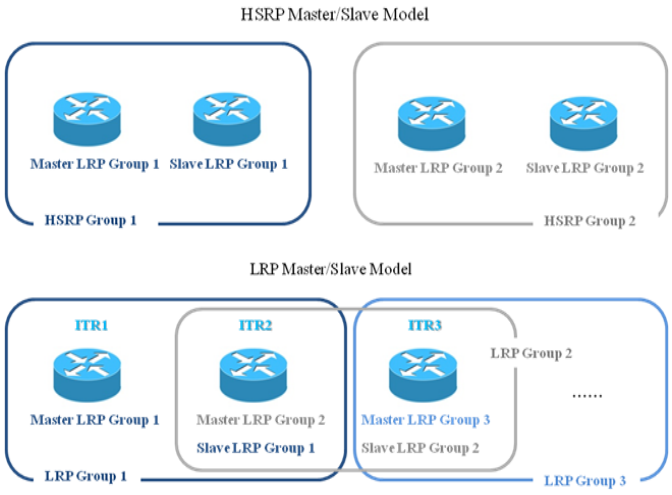


Figure 4.7: Master/Slave Model HSRP vs. LRP.

automatically detects this event and (in real time) forwards all the incoming traffic to the other ITR (ITR2) belonging to its LRP group (step 3). ITR2 has the correspondent mapping since it shares a virtual group with ITR1 and now is in charge of encapsulating and sending this traffic to its destination (step 4). On the other hand, by means of the internal routing protocol (running in the AS), the failure of the inter-domain link is notified to update the routing tables in order to reroute the traffic (step 5). Meanwhile, the LISP Control Box (LCB) –the entity in charge of the configuration of xTRs related to LISP capabilities – would reconfiguring the mapping of the different ITRs with the aim of load balancing the outbound traffic (step 6). Finally, the restored traffic is rerouted according to the internal routing policies (step 7).

Handling ITR failures with LRP

The following lines describes how LRP is capable to prevent traffic loss while minimizing the amount of signaling overhead required to so. As shown in Fig 4.9 (step 1), a large volume of traffic is sent to border router ITR1, which is responsible for encapsulating this traffic and send it to its final destination. In case of a failure affecting ITR1 (step 2), the resilience mechanism (HSRP) deployed in the network automatically selects ITR2 for the role of Master. This is done i a short-term basis by HSRP, nearly 3 seconds, which it is much faster than the reaction of any routing protocol against a failure event. Moments later, the routing protocol deployed in the network notifies the failure of ITR1 (step 4), updating in this way the routing tables and allowing the traffic to be rerouted. On the other hand, the LCB would be responsible for reconfiguring the mapping of the different ITR to balance the outbound traffic load (step 5). Finally, the traffic is rerouted according to the routing protocol policies (step 6).

Based on the illustrative examples described in this section related to failures in a LISP scenario, it can be stated that LRP prevents packet loss and in particular the sending of data-probes, i.e., reduces the signaling overhead under the presence of links and nodes failures.

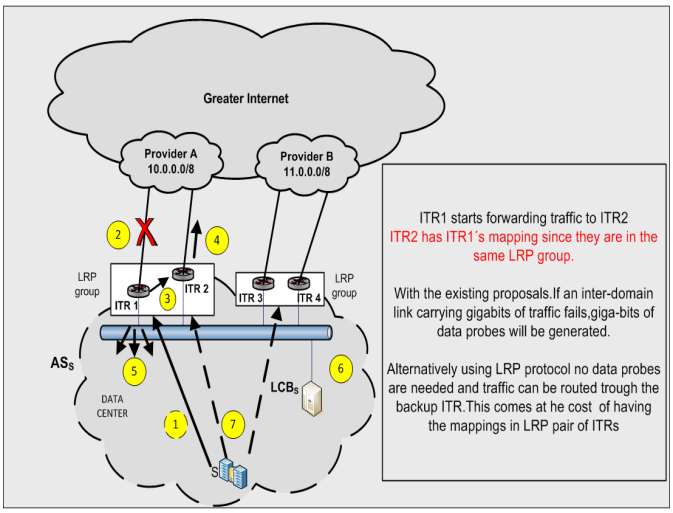


Figure 4.8: Dealing with Inter-domain link failures by means of LRP.

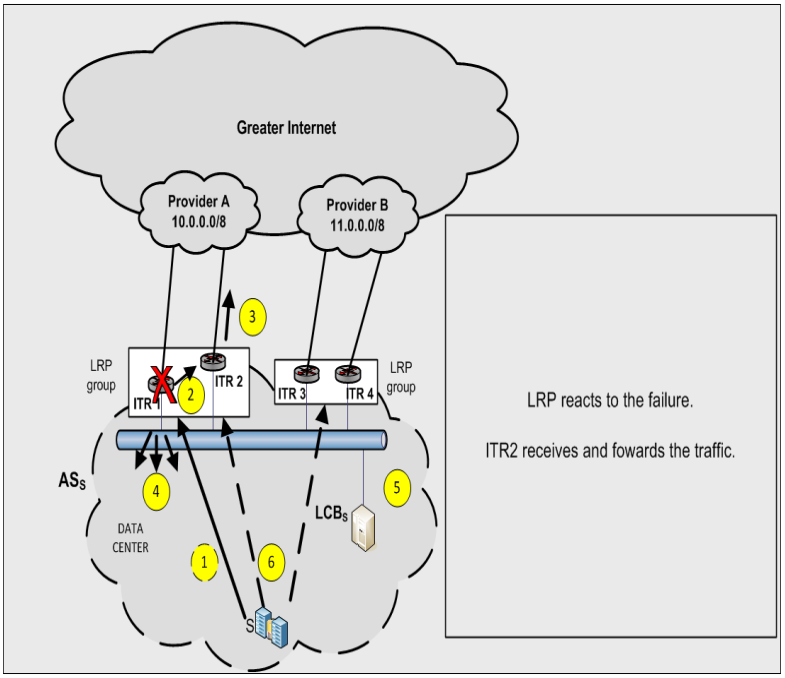


Figure 4.9: Dealing with ITR failures by means of LRP.

4.4 Context-Aware PCE

In summary, the network is facing a new highly demanding scenario, where technology must evolve fast enough to support the ever growing users' demands. Indeed, the pendulum has swung from location/host to information/context-oriented communication models. In this thesis, we push for positioning the PCE concept into a context-aware communication model. To this end, we introduce the novel concept of Context-Aware PCE, based on the synergy between conventional PCE architectures and ILSA schemes. Contrary to a conventional PCE where the endpoints of a Path Computation Request (PCReq) are location dependent, i.e., host-oriented PCE, in a context-aware PCE scenario, the endpoints are IDs (identifiers). The IDs are finally mapped to locators (LOCs) by using an ILSA scheme according to a given "context".

Context-aware networking can be adopted in PCE scenarios by the following: 1) enabling the interaction between PCEs and ILSA schemes, 2) Coupling the endpoints of a PCReq to IDs, and; 3) Perform the mapping between an ID and a LOC according to a given context, e.g., location, time, traffic volume, etc. On this basis, the amount of network resources (optical wavelength, cross-layer connections) allocated to a connection might change during its holding time according to the context specified by a Path Computation Client (PCC)– a PCC is the entity requesting a path computation. Indeed, the paradigm of a context-aware PCE stems in the fact that PCCs are solely interested in setting up a connection to a certain endpoint or content no matter its location. The relevant is the *"What"* ("connection to something") rather than the *"Where"* ("connection to a location").

Moreover, bear in mind that coupling the endpoints of a connection to IDs take off from the PCCs several tasks such as keeping state of nodes location, validation of node status (availability), and opens the possibility to the re-optimization of connections in a proactive manner –referred to as Active-PCE by authors in [151] .

In the following lines, we illustrate how the synergy between PCE and ILSA schemes can enhance the current location/host oriented communication model in order to augment distinct network features such as resilience, or traffic-oriented routing.

Figure 4.10 shows a possible use-case of collaboration between a context-aware PCE and an ILSA scheme in order to augment the resilience level in a multi-domain optical scenario. As shown in Fig. 4.10a, a PCC within domain 1 sends a PCReq for a LSP with LOC *A* and ID *Building-Domain 2* as endpoints (step 1). A context-aware PCE communicates with an ILSA scheme to obtain the set of LOCs corresponding to the ID *Building-Domain2*, which are *B* and *C* (step 2). The PCE computes two LSPs (one for each LOC). Then the context-aware PCE sends the Path Computation Reply (PCRep) to the PCC (step 3). Based on a policy defined by the administrator of domain 1, the LSP for Wavelength Router (WR) *B* is configured as a primary connection, whereas the path to WR *C* is configured as a backup connection (step 4).

In case of a failure affecting WR *B*, see Fig. 4.10b, WR *A* can switch the traffic to WR *C*.

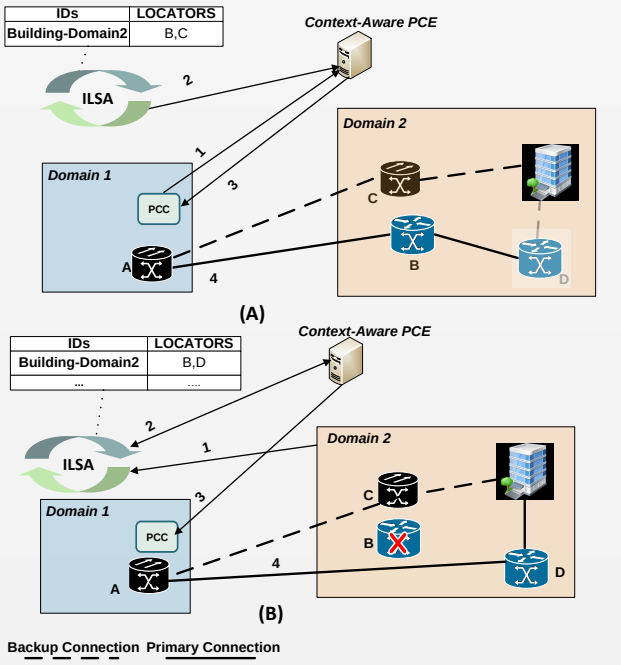


Figure 4.10: A Context-Aware PCE for augmenting the network resilience level: a) LSP computation; b) LSP re-optimization.

As a consequence of this failure, a pre-configured recovery action triggers the provisioning of WR *D* as the backup for WR *C*. Domain 2 sends update information to the ILSA scheme specifying that ID *Building-Domain 2* is coupled to LOCs *C* and *D* (step 1). Then, the ILSA scheme sends this update to the PCE (step 2). The PCE re-optimizes LSP *(A, Building-Domain 2)* by computing a path between WRs *A* and *D* and send a Path Computation Update message (PCRUpd) to the PCC (step 3). Finally, the PCC triggers the provisioning of LSP (*A*, *D*), and the affected traffic is once again is routed along the primary path recently provisioned (step 4).

In order to support the scenario shown in Fig. 4.10 in the conventional host-oriented model, it is required a control plane capable of distributing location information among PCCs within different domains. Foremost, the deployment of a multi-domain control plane can be arduous due to scalability, confidentially and technical issues [152]. As such, ILSA schemes should be considered as scalable solution to disseminate control information such as location data.

Moreover, Fig. 4.11 shows another use-case of collaboration between a context-aware PCE and ILSA scheme, where an Open-Data Skateholder (Domain 1) collects information from Data Repositories (Domain 2, and 3) by means of an Open-Data Middleware [153]. The rationale of this scenario is to illustrate how network operators can optimize the allocation of resources in their networks by: 1) interacting with both ILSA and PCE schemes; and 2) user behavior characterization, i.e., analyze the trending topic or traffic behavior.

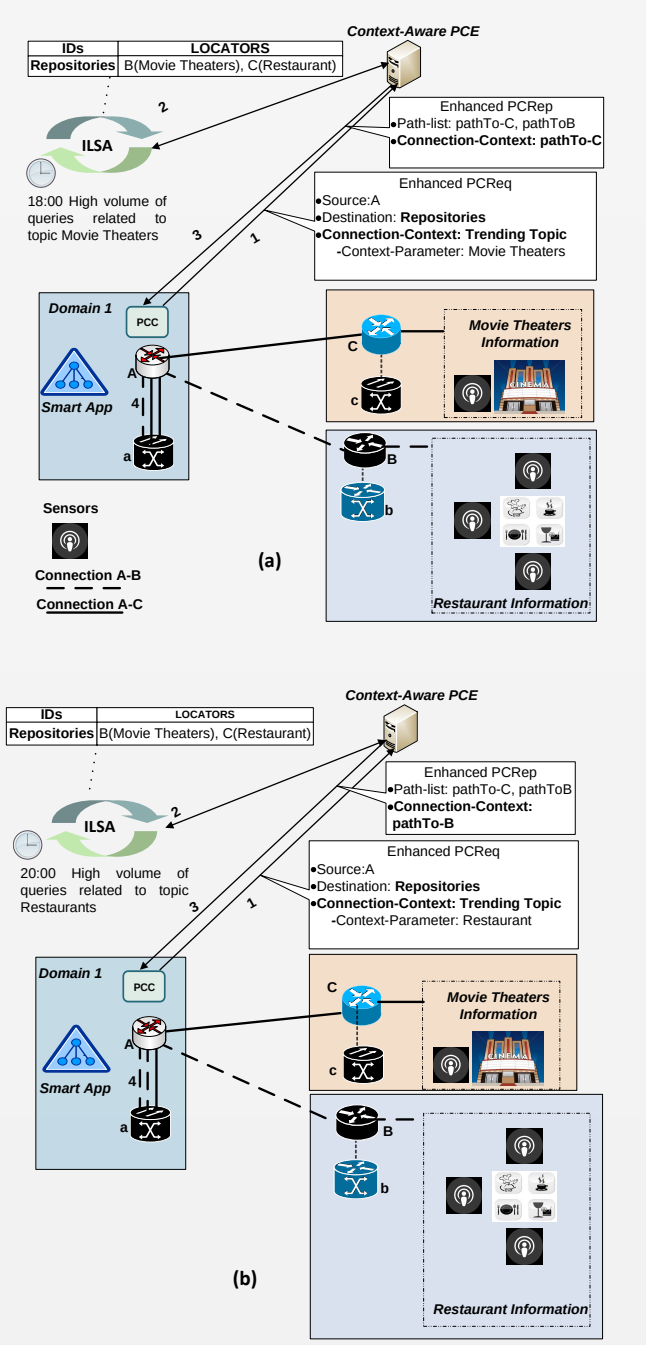


Figure 4.11: A Context-Aware PCE for defining context-aware connections: a) Trending-Topic: Movie Theaters; b) Trending-Topic: Restaurants.

As shown in Fig. 4.11a, a PCC sends a PCReq for an LSP with LOC A and ID *Repositories* as endpoints (step 1). Notice that the conventional PCReq and PCRep are enhanced with a new parameter so-called Connection-Context. The purpose of the Connection-Context parameter is to define the main drivers of a connection. For the scenario shown in Fig.4.11,

a PCC is interested on establishing a connection to all devices mapped by ID *Repositories* with the constraint that more network resources must be allocated to the connection whose destination is the device corresponding to the trending-topic *"Movie Theaters"*.

By means of interaction with an ILSA scheme, ID *Repositories* is mapped to LOCs B and C (step 2). Then a Context-Aware sends a PCRep to the PCC specifying two paths ($pathToB$, $pathToC$) corresponding to the given Connection-Context (step 3). Finally, the PCC triggers the provision of a connection between LOCs $A-B$ and $A-C$, but more network resources–cross-layer connections or transponders– are allocated to connection $A-C$ (step 4).

As a final example of interaction between PCE and ILSA scheme we consider the scenario shown in Fig, 4.12a, which depicts a green-networking scenario, where a certain domain (Domain 1) wants to set up a connection with a Geographic Information System (GIS) provider using the sun as its main energy source (solar energy). To this end, a PCC within Domain 1 sends a PCReq with ID $GIS-Provider$ as destination and "Energy Source" as the Connection-Context (step 1). Hereinafter, an ILSA scheme maps ID $GIS-Provide$r to a LOC belonging to a router within domain $GIS-Provider1$ (step 2). Finally, a connection is established between $routerA$ and $routerC$.

The $GIS-Provider1$ switches to conventional energy (oil energy), and $GIS-provider2$ switches to solar energy, see Fig. 4.12b. This is informed to a context-aware PCE (step 1), this one re-optimizes path $A-GIS-Provider$ (step 2). Then, a PCC tear down connection $A-C$, and triggers the setup of connection $A-B$ (step 3).

It can be assumed that based on time information, an ILSA scheme can recognize which domain is using solar energy, hence no signaling is required. However, NSI such as energy source can be sent/requested for each domain in order to update the mapping entries (information concerning IDs and LOCs) stored by an ILSA scheme.

The scenarios depicted in Fig. 4.10, 4.11 and 4.12, drive us to encourage the practice of context-aware connections, where the establishment and the network resources allocated to a connection are done according to a given context-information. Furthermore, we consider that for suitable handling the IoT requirements neither PCCs nor PCEs should manage any location information concerning a connection destination. As it is stated in the literature, for the purpose of managing location information ILSA schemes are second to none.

On the other hand, in the following lines, we describe how the interaction between a context-aware PCE and ILSA scheme can address several of the requirements of an IoT model. To this end, we introduce the concept of a Context-Aware Graph, which can enhance the capabilities of a context-aware PCE.

Driven by the advent of new technologies related to transport capacity and both diversity and capabilities of end-devices, there is an increasing deployment of smart devices (sen-

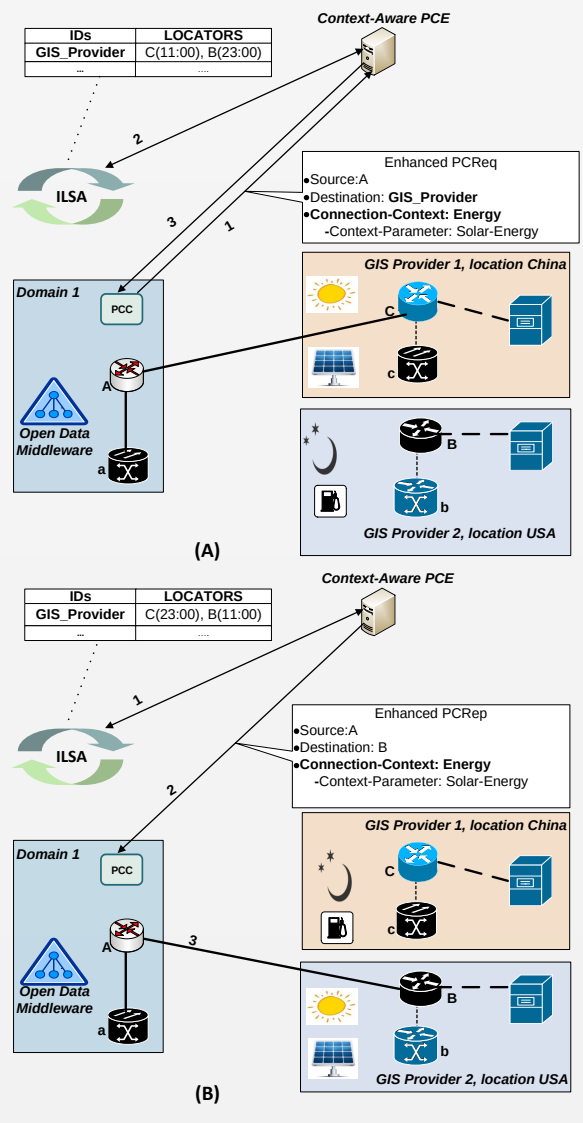


Figure 4.12: A Context-Aware PCE in an green-networking scenario.

sors, etc), and open data gathering servers, coining the so-called Smart Cities –a concept encompassed by the IoT. A Smart City infrastructure embodies a massive deployment of smart-services such as energy efficiency systems, urban transportation optimization, and interactive information systems, among others. Consequently, city councils and local governments are starting to develop new internet applications that exploit Smart City services.

Upon the composition of a smart-service, i.e., a service that collects information from several data sources, physical connectivity to the data sources must be established. Traditionally, both service composition and path computation have been two processes independently from each other. Nonetheless, we believe that this myopic practice limits the scalability and performance of Apps devoted to service composition. This issue motivated us to intro-

duce the concept of an overlay graph for modeling context-aware connections, termed as Context-Aware Graph.

Our intention is to combine Apps devoted to service composition and a context-aware PCE scheme into a collaborative ecosystem, where a context-aware PCE computes a path based on both the context-aware Graph and Transport Network Graph. The context-aware graph is used for service composition, i.e., forward a request across different data sources in order to compose a smart-service. Once a smart-service is composed, the transport network graph is used to compute a path in order to establish physical connectivity to the data repositories required by the smart-service.

To illustrate the features of the context-aware graph, imagine that a network carrier wants to offer to its clients a smart-service through they can: select a movie to watch, pick the movie theater of their preference, along with both street map and transportation information related to the selected movie-theater. This scenario is depicted in Fig. 4.13. As shown in Fig. 4.13, an App –as the role of a PCC– sends a PCReq with ID "Entertainment" as destination (step 1). Then, in order to compose the requested smart service, the context-aware PCE computes two paths to ID "Entertainment", both having the same cost on the context-aware graph: (Movies-Theaters DB, GIS Provider 1, Taxis-DB), (Movies-Theaters DB, GIS Provider 2, Taxis-DB), step 2. In this example, the number of hops along a path is used as the cost metric.

A tie breaker process is required to select only one of the computed paths. As a consequence, the context-aware PCE requests to an ILSA the LOCs of each node belonging to the computed paths (step3). Then, the context-aware PCE computes the shortest-path based on the transport network graph, and sends the selected path (Movies-Theaters DB, GIS Provider 1, Taxis-DB) to the App (step 4).

With the scenario described in Fig. 4.13 we intend to position the context-aware PCE as the role of not only computation of physical paths, but also virtual paths representing a service composition. It is worth mentioning that the context-aware graph introduced in this section is modeled as $G(V,E)$, where V is the of nodes representing an Open Data provider, and E is the set of edges representing that data from a source vertex is required by the terminal vertex of an edge in order to provide certain information (service). However, the context-aware graph can represent other types of graphs such as a Web-scale workflow system [154].

4.5 Validation of the Context-Aware PCE

In this section, we provide numerical results related to the time required to provision a path (T_{pp}) by the conventional PCE and the proposed context-aware PCE in an IoT scenario. In addition, in order to ensure realistic findings, we present real data that provides some light on both periodic and highly dynamic behavior regarding trending-topics, which is one of the use-cases presented in the previous section.

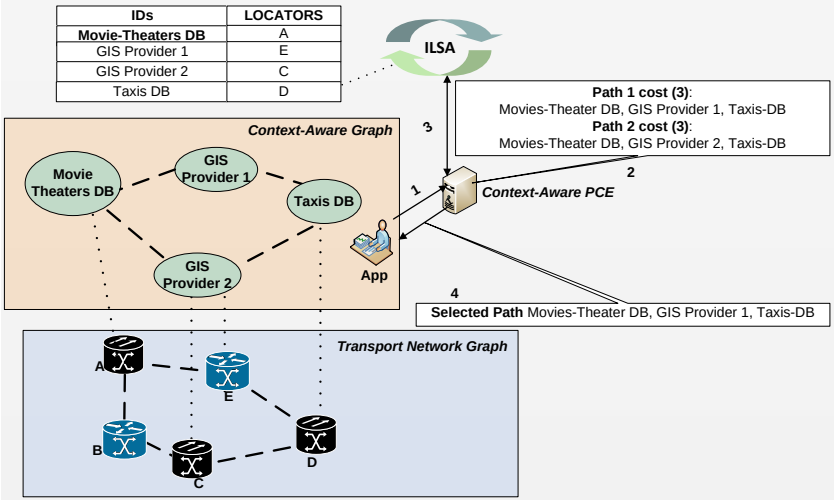


Figure 4.13: A Context-Aware Graph.

We compare both the proposed context-aware and the conventional PCE schemes in order to evaluate T_{pp} in an IoT scenario. To this end, T_{pp} is obtained as shown in Equation (4.1), where T_{mp} is the time required to obtain the mapping between a ID-LOC pair, commonly known as the mapping phase, and T_{sc} is the time to set up a circuit.

$$T_{pp} = T_{mp} + T_{sc} \tag{4.1}$$

T_{sc} is computed in the same manner for both context-aware and conventional PCE schemes as shown in Equation 4.2), where H is the average path length, $H = 4$ hops; d is the average delay between two nodes under uniform traffic, $d = 50ms$; C is the time to set up a cross-connect, $C = 500us$; finally, P is the message processing time, $P = 10\mu s$.– values extracted from [155]. Under these assumptions $T_{sc} = 351ms$.

$$T_{sc} = (2H - 1) \times d + C + 2 \times H \times P \tag{4.2}$$

The time required by the mapping phase for the context-aware PCE scenario (T^1_{mp}) is computed as shown in Equation (4.3), where $T_{chord-nodes}$ is the number of nodes (Chord Nodes) forming the ILSA scheme, and d^1 is average delay between two nodes in ILSA scheme, $d^1 = 10ms$. We assume that the mapping phase engine of the evaluated ILSA scheme is based on the Chord algorithm, similar to authors in [156]. The rationale driving this assumption boils down to the high performance achieved by ILSA schemes based on DHT mapping systems related to lookup times [157],[156]. This makes them reliable for the mapping system of an ILSA scheme. It is worth mentioning that there are DNS-based mapping systems available in

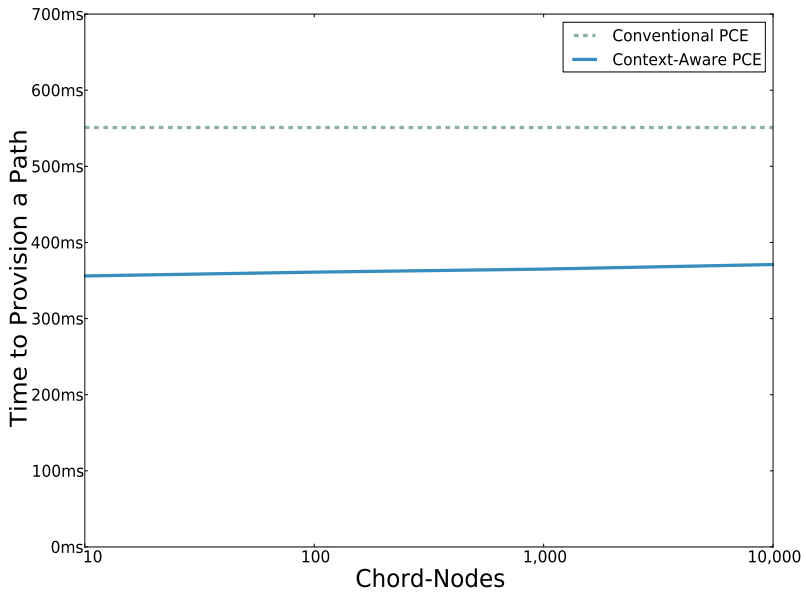


Figure 4.14: Time required to provision a path for context-aware and the conventional PCE schemes in an IoT scenario.

the literature [158]. Nevertheless, they are positioned in intra-domain scenarios where the amount of possible identifiable NEs is far less than the expected in an IoT scenario.

$$T^1_{mp} = 0.5 \times log \times T_{chord-nodes} \times d^1 \tag{4.3}$$

To obtain the time required by the mapping phase for the conventional PCE scenario (T^2_{mp}), we assume that the mapping phase related to ID-LOCs relies on the current DNS architecture for its operation. According to authors in [159], the mapping phase for DNS systems varies from 100 ms up to 1 second according to parameters such as Time to Live (TTLs) and cache sharing. Thereby, we assume that on average $T^2_{mp} = 200ms$. Fig.4.14, shows the T_{pp} for both context-aware and conventional PCE schemes. As it can be observed, the T_{pp} for the context-aware PCE is significantly less even for large network scenarios (high number of control/Chord nodes). In addition notice, that the performance of the conventional PCE could be worse, if it is assumed than the DNS-mapping system needs to perform the conventional DNS lookup in addition the ID-LOC mapping.

The obtained results were expected since several studies already available in network research claim that in an IoT scenario, the mapping phase should be decoupled from the network elements, and placed in a dedicated element for its execution, namely an ILSA scheme.

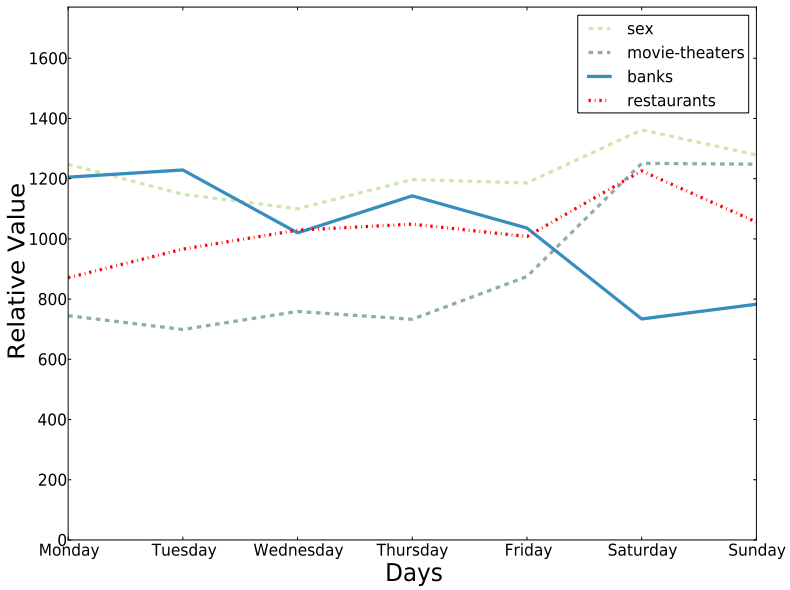


Figure 4.15: Daily queries distribution for search topics restaurants, banks, erotic-content and movie theaters.

Moreover, with the aim of demonstrating that it is not unrealistic to assume a periodic and highly-dynamic behavior regarding trending-topics, we consider the numerical results shown in Fig. 4.15, which depict the daily queries distribution concerning the following search topics: restaurants, banks, erotic-content, and movie theaters, between June and September in Barcelona, Spain.

The numerical results shown in Fig. 4.15 are relative values extracted from [160]. Notice that certain search topics such as movie theaters exhibit a drastic increase during weekends, nearly 41%; the opposite occurs with search topics such as banks, a reduction up to 48%. The obtained results provide some light on a periodic and dynamic behavior that regularly follow search topics.

It can be stated that the results presented in this section fuel the development of new applications which jointly with both a context-aware PCE and ILSA schemes can support the requirements of an IoT scenario.

5 Conclusions and Future Work

This thesis focuses on the study of routing and protection algorithms with the aim of providing solutions that can enhance the resilience of Carrier-Grade Networks. To this end, two technical objectives were defined: 1) Proposing and validating innovative Network Coding Protection (NCP)-based strategies to enhance network resilience considering both planning and dynamic scenarios; and 2) Evaluating the benefits of new network paradigms such as Path Computation Element (PCE) and context-aware communications to enhance the performance of the proposed routing and protection algorithms.

This thesis considers NCP as a technical solution contributing to improve overall network protection because of the low network resources consumption and the fast recovery times as obtained by this type of protection scheme. With regard to the study of NCP strategies, this thesis evaluates NCP schemes in both single and multi-layer scenarios considering planning as well as dynamic scenarios. For single-layer scenarios it was presented a techno-economic study for assessing the protection cost obtained when employing three proactive protection strategies, namely Dedicated Protection (DP), network coding with a DP scheme (DPNC), and multiple-coding with a DP scheme (DPNC*). It was assumed that the evaluated protection schemes were deployed either at the IP/MPLS or at the Optical layer of a multi-layer network.

Based on the obtained results, we conclude that the use of DPNC, specifically the multiple-coding feature (DPNC*), can significantly reduce both CAPEX and OPEX, independently of the network layer where they are deployed, in comparison with conventional protection proactive schemes and despite of the cost associated to enable NC capabilities. An average of 60.5% of CAPEX reduction can be achieved independently of the network layer technology. Indeed, 49% and 72% of CAPEX reduction is obtained when deploying DPNC* at the Optical and the IP/MPLS layers respectively. In addition, a 52% of OPEX reduction is obtained at the IP layer.

Moreover, we evaluate the benefits of multiple-coding in flexible-grid scenarios. Flexible-grid provides a substantial reduction on both network resources and power consumption. These advantages cannot be neglected in new scenarios such as Data Center Networks (DCNs) where the available transport capacity as well as other issues such as control signaling and

power consumption significantly affects the scalability of a protection scheme.

We conclude that by means of multiple-coding combined with the benefits provided by a flexible-grid, the network resources required for protection tasks can be substantially reduced. Indeed, multiple-coding with flexible-grid outperforms conventional proactive protection schemes as well as multiple-coding deployed in a conventional fixed-grid scenario.

On the other hand, for multi-layer scenarios, we propose a novel NCP scheme, referred to as DPNC+. The novelty of DPNC+ is that it leverages Network Coding (NC) techniques, backup path provisioning, and cross-layer Network State Information (NSI), in order to reduce the network resources allocated to link protection. Simulation results obtained using real network topologies show that the proposed scheme provides a significant reduction (about 50%) of both IP/MPLS and Optical bandwidth required for link protection, in comparison with other proactive protection schemes. We believe that network operators should consider NCP schemes combined with cross-layer information as an appealing solution to design efficient proactive protection schemes.

Motivated by the good performance of NCP schemes in planning scenarios, this thesis also evaluates NCP schemes in dynamic scenarios. For this purpose, we present a novel proactive protection scheme, referred to as Predictive Network Coding Protection (PNCP). PNCP is a source Routing and Wavelength Assignment (RWA) Algorithm devised to mitigate the negative effects of inaccurate NSI on the blocking probability in dynamic protected scenarios. In dynamic scenarios there are several sources leading to have inaccurate NSI such as 1) non-neglected delay propagation; 2) updating policies; 3) high level of aggregation imposed by a hierarchical network design; 4) frequent CRs arrivals, and; 5) control messages failures. This thesis focuses on the study of inaccurate NSI caused by updating policies in both single and multi-fiber networks since this source of inaccuracy is dominant compared with the other potential sources. To the best of our knowledge, this thesis is the first work related to the study of NCP schemes in dynamic scenarios considering inaccurate NSI.

It is worth mentioning that we first evaluated Prediction techniques –considering fine and coarse granularity predictive counters– in unprotected scenarios, in comparison with other RWA algorithms. The evaluated Prediction techniques consist in using predictive counters in order to model a route availability.

Indeed, the main mechanism of PNCP is a Predictive RWA algorithm based on the use of fine-granularity predictive counters for the purpose of predicting route availability contrary to other proposals adopting coarse-granularity counters. Based on the performance yielded by PNCP, we conclude that indeed, fine-granularity predictive counters can substantially mitigate the negative effects of inaccurate NSI while avoiding periodic dissemination of NSI. In addition, we also conclude that PNCP significantly improves the performance obtained by conventional protection schemes in network scenarios with routing inaccuracy, as well as it yields a lower utilization of the network resources dedicated for protection.

Finally, this thesis focuses on the analysis and evaluation of new trends for routing and resilience. The rationale driven this objective, is the advent of new network architectures such as the PCE. PCE-based schemes can substantially overcome the weaknesses related to path computation of source RWA strategies in the current location/host-oriented communication model.

However, there is a wide consensus in the network research community pushing for the demise of location/host-oriented communication models in the coming years, specifically for new network paradigms, such as the so-called Internet of Things (IoT). Indeed, an IoT scenario is raising new challenges requiring novel research efforts that turned into new network architectures, such as ID/LOC Split Architectures (ILSAs) or information-centric or context-aware communication models. Therefore, this thesis also benefits from the knowledge on ILSA schemes, and introduces a novel PCE-like architecture, so-called context-aware PCE, that enhances conventional PCE schemes to fully exploit the advantages provided by new communication models.

The building block of a Context-Aware PCE is the synergy with ILSA schemes. ILSA schemes are a widespread strategy to address the problem related to the double functionality of addresses, while avoiding Clean-Slate approaches. This problem is well-known in network research because it hinders the deployment of new application and services as well as the scalability of the current addressing scheme.

Contrary to the conventional PCE, in a context-aware PCE, the endpoints of a path computation request are coupled to an Identifier (ID). The mapping of an ID to a IP based locator is done according to a given context, traffic volume, trending topic. We believe that in order to address the requirements–mobility, traffic engineering, green-networking, smart internet applications– of future internet architectures such as the IoT, interaction between PCE and ILSA schemes would be eventually required.

The following areas of work might be set from this thesis. The first, to evaluate distinct coding strategies that can potentially improve the performance of NCP schemes. In particular, we consider that four main issues related to the performance of NCP schemes which are: 1) the amount of network resources required for protection; 2) dealing with Shared Risk Link Groups (SRLG), i.e., optical link failures affecting more than one virtual (IP/MPLS) link; 3) NCP schemes that can operate jointly with restoration schemes; and 4) the deployment of NCP schemes on multi-domain scenarios. We consider that the third issue is highly important within multi-layer large scenarios where there can be several potential failure sources. As a matter of fact, we consider that protection strategies such as NCP schemes cannot be the only resilience strategy deployed in a network, i.e., protection and restoration schemes deployed at all network layers (IP /MPLS and Optical) are required.

Another future line of work inferred from this thesis is the study related to the performance of protection algorithms under inaccurate NSI, considering both flexible-spectrum grid and dynamic scenarios. In recent years, the pendulum has swung from the study of RWA algorithms

to Routing Spectrum Assignment Algorithms (RSA) motivated by benefits of a flexible-grid. RSA algorithms must consider additional constraints such as the Spectrum Contiguous Constraint in order to avoid the negative effects caused by spectrum fragmentation. To the best of our knowledge, there is not any study already available dealing with the inaccurate NSI in dynamic flexible-grid scenarios.

Finally, we consider that the novel concept of a context-aware PCE presented in this thesis can be further elaborated to generate consensus and wide visibility on the community while also providing new evaluation results supporting this concept.

Bibliography

[1] Cisco Systems Inc. Converge IP and DWDM Layers in the Core Network, 2007. White Paper.

[2] T. Morioka, Y. Awaji, R. Ryf, P. Winzer, D. Richardson, and F. Poletti. Enhancing optical communications with brand new fibers. *Communications Magazine, IEEE*, 50(2):s31–s42, 2012.

[3] Yang Qin, L. Mason, and Ke Jia. Study on a joint multiple layer restoration scheme for IP over WDM networks. *Network, IEEE*, 17(2):43 – 48, mar/apr 2003.

[4] A. Sano, T. Kobayashi, S. Yamanaka, A. Matsuura, H. Kawakami, Y. Miyamoto, K. Ishihara, and H. Masuda. 102.3-Tb/s (224 * 548-Gb/s) C- and extended L-band all-Raman transmission over 240 km using PDM-64QAM single carrier FDM with digital pilot tone. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3.

[5] http://metroethernetforum.org/, [Online; accessed August-2014].

[6] IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks, Amendment 4: Provider Bridges. *IEEE Std 802.1ad-2005 (Amendment to IEEE Std 8021Q-2005)*, pages 1 –60, 2006.

[7] IEEE Standard for Local and metropolitan area networks – Virtual Bridged Local Area Networks Amendment 7: Provider Backbone Bridges. *IEEE Std 802.1ah-2008 (Amendment to IEEE Std 802.1Q-2005)*, pages 1 –110, 14 2008.

[8] IEEE Draft Standard for Local and Metropolitan Area Networks—Virtual Bridged Local Area Networks Amendment: Provider Backbone Bridge Traffic Engineering (PBB-TE) Infrastructure Segment Protection. *IEEE P802.1Qbf/D1.3, February 2011*, pages 1–84, 2011.

[9] IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management. *IEEE Std 802.1ag - 2007 (Amendment to IEEE Std 802.1Q - 2005 as amended by IEEE Std 802.1ad - 2005 and IEEE Std 802.1ak - 2007)*, pages 1 –260, 2007.

[10] M. Bocci et al. A Framework for MPLS in TransportNetworks, IETF RFC 5921, July 2010.

[11] C. Kachris, K. Kanonakis, and I. Tomkos. Optical interconnection networks in data centers: recent trends and future challenges. *Communications Magazine, IEEE*, 51(9):39–45, September 2013.

[12] Theophilus Benson, Aditya Akella, and David A. Maltz. Network Traffic Characteristics of Data Centers in the Wild. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, pages 267–280, New York, NY, USA, 2010. ACM.

[13] Nathan Farrington, George Porter, Sivasankar Radhakrishnan, Hamid Hajabdolali Bazzaz, Vikram Subramanya, Yeshaiahu Fainman, George Papen, and Amin Vahdat. Helios: a hybrid electrical/optical switch architecture for modular data centers. *SIGCOMM Comput. Commun. Rev.*, 41(4):–, August 2010.

[14] Roch A. Guerin, , and Ariel Orda. QoS Routing in Networks with Inaccurate Information: Theory and Algorithms. *IEEE/ACM Transactions on Networking*, 1997.

[15] A. Haider and R. Harris. Recovery techniques in next generation networks. *Communications Surveys Tutorials, IEEE*, 9(3):2 –17, quarter 2007.

[16] Dongyun Zhou and S.S. Subramaniam. Survivability in optical networks. *Network, IEEE*, 14(6):16–23, 2000.

[17] Ahmed E. Kamal. 1 + N network protection for mesh networks: network coding-based protection using p-cycles. *IEEE/ACM Trans. Netw.*, 18(1):67–80, February 2010.

[18] A.H.A. Muktadir, A.A. Jose, and E. Oki. An Optimum Mathematical Programming Model for Network-Coding Based Routing with 1+1 Path Protection. In *World Telecommunications Congress (WTC), 2012*, pages 1 –5, march 2012.

[19] E. Marin-Tordera, X. Masip-Bruin, S. Sanchez-Lopez, J. Sole-Pareta, and J. Domingo. The prediction-based routing in optical transport networks. *Computer Networks*, 29:865–878, 2006.

[20] I Chlamtac, A Ganz, and G. Karmi. Lightpath communications: an approach to high bandwidth optical WAN's. *Communications, IEEE Transactions on*, 40(7):1171–1182, Jul 1992.

[21] ITU-T. Architecture for the Automatically Switched Optical Network (ASON) - Rec. 8080/Y.1304, 2001.

[22] IETF. Generalized Multi-Protocol Label Switching (GMPLS) Architecture (RFC 3945), October 2004.

[23] Swarup Acharya, Yuh-Jye Chang, Bhawna Gupta, P. Risbood, and Anurag Srivastava. Precomputing high quality routes for bandwidth guaranteed traffic. In *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, volume 2, pages 1202–1207 Vol.2, Nov 2004.

[24] Norihito Fujita and A Iwata. Adaptive and efficient multiple path pre-computation for QoS routing protocols. In *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, volume 4, pages 2215–2219 vol.4, 2001.

[25] Jane M. Simmons. Diversity requirements for selecting candidate paths for alternative-path routing. In *Optical Fiber Communication (OFC), collocated National Fiber Optic Engineers Conference, 2010 Conference on (OFC/NFOEC)*, pages 1–3, March 2010.

[26] Kit-Man Chan and T.P. Yum. Analysis of least congested path routing in WDM lightwave networks. In *INFOCOM '94. Networking for Global Communications., 13th Proceedings IEEE*, pages 962–969 vol.2, Jun 1994.

[27] Ling Li and AK. Somani. Dynamic wavelength routing using congestion and neighborhood information. *Networking, IEEE/ACM Transactions on*, 7(5):779–786, Oct 1999.

[28] Hui Zang, L. Sahasrabuddhe, J.P. Jue, S. Ramamurthy, and B. Mukherjee. Connection management for wavelength-routed WDM networks. In *Global Telecommunications Conference, 1999. GLOBECOM '99*, volume 2, pages 1428–1432 vol.2, 1999.

[29] E. Karasan and E. Ayanoglu. Effects of wavelength routing and selection algorithms on wavelength conversion gain in WDM optical networks. In *Advanced Applications of Lasers in Materials Processing/Broadband Optical Networks/Smart Pixels/Optical MEMs and Their Applications. IEEE/LEOS 1996 Summer Topical Meetings:*, pages 43–44, Aug 1996.

[30] R. Barry and S.S. Subramaniam. The MAX SUM wavelength assignment algorithm for WDM ring networks. In *Optical Fiber Communication. OFC 97., Conference on*, pages 121–122, Feb 1997.

[31] Shizhong Xu, Lemin Li, Sheng Wang, and Chibiao Chen. Wavelength assignment for dynamic traffic in WDM networks. In *Networks, 2000. (ICON 2000). Proceedings. IEEE International Conference on*, pages 375–379, 2000.

[32] A Birman and A Kershenbaum. Routing and wavelength assignment methods in single-hop all-optical networks with blocking. In *INFOCOM '95. Fourteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Bringing Information to People. Proceedings. IEEE*, pages 431–438 vol.2, Apr 1995.

[33] Hui Zang and Jason P. Jue. A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks. *Optical Networks Magazine*, 1:47–60, 2000.

[34] Radia Perlman. *Interconnections (2Nd Ed.): Bridges, Routers, Switches, and Internetworking Protocols*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2000.

[35] Jian Liu, Gaoxi Xiao, Kejie Lu, and I Chlamtac. An Evaluation of Distributed Parallel Reservations in Wavelength-Routed Networks. *Selected Areas in Communications, IEEE Journal on*, 25(9):27–39, December 2007.

[36] S. Shen, Gaoxi Xiao, and Tee Hiang Cheng. The Performance of Periodic Link-State Update in Wavelength-Routed Networks. In *Broadband Communications, Networks and Systems, 2006. BROADNETS 2006. 3rd International Conference on*, pages 1–10, 2006.

[37] O. Komolafe and J. Sventek. Impact of GMPLS Control Message Loss. *Lightwave Technology, Journal of*, 26(14):2029–2036, July 2008.

[38] Jun Zhou and Xin Yuan. A Study of Dynamic Routing and Wavelength Assignment with Imprecise Network State Information. In *in Proceedings of International Conference on Parallel Processing Workshops (ICPPW 02)*, pages 202–207, 2002.

[39] X. Masip-Bruin, S. Sanchez-Lopez, and D. Colle. Routing and wavelength assignment under inaccurate routing information in networks with sparse and limited wavelength conversion. In *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, volume 5, pages 2575–2579 vol.5, Dec 2003.

[40] Bo Li and Xiaowen Chu. Routing and wavelength assignment vs. wavelength converter placement in all-optical networks. *Communications Magazine, IEEE*, 41(8):S22–S28, Aug 2003.

[41] A. Agarwal and M. Charikar. On the advantage of network coding for improving network throughput. In *Information Theory Workshop, 2004. IEEE*, pages 247–249, 2004.

[42] H. Overby, G. Biczok, P. Babarczi, and J. Tapolcai. Cost comparison of 1+1 path protection schemes: A case for coding. In *Communications (ICC), 2012 IEEE International Conference on*, pages 3067–3072.

[43] S. Nazim and E. Ayanoglu. Network Coding-Based Link Failure Recovery over Large Arbitrary Network. Globecom, 2013.

[44] S.N. Avci, Xiaodan Hu, and E. Ayanoglu. Hitless recovery from link failures in networks with arbitrary topology. In *Information Theory and Applications Workshop (ITA), 2011*, pages 1–6, 2011.

[45] W. Ramirez, X. Masip-Bruin, M. Yannuzzi, R. Serral-Gracia, and A. Martinez. An Efficient Protection Strategy Using Multiple Network Coding. In *International Workshop on Network Management Innovations, SACONET 2013, Paris, France*, June 2013.

[46] "Python web page", http://www.python.org/.

[47] http://networkx.github.io/, [Online; accessed August-2014].

[48] Spectral grids for WDM applications: DWDM frequency grid, ITU-T Recommendation G.694.1 available at http://www.itu.int.

[49] O. Gerstel, M. Jinno, A. Lord, and S. J B Yoo. Elastic optical networking: a new dawn for the optical layer? *Communications Magazine, IEEE*, 50(2):s12–s20, 2012.

[50] M. Jinno, Takuya Ohara, Y. Sone, A. Hirano, O. Ishida, and M. Tomizawa. Elastic and adaptive optical networks: possible adoption scenarios and future standardization aspects. *Communications Magazine, IEEE*, 49(10):164–172, 2011.

[51] Shuqiang Zhang and B. Mukherjee. Energy-efficient dynamic provisioning for spectrum elastic optical networks. In *Communications (ICC), 2012 IEEE International Conference on*, pages 3031–3035, 2012.

[52] A. Klekamp, Gebhard, and F. Ilchmann. Efficiency of adaptive and mixed-line-rate IP over DWDM networks regarding CAPEX and power consumption [Invited]. *Optical Communications and Networking, IEEE/OSA Journal of*, 4(11):B11–B16, 2012.

[53] A. Muhammad, P. Monti, I. Cerutti, L. Wosinska, P. Castoldi, and A. Tzanakaki. Energy-Efficient WDM Network Planning with Dedicated Protection Resources in Sleep Mode. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–5, 2010.

[54] J. Lopez, Yabin Ye, V. Lopez, F. Jimenez, R. Duque, P.M. Krummrich, F. Musumeci, M. Tornatore, and A. Pattavina. Traffic and power-aware protection scheme in Elastic Optical Networks. In *Telecommunications Network Strategy and Planning Symposium (NETWORKS), 2012 XVth International*, pages 1–6, 2012.

[55] Peter Babarczi, Gergely Biczok, Harald Overby, Janos Tapolcai, and Peter Soproni. Realization strategies of dedicated path protection: A bandwidth cost perspective. *Computer Networks*, 57(9):1974 – 1990, 2013.

[56] M. Manzano, J.-L. Marzo, E. Calle, and A. Manolovay. Robustness analysis of real network topologies under multiple failure scenarios. In *Networks and Optical Communications (NOC), 2012 17th European Conference on*, pages 1–6, 2012.

[57] A. Kwasinski. Effects of notable natural disasters from 2005 to 2011 on telecommunications infrastructure: Lessons from on-site damage assessments. In *Telecommunications Energy Conference (INTELEC), 2011 IEEE 33rd International*, pages 1–9, 2011.

[58] Department of Homeland-Security. Pandemic Influenza Impact on Communications Networks Study, unclassified, Department of Homeland Security (DHS). December 2007.

[59] Telefonica network planning report [Internal]. Technical report, 2013.

[60] S. Borgatti and M. Everett. A Graph-theoretic perspective on centrality. *Social Networks*, 28(4):466–484, October 2006.

[61] J.P Vasseur et al. *Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*. David Clark, 2004.

[62] F. Rambach, B. Konrad, L. Dembeck, U. Gebhard, M. Gunkel, M. Quagliotti, L. Serra, and V. Lopez. A multilayer cost model for metro/core networks. *Optical Communications and Networking, IEEE/OSA Journal of*, 5(3):210–225, 2013.

[63] D. Kong, Y. Li, H. Wang, S. Zhou, J. Zang, J. Zhang, J. Wu, and J. Lin. All-optical XOR gates for QPSK signal based optical networks. *Electronics Letters*, 49(7):486–488, 2013.

[64] T. Borgar and K. Stordahl. Models for forecasting cost evolution of components and technologies, 2004.

[65] C. Dorize, W. Van Heddeghem, F. Smyth, E. Le Rouzic, and B. Arzur. draft report on baseline power consumption, version 1.8. Technical report, Greentouch, 2011.

[66] O. Gerstel and R. Ramaswami. Optical layer survivability: a post-bubble perspective. *Communications Magazine, IEEE*, 41(9):51–53, Sept 2003.

[67] Min Zhang, Ling Wang, and Peida Ye. All optical XOR logic gates: technologies and experiment demonstrations. *Communications Magazine, IEEE*, 43(5):S19–S24, 2005.

[68] Yong-Kee Yeo, Jianjun Yu, and Gee-Kung Chang. Performance of DPSK and NRZ-OOK signals in a novel folded-path optical packet switch buffer. In *Optical Fiber Communication Conference, 2005. Technical Digest. OFC/NFOEC*, volume 3, pages 3 pp. Vol. 3–, March 2005.

[69] Y. Zhang, H. Zhang, and M. Yao. Modular and intelligently controlled optical buffer with last-in-first-out priority. *Electronics Letters*, 47(22):1236–1238, Oct 2011.

[70] H. Kiuchi. Highly Stable Millimeter-Wave Signal Distribution With an Optical Round-Trip Phase Stabilizer. *Microwave Theory and Techniques, IEEE Transactions on*, 56(6):1493–1500, June 2008.

[71] Kejie Lu, Gaoxi Xiao, and I. Chlamtac. Analysis of blocking probability for distributed lightpath establishment in WDM optical networks. *Networking, IEEE/ACM Transactions on*, 13(1):187–197, Feb 2005.

[72] X. Masip-Bruin, S. Sanchez-Lopez, J. Sole-Pareta, and J. Domingo-Pascual. QoS routing algorithms under inaccurate routing for bandwidth constrained applications. In *Communications, 2003. ICC '03. IEEE International Conference on*, volume 3, pages 1743–1748 vol.3, 2003.

[73] A. Al-Fuqaha, G. Chaudhry, C. Beard, M. Guizani, M. Labrador, and I. Habib. Link-state update policies for all-optical DWDM transport networks. In *Communications, 2004 IEEE International Conference on*, volume 3, pages 1831–1835 Vol.3, June 2004.

[74] S. Shen, Gaoxi Xiao, and Tee-Hiang Cheng. A novel method of link-state update in wavelength-routed networks. *Lightwave Technology, Journal of*, 24(3):1112–1120, March 2006.

[75] Jun Zheng and H.T. Mouftah. Distributed lightpath control based on destination routing for wavelength-routed WDM networks. In *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, volume 3, pages 1526–1530 vol.3, 2001.

[76] James E. Smith. A Study of Branch Prediction Strategies. In *Proceedings of the 8th Annual Symposium on Computer Architecture*, ISCA '81, pages 135–148, Los Alamitos, CA, USA, 1981. IEEE Computer Society Press.

[77] E. Ahvar, E. Marin-Tordera, M. Yannuzzi, X. Masip-Bruin, and S. Ahvar. FRA: A new fuzzy-based routing approach for optical transport networks. In *Networks and Optical Communications (NOC), 2012 17th European Conference on*, pages 1–6, 2012.

[78] K. Kompella and Y. Rekhter. OSPF extensions in support of generalized multi-protocol label switching (GMPLS), RFC 4203.

[79] Cisco. *Cisco Data Center Infraestructure 2.5 Design Guide*. 2010.

[80] Software-defined networking: the service provider perspective. *Ericsson Review*, 2013.

[81] Open Networking foundation. Openflow Switch Specification Version 1.1.0 http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf.

[82] Marta Cuaresma Saturio, Victor Lopez, Oscar Gonzalez Dios, Fernando Munoz del Nuevo, and JuanPedro Fernandez-Palacios. Implementation and Assessment of Pre-reservation Mechanism for PCE Environments. *Journal of Network and Systems Management*, pages 1–21, 2013.

[83] X. Masip-Bruin, E. Marin-Tordera, M. Yannuzzi, R. Serral-Gracia, and S. Sanchez-Lopez. Reducing the effects of routing inaccuracy by means of prediction and an innovative link-state cost. *Communications Letters, IEEE*, 14(5):492–494, 2010.

[84] A. Zapata-Beghelli and P. Bayvel. Dynamic Versus Static Wavelength-Routed Optical Networks. *Lightwave Technology, Journal of*, 26(20):3403–3415, 2008.

[85] http://www.omnetpp.org/, [Online; accessed September 2014].

[86] R.A. Barry and P.A. Humblet. Models of blocking probability in all-optical networks with and without wavelength changers. *Selected Areas in Communications, IEEE Journal on*, 14(5):858–867, Jun 1996.

[87] Konstantinos V. Miliotis, Georgios I. Papadimitriou, and Andreas S. Pomportsis. On the use of adaptive weight functions in wavelength-continuous WDM multi-fiber networks under dynamic traffic. *Communications and Networks, Journal of*, 7(4):499–508, Dec 2005.

[88] M. Tornatore, A. Baruffaldi, Hongyue Zhu, B. Mukherjee, and A. Pattavina. Holding-Time-Aware Dynamic Traffic Grooming. *Selected Areas in Communications, IEEE Journal on*, 26(3):28–35, April 2008.

[89] G Shen, T.H Cheng, S.K Bose, C Lu, T.Y Chai, and H.M.M Hosseini. Approximate analysis of limited-range wavelength conversion all-optical {WDM} networks. *Computer Communications*, 24(10):949 – 957, 2001.

[90] W. Ramirez, X. Masip-Bruin, M. Yannuzzi, D. Montero, A. Martinez, and V. Lopez. Network coding-based protection scheme for Elastic Optical Networks. In *Design of Reliable Communication Networks (DRCN), 2014 10th International Conference on the*, pages 1–8, April 2014.

[91] C. Pinart. A multilayer fault localization framework for IP over all-optical multilayer networks. *Network, IEEE*, 23(3):4 –9, may-june 2009.

[92] A. Zinin A. Atlas. Basic Specification for IP Fast Reroute: Loop-Free Alternates, 2008.

[93] A.F. Hansen, A. Kvalbein, T. Cicic, S. Gjessing, and O. Lysne. Resilient routing layers for recovery in packet networks. In *Dependable Systems and Networks, 2005. DSN 2005. Proceedings. International Conference on*, pages 238 – 247, june-1 july 2005.

[94] G. Enyedi A. Csaszar J. Tantsura M. Konstantynowicz R. White M. Shand A. Atlas, R. Kebler. An Architecture for IP/LDP Fast-Reroute Using Maximally Redundant Trees, 2013.

[95] C. Hopps D. Thaler. Multipath Issues in Unicast and Multicast Next-Hop Selection, 2000.

[96] Muriel Medard, Richard A. Barry, Steven G. Finn, Wenbo He, and Steven S. Lumetta. Generalized loop-back recovery in optical mesh networks. *IEEE/ACM Trans. Netw.*, 10(1):153–164, February 2002.

[97] J. Zheng and H.T. Mouftah. Destination-initiating path restoration protocol for wavelength-routed WDM networks. *Communications, IEE Proceedings-*, 149(1):18 –22, feb 2002.

[98] Jitender S. Deogun L. Kong, M. Ali. Building Redundant Multicast Trees for Preplanned Recovery in WDM optical networks. *High Speed Networks*, 15:379–398, 2006.

[99] W. D. Grover and D. Stamatelakis. Bridging the Ring-Mesh Dichotomy With P-Cycles, 2000.

[100] Hong Huang and J. Copeland. Hamiltonian cycle protection: a novel approach to mesh WDM optical network protection. In *High Performance Switching and Routing, 2001 IEEE Workshop on*, pages 31 –35, 2001.

[101] A. V. Sichani and H. T. Mouftah. Maximum Mutual Links-kth Shortest Path Protocol A Survivable Routing Scheme for WDM Mesh Networks. In *Proceedings of IST 05*, number 109-116, 2005.

[102] A. Giorgetti, L. Valcarenghi, F. Cugini, and P. Castoldi. PCE-Based Dynamic Restoration in Wavelength Switched Optical Networks. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1 –6, may 2010.

[103] R. Shenai, C. Maciocco, M. Mishra, and K. Sivalingam. Threshold based selective link restoration for optical WDM mesh networks. In *Design of Reliable Communication Networks, 2003. (DRCN 2003). Proceedings. Fourth International Workshop on*, pages 31 – 38, oct. 2003.

[104] Mohamed Mostafa A. Azim, Xiaohong Jiang, and Susumu Horiguchi. On Modeling the Restoration Probability of Active Restoration-Based Optical Networks with Correlation Among Backup Routes. In *Frontier of Computer Science and Technology, 2006. FCST '06. Japan-China Joint Workshop on*, pages 23 –38, nov. 2006.

[105] B.T. Mangara and F.W. Leuschner. Automated dynamic protection planning for improved survivability of DWDM optical networks: performance analysis for distributed restoration algorithms. In *AFRICON, 2004. 7th AFRICON Conference in Africa*, volume 2, pages 977 – 979 Vol.2, sept. 2004.

[106] J. Perello, S. Spadaro, F. Agraz, M. Angelou, S. Azodolmolky, Y. Qin, R. Nejabati, D. Simeonidou, P. Kokkinos, E. Varvarigos, S. Al Zahr, M. Gagnaire, and I. Tomkos. Experimental evaluation of centralized failure restoration in a dynamic impairment-aware all-optical network. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1 –3, march 2011.

[107] F. Munoz, V. Lopez, O.G. de Dios, and J.P. Fernandez-Palacios. Multi-layer restoration in hierarchical IP/MPLS over WSON networks. In *Networks and Optical Communications (NOC), 2012 17th European Conference on*, pages 1–6, 2012.

[108] Lei Lei, Aibo Liu, and Yuefeng Ji. A joint resilience scheme with interlayer backup resource sharing in IP over WDM networks. *Communications Magazine, IEEE*, 42(1):78 – 84, jan 2004.

[109] P. Pacharintanakul. The effects of multi-layer traffic on the survivability of IP-over-WDM networks. In *Proceedings of the 2009 IEEE international conference on Communications*, ICC'09, pages 2354–2359, Piscataway, NJ, USA, 2009. IEEE Press.

[110] E. Palkopoulou, D.A. Schupke, and T. Bauschert. Shared backup router resources: realizing virtualized network resilience. *Communications Magazine, IEEE*, 49(5):140 –146, may 2011.

[111] K. Kompella, Y. Rekhter, A. Banerjee, J. Drake, G. Bernstein, D. Fedyk, E. Mannie, D. Saha, and V. Sharma. OSPF Extensions in Support of Generalized MPLS, draft-kompella-ospf-gmpls-extensions-02.txt, January 2002.

[112] K. Sato, N. Yamanaka, Y. Takigawa, M. Koga, S. Okamoto, K. Shiomoto, E. Oki, and W. Imajuku. GMPLS-based photonic multilayer router (Hikari router) architecture: an overview of traffic engineering and signaling technology. *Communications Magazine, IEEE*, 40(3):96 –101, mar 2002.

[113] M. Yannuzzi, X. Masip-Bruin, O.G. de Dios, C.G. Argos, M. Maciejewski, and J. Altmann. Bridging the interoperability gap between the Internet and optical network management systems. In *Networks and Optical Communications (NOC), 2011 16th European Conference on*, pages 177 –180, july 2011.

[114] L. Yang, R. Dantu, T. Anderson, and R. Gopal. Forwarding and Control Element Separation (ForCES) Framework, RFC 3746, Apr 2004.

[115] O. de Dios, V. Lopez, M. Cuaresma, F. Munoz, M. Chamania, and A Jukan. Coordinated computation and setup of multi-layer paths via inter-layer PCE communication: standards, interoperability and deployment. *Communications Magazine, IEEE*, 51(12):144–154, December 2013.

[116] J. Moy. OSPF Version 2, April 1998.

[117] M. Bjorklund. YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF), October 2010.

[118] Jijun Zhao, Lei Lei, Yuefeng Ji, and Daxiong Xu. Integrated multilayer survivability strategy with inter-layer signaling. In *Communication Technology Proceedings, 2003. ICCT 2003. International Conference on*, volume 1, pages 612 – 616 vol.1, april 2003.

[119] D. Coudert, S. Perennes, H. Rivano, and Marie-Emilie Voge. Shared Risk Resource Groups and Survivability in Multilayer Networks. In *Transparent Optical Networks, 2006 International Conference on*, volume 3, pages 235 –238, june 2006.

[120] G. Booker, A. Sprintson, E. Zechman, C. Singh, and S. Guikema. Efficient traffic loss evaluation for transport backbone networks. *Comput. Netw.*, 54(10):1683–1691, July 2010.

[121] J. Lang. Link Management Protocol (LMP), RFC-4204, 2005.

[122] Choonho Son, Junsuk Oh, Kyoung-Ho Lee, Kieung Kim, and Jaehyung Yoo. Efficient physical topology discovery for large OSPF networks. In *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE*, pages 325–330, 2008.

[123] A. Farrel and J.-P Vasseur. A Path Computation Element (PCE)-Based Architecture, RFC 4655, August 2006.

[124] Y. Vasseur, JP. Ikejiri and R. Zhang. OSPF Protocol Extensions for Path Computation Element (PCE) Discovery, Jannuary 2008.

[125] G. Lee, Y. Bernstein and W. Imajuku. Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs), RFC 6163, April 2011.

[126] JL. Le Roux JP. Vasseur. Path Computation Element (PCE) Communication Protocol (PCEP), March 2009.

[127] M. Chamania, O.G. de Dios, V. Lopez, M. Cuaresma, M. Drogon, A. Jukan, X. Masip-Bruirr, and M. Yannuzzi. Coordinated computation of multi-layer paths via inter-layer PCE communication: Standards, interoperability and deployment. In *Communications (ICC), 2012 IEEE International Conference on*, pages 6926–6931, 2012.

[128] D.A. Schupke and F. Rambach. Path computation element; An enabler for automated network operation. In *Networks and Optical Communications (NOC), 2011 16th European Conference on*, pages 12 –15, july 2011.

[129] http://www.topology-zoo.org/, [Online; accessed September 2014].

[130] http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/cdp/configuration/15-mt/cdp-15-mt-book/nm-cdp-discover.html.

[131] http://www.nongnu.org/quagga/, [Online; accessed September 2014].

[132] http://www.nrl.navy.mil/itd/ncs/products/mgen, [Online; accessed August-2014].

[133] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of Things: A survey. *Computer Networks*, 54(15):2787 – 2805, 2010.

[134] Luca Cittadini, Wolfgang Muhlbauer, Steve Uhlig, Randy Bush, Pierre Francois, and Olaf Maennel. Evolution of internet address space deaggregation: myths and reality. *IEEE J.Sel. A. Commun.*, 28(8):1238–1249, October 2010.

[135] D. Wilson N. Murphy. The End of Eternity Part One: IPv4 Address Exhaustion and Consequences. *The Internet Protocol Journal*, 11, 2008.

[136] Xiaoliang Zhao, D.J. Pacella, and J. Schiller. Routing Scalability: An Operator's View. *Selected Areas in Communications, IEEE Journal on*, 28(8):1262 –1270, october 2010.

[137] M. Yannuzzi, X. Masip-Bruin, and O. Bonaventure. Open issues in interdomain routing: a survey. *Network, IEEE*, 19(6):49 – 56, nov.-dec. 2005.

[138] G. Huston. The Changing Foundation of the Internet: Confronting IPv4 Address Exhaustion. *Internet Protocol Journal*, 11, 2008.

[139] T.Hain. A Pragmatic Report on IPv4 Address Space Consumption. *The Internet Protocol Journal*, 2008.

[140] G.Huston. IPv4 Address Depletion. *The Internet Protocol Journal*, 3, 2010.

[141] http://www.cisco.com/go/lisp.

[142] D. Meyer D. Lewis. D. Farinacci, V. Fuller. Locator/Id Separation Protocol (LISP), November 2012.

[143] J. Noel Chiappa. Endpoints and Endpoint Names: A proposed enhancement to the internet Architecture, 1999.

[144] General requirements for ID/locator separation in NGN.

[145] L. Zhang D. Meyer and K. Fall. Report from the IAB Workshop on Routing and Addressing, September 2007.

[146] T. Bu, Lixin Gao, and D. Towsley. On characterizing BGP routing table growth. In *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, volume 3, pages 2185–2189 vol.3, Nov 2002.

[147] P. Nikander R. Moskowitz. Host Identity Protocol (HIP) Architecture,RFC 4423.

[148] Bruno Quoitin, Luigi Iannone, Cédric de Launois, and Olivier Bonaventure. Evaluating the Benefits of the Locator/Identifier Separation. In *Proceedings of 2Nd ACM/IEEE International Workshop on Mobility in the Evolving Internet Architecture*, MobiArch '07, pages 5:1–5:6, New York, NY, USA, 2007. ACM.

[149] M. Yannuzzi, X. Masip-Bruin, E. Grampin, R. Gagliano, A Castro, and M. German. Managing interdomain traffic in Latin America: a new perspective based on LISP. *Communications Magazine, IEEE*, 47(7):40–48, July 2009.

[150] https://www.ietf.org/rfc/rfc2281.txt [Online; accessed September 2014].

[151] D. King. A. Farrel. Unswered Questions in the Path Computation Element Architecture draft-ietf-pce-questions-01.txt., October. 2013.

[152] N. Ghani, Qing Liu, D. Benhaddou, N.S.V. Rao, and T. Lehman. Control plane design in multidomain/multilayer optical networks. *Communications Magazine, IEEE*, 46(6):78–87, June 2008.

[153] X. Masip, Ren Guang-Jie, R. Gracia, and M. Yannuzzi. Unlocking the Value of Open Data with a Process-based Information Platform. In *presented at IEEE CBI 2013, Vienna, Austria.*, 2013.

[154] S. Dustdar and M. Gaedke. The Social Routing Principle. *Internet Computing, IEEE*, 15(4):80–83, July 2011.

[155] Hui Zang, J.P. Jue, L. Sahasrabuddhe, R. Ramamurthy, and B. Mukherjee. Dynamic lightpath establishment in wavelength routed WDM networks. *Communications Magazine, IEEE*, 39(9):100–108, Sep 2001.

[156] A. Martinez, X. Masip-Bruin, W. Ramirez, R. Serral-Gracia, E. Marin-Tordera, and M. Yannuzzi. Toward a new addressing scheme for a service-centric Internet. In *Communications (ICC), 2012 IEEE International Conference on*, pages 6463–6467, June 2012.

[157] Hongbin Luo, Yajuan Qin, and Hongke Zhang. A DHT-Based Identifier-to-Locator Mapping Approach for a Scalable Internet. *Parallel and Distributed Systems, IEEE Transactions on*, 20(12):1790–1802, Dec 2009.

[158] L. Jakab, A Cabellos-Aparicio, F. Coras, D. Saucez, and O. Bonaventure. LISP-TREE: A DNS Hierarchy to Support the LISP Mapping System. *Selected Areas in Communications, IEEE Journal on*, 28(8):1332–1343, October 2010.

[159] Jaeyeon Jung, Emil Sit, Hari Balakrishnan, and Robert Morris. DNS Performance and the Effectiveness of Caching. *IEEE/ACM Trans. Netw.*, 10(5):589–603, October 2002.

[160] http://www.google.com/trends/, [Online; accessed September 2014].

[161] S.A Aly, V. Kapoor, Jie Meng, and A Klappenecker. Bounds on the Network Coding Capacity for Wireless Random Networks. In *Information Theory and Applications Workshop, 2007*, pages 231–236, Jan 2007.

[162] R. Bassoli, H. Marques, J. Rodriguez, K.W. Shum, and R. Tafazolli. Network Coding Theory: A Survey. *Communications Surveys Tutorials, IEEE*, 15(4):1950–1978, Fourth 2013.

[163] R.C. Menendez and J.W. Gannet. Efficient, Fault-Tolerant All-Optical Multicast Networks via Network Coding. In *Optical Fiber communication/National Fiber Optic Engineers Conference, 2008. OFC/NFOEC 2008. Conference on*, pages 1–3, Feb 2008.

[164] S.A. Aly and A.E. Kamal. Network Coding-Based Protection Strategy Against Node Failures. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1–5, 2009.

[165] S.N. Avci and E. Ayanoglu. Network coding-based link failure recovery over large arbitrary networks. In *Global Communications Conference (GLOBECOM), 2013 IEEE*, pages 1519–1525, Dec 2013.

[166] IB. Barla, Franz Rambach, D.A Schupke, and M. Thakur. Network Coding for Protection against Multiple Link Failures in Multi-Domain Networks. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–6, May 2010.

[167] C. Gkantsidis and P.R. Rodriguez. Cooperative Security for Network Coding File Distribution. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1–13, April 2006.

[168] Kejie Lu, Shengli Fu, Yi Qian, and Tao Zhang. On the Security Performance of Physical-Layer Network Coding. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1–5, June 2009.

[169] A Martinez, C. Garcia-Meca, R. Ortuno, F.J. Rodriguez-Fortuno, and J. Marti. Metamaterials for optical security. *Applied Physics Letters*, 94(25):251106–251106–3, Jun 2009.

¡

Publications

- A. Martinez, W. Ramirez, M. German, R. Serral, E. Marin, M. Yannuzzi, and X. Masip-Bruin. "An approach to a fault tolerance LISP architecture", In Proceedings of the 9th IFIP TC 6 international conference on Wired/wireless internet communications (WWIC'11), Xavier Masip-Bruin, Marcelo Yannuzzi, Dominique Verchere, and Vassilis Tsaoussidis (Eds.). Springer-Verlag, Berlin, Heidelberg, 338-349.

- W. Ramirez, A. Martinez, R.Serral, E. Marin-Tordera, M. Yannuzzi, X. Masip-Bruin, "Dealing with availability and reachability issues in a LISP architecture", In Proceedings of 1th W-FIERRO 2011 July 7-8, 2011 Cartagena, Spain.

- A. Martinez, X. Masip-Bruin, W. Ramirez, M. Yannuzzi, R. Serral-Gracia, E. Marin-Tordera, "Toward a New Addressing Scheme for a Service-Centric Internet",3rd IEEE International Workshop on Smart Communications in Network Technologies (SaCoNeT-III), presented in ICC 2012, Ottawa, Canada, June 2012.

- W. Ramirez, X. Masip-Bruin, Yannuzzi, R. M. Serral-Gracia, and A. Martinez. An Efficient Protection Strategy Using Multiple Network Coding. In International Workshop on Network Management Innovations, SACONET 2013, Paris, France, June 2013.

- W. Ramirez, X. Masip-Bruin, M. Yannuzzi, R. Serral-Gracia , A. Martinez, S. Siddiqui, "A survey and taxonomy of ID/Locator Split Architectures, Computer Networks", February 2014, pp. 13-33.

- W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, M. Yannuzzi, MS. Siddiqui, A. Martinez, V. Lopez, "Improving Reliability in Multi-Layer Networks With Network Coding Protection", International Conference on Optical Networking Design and Modeling (ONDM), 2014.

- W. Ramirez, X. Masip-Bruin, M. Yannuzzi, D. Montero, A. Martinez, V. Lopez, "Network Coding-Based Protection Scheme for Elastic Optical Networks", International Conference on Design of Reliable Communication Networks, 2014.

- A. Martinez, M. Yannuzzi, V. Lopez, D. Lopez, M. Chamania, A. Jukan, W. Ramirez, X. Masip-Bruin, R. Serral-Gracia, M. Maciejewski, J. Altmann, "A Survey on Carrier-Grade Network Management: Interoperability Challenges and Trends in Multi-Layer and Multi-Vendor Settings", IEEE Communications Surveys & Tutorials, 2014.

- W. Ramirez, X. Masip-Bruin, M. Yannuzzi, E. Marin-Tordera, A. Martinez, V. Lopez, “An Hybrid Prediction-based Routing Approach for Reducing Routing Inaccuracy in Optical Transport Networks”, 19th European Conference on Networks and Optical Communications, 2014.

- M.S. Sidiqui, D. Montero, M. Yannuzzi, R. Serral-Garcia, X. Masip-Bruin, W. Ramirez, “Route Leak Detection Using Real-Time Analytics on local BGP Information”, Proceedings of IEEE GLOBECOM 2014 .

- W. Ramirez, X. Masip-Bruin, M. Yannuzzi, E. Marin-Tordera, A. Martinez, V. Lopez, “A Techno-Economic Study of Network Coding Protection Schemes”, Proceedings of IEEE GLOBECOM 2014” .

- W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, “A Context-Aware PCE”, PACE Workshop, In International Workshop on Network Management Innovations”, SACONET 2014, Vila Nova i la Geltru, Spain, June 2014.

Ongoing Work

- W. Ramirez, X. Masip-Bruin, M. Yannuzzi, A. Martinez, D. Montero, V. Lopez, “Managing Resilience in Carrier-Grade Networks: Open Issues and Trends. Submitted to Computer Communications” 2014.

- W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, M. Yannuzzi, A. Martinez, “On the Routing and Wavelength Assignment in the Presence of Inaccuracy: A New Challenge for the Design of Data Center Networks”, Submitted to Optical Switching and Networking, 2014.

- Wilson Ramirez, Xavier Masip; Eva Marin-Tordera, Marcelo Yannuzzi, Anny Martinez; Victor Lopez, “Dynamic Network Coding Protection under Inaccurate Network State Information”, Submitted to Computer Networks, Special Issue: Fault Tolerant Networks, 2014.

- W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, R. Casellas, “A Context-Aware PCE: Where PCE meets IoT”, to be submitted to Optical Fiber Communication (OFC) Conference, 2015.

- W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, R. Casellas, “Positioning the PCE in an IoT scenario”, to be to submitted to Communication Magazine.

- A. Martinez, M. Yannuzzi, R. Serral-Gracia and W. Ramirez. "Ontology Based Information Extraction from the Command Line of Network Routers", International Conference on Mining Intelligence and Knowledge Exploration, 2014 (submitted, under review)

- A. Martinez, M. Yannuzzi, R. Serral-Gracia and W. Ramirez. "An Ontology-Based Information Extraction System to complement OF-CONFIG for SDN-based Network Man-

agement", IFIP/IEEE International Symposium on Integrated Network Management 2015. (to be submitted)

Overview of Network Coding

This appendix section presents in a nutshell the state of the art regarding the use of NC for enabling resilience in wired networks.

Coding theory is related to the study of the codes and their use for specific applications. Codes can be used used for data compression, cryptography, error-correction or for Network Coding (NC).

NC is a technique commonly used for throughput reduction specifically in multicast and wireless network scenarios [41],[161]. This technique consists in the aggregation (compression) of several data streams by means of coding strategies (typically an Exclusive-Or operation) in order to reduce the network resources allocated to data transmission. There are several coding techniques such as Variable-rate Linear NC, Random Linear NC, Vector NC, among others. Due to its low complexity, Linear NC is commonly assumed in network research related to wired networks. For more information related to coding techniques the reader is referred to [162]. In recent years, there is a trend in network research that consists in using NC to enable resilience in wired networks. By means of NC a protection scheme can improve network throughput and thereby reduce power consumption.

Early works related to NCP for wired networks can be found in [163], [164]. It is worth noting that these works do not provide a evaluation of NCP schemes considering distinct network topologies configurations. Other works such as [17] proposed the use of NC combined with a 1+N protection strategy on p-cycles. Moreover, the studies available in [42], [18] and [45] proposed to enhance conventional DP schemes with NC features. Another work dealing with NCP for wired networks can be found in [165], which presented a novel coding structure and discussed design issues of an NCP scheme.

Moreover, contrary to the works described above proposing to use NC for link/path protection in single or multiple failure scenarios in single domain networks, the work presented in [166] addressed the use of network coding to endow multidomain networks with resilience capabilities.

On the other hand, there are several studies in network research that focus on offering security features to NC schemes, specifically concerning the integrity and confidentiality of the coded data. The latest is related to guarantee that coded data can be only accessed or decoded

by trusted nodes, while the former is necessary to verify that coded data is not polluted by malicious nodes. For more information about securing network coding the reader is referred to [167], [168]. These works are focused on P2P, wireless, or ad-hoc network scenarios, hence, they do not address network coding security for intra-domain or interdomain wired networks which are also vulnerable to security threats. Moreover, the algorithms proposed in these works operate electrically, thus, if an NCP scheme is endowed with security features this one must be deployed at the IP/MPLS network layer (rather than at the Optical network layer). This may be an issue for a network operator that opted for Optical layer protection, because the optical implementation of security operations is harder compared to electrically. Nevertheless, in recent years there are works in the research literature that attempt to provide information security purely optical [169]. In summary, most of the works introduced in this section present evaluations of NCP in distinct network topologies in an agnostic manner. For instance, it is not considered the network layer technology, IP or Optical. Indeed, there is limited information in network research regarding the performance of NCP deployed over Optical and IP topologies, and the advantages that this may bring to a network provider concerning its CAPEX investments. This thesis addresses this issue.

On other hand, it must be remarked that there are two concepts which must be considered on any NCP scheme as relevantly impacting on their performance. 1) Protection groups: A protection group is the possible combination of links that are suitable for network coding protection. This thesis assumes the formation of protection group with data streams that have common terminal vertices is suitable to reduce the complexity of an NCP scheme. 2) Coding Paths. A coding path is defined as the path carrying the coded (protected) traffic for a particular set of primary data streams to be protected.

Personal Information

First Name: Wilson.

Last Name: Ramirez Almonte

Birth Date: 18/12/1986

Email: wramirez@ac.upc.edu

Main Research Topics:

- **Routing algorithms**: Routing for unprotected and protected scenarios consisting in IP, Optical and Wireless technologies.
- **Future Internet Architectures.** Evaluate the benefits of new network paradigms such as Optical Wireless, context-aware communication, Smart Cities, Data Center Networks.
- **Network Optimization:** Augment network features in an optimal manner (network resources, cost), e.g., protection, mobility, and energy-efficiency

Academic Status

University: Polytechnic University of Catalonia (UPC), Spain.

Department: Computer Architecture.

Academic Formation

Bachelor's Degree: Electronics and Telecommunication Engineer, Technological Institute of Santo Domino (INTEC), Dominican Republic, 2009.

Master's Degree: Communication and Information Technology, Polytechnic University of Cartagena (UPCT), Spain,2010.

Ph.D's Degree: Polytechnic University of Catalonia (UPC), Advanced Network Architectures Lab (CRAAX), Computer Architecture Spain, ***(ongoing, final year).***

Technical Formation

- Cisco Certified Network Associate Routing & Switching (CCNA) , 2008.
- Seminar of Microwave and Radio-frequency: Harry Stratex, 2009.
- Certification Alvarium (VL_Brezze Max Access), 2009.
- Certification Wimax Networks: (Network Overview,Configuration,Security and Faults Nortel), 2009.
- Certification Motorola PTP-600, PTP 100, Moto Wi4, 2009.
- Seminario Arquitecturas de software un enfoque en programación dirigida por modelos, Universidad Politécnica de Cartagena, 2010
- Curso procesamiento de señales aplicadas a imágenes médicas, Universidad Politécnica de Cartagena, 2010.

Main Skills

- IP Networks: Management, Configuration and Design.
- Wireless Networks: Management, Configuration and Design.
- Programming: C++, Python.
- High experience related to the management and preparation of research and technical projects.
- Management of working groups and human resources.

Honors and Awards

- Formación de Personal Investigador --Research Personnel Training--(FPI-MICINN) granted by the Ministry of Science and Innovation of Spain,2010 to 2014.
- Beca Estudios Master en el extranjero (Master Scholarship), granted by Secretary of Higher Education, Santo Domingo, Dominican Republic, 2009.
- Fellowship graduate training section, Polytechnic University of Cartagena, 2010.

- Fellowship for Course of Digital Signal Processing Applied to Medical Images, Polytechnic University Cartagena (UPCT), 2010.
- Research Support, Polytechnic University of Catalonia (UPC), 2013.

Teaching Experience

April,2013. Teaching Asssistant, Polytechnic University of Catalonia (UPC)

Course: Future Internet, Graduate course.

Professional Experience

- 18/04/2008 - 14/08/2009, at Wind telecom (Internet Service Provier), Dominican Republic.
- Job Position: Maintenance Engineer (WiMAX and Microwave network).

Languages

- **Language 1:** Spanish, Native (Spoken and Written).
- **Language 2:** English, High Level (Spoken, and Written).