Universitat de Girona

MULTI-LAYER SURVIVABILITY: ROUTING SCHEMES FOR GMPLS-BASED NETWORKS

Anna URRA FÀBREGAS

ISBN: 978-84-690-4373-8
Dipòsit legal: GI-283-2007

Multi-Layer Survivability: Routing Schemes for GMPLS-based Networks

A Thesis
Presented to
The Academic Faculty

by

Anna Urra

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Department of Electronics, Computer Science and Automatic Control
Universitat de Girona
Girona, July 2006

Multi-Layer Survivability: Routing Schemes for GMPLS-based Networks

Approved by:

Date Approved: ______________________

To,

My mother Roser and my father Gonçal

for their patience and support for all these years. To,

My husband Apu

for his endless support, his friendship.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my deepest sense of gratitude to my advisor, Dr. Jose L. Marzo and co-advisor Dr. Eusebi Calle, for being my mentors in all respects of research. They have provided me with invaluable guidance and support, throughout the course of my Ph. D. studies.

I wish to convey my heartfelt thanks to Dr. Ian Akyildiz who acted not just as my guide but also as a well-wisher and shaped up my life. I will always be grateful to him. I would like to thank Dr. Caterina Scoglio and Dr. Tricha Anjali with whom I had the privilege of working with and learning a great deal, Dr. David Harle, who supported me immensely during my stay in Glasgow and devoted his valuable time to review my thesis.

I would like to thank Saud Albarrak and other colleagues at the University of Glasgow, my fellow researchers at the University of Girona and Georgia Tech for their support and friendship.

I would like to thank Departament d'Universitats, Recerca i Societat de la Informació (The Ministry Universities, Research and Information Society) (DURSI) Initiative of the Catalonia Government and European Social Funds for funding my research projects

All the successes in my life are the result of the values imbibed in me by my parents. I am immensely indebted to them who, in spite of all their hardships and sufferings have taught me the values of truth, endurance and patience and who always encouraged me to pursue my goals. I would not be what I am today, without the unfaltering faith and love of all of them. I would like to offer special thanks to my husband Dr. Apurva N. Mody for his end-less support, patience and kindness specially during the last phase of this research.

Finally, I would like to acknowledge the support and love of my family and especially Maria and Pere, for constantly sending me words of encouragement, and supporting me through thick and thin.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATM	Asynchronous Transfer Mode
BP	Backup Path
DWDM	Dense Wavelength Division Multiplexing
FIS	Fault Indication Signal
GMPLS	Generalized Multi-Protocol Label Switching
IETF	Internet Engineering Task Force
IP	Internet Protocol
IS-IS	Intermediate System to Intermediate System
ISP	Internet Services Providers
LFP	Link Failure Probability
LSC	Label Switch Capable
LSP	Label Switched Path
LSP_FP	Label Switched Path Failure Probability
LSR	Label Switching Router
MIRA	Minimum Interference Routing Algorithm
MPLS	Multi-Protocol Label Switching
MTBF	Mean Time Between Failures
MTTR	Mean Time To Repair
OSPF	Open Shortest Path First
OXC	Optical Cross Connect
PFP	Path Failure Probability
PSL	Path Switch LSR
PML	Path Merge LSR
PSC	Packet Switch Capable
QoS	Quality of Service
RFP	Residual Failure Probability

RSVP	Reservation Protocol
RSVP-TE	Reservation Protocol Traffic Engineering
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Network
SRLG	Shared Risk Link Group
TE	Traffic Engineering
WDM	Wavelength Division Multiplexing

SUMMARY

The use of optical technology in core networks combined with IP/Multi-Protocol Label Switching (MPLS) solution has been presented as a suitable choice for the next generation Internet architecture. The integration of both layers is facilitated by the development of Generalized MPLS (GMPLS). In this network architecture, a single fibre failure can result in potentially huge data losses as the effects propagate up and through the network causing disruptions in the service of many applications. Thus, survivability has become a key issue to improve and satisfy the increasing requirements of reliability and Quality of Service (QoS) of these applications. Fault recovery schemes have been adopted in the network in order to provide such survivability. These schemes are based on switching the traffic affected by the failure to an alternative path. The computation of the working and alternative path is a crucial step to offer the required QoS to the traffic services. Some relevant parameters, such as resource consumption and recovery time, could be affected negatively if suitable routing algorithms are not used. According to the timing of backup path computation, recovery mechanisms are classified in protection and restoration. Although restoration is flexible in terms of resource consumption, it offers low recovery time and the recovery action may not be successful because there are insufficient network resources. Protection describes recovery schemes that are pre-planned for both spare capacity and backup paths achieving the shortest recovery time and providing high availability against network failures. The accuracy and performance of QoS with Protection (QoSP) routing algorithms in terms of resource consumption depends on the available network information. The availability of full or partial network information influences the management of the network capacity. The reduction of the recovery time is another parameter to be considered for backup path selection and it is achieved by applying segment or local backup path methods instead of path protection.

Nowadays different existing QoSP routing algorithms are oriented towards offering the reliability required by the traffic services. However, they operate in a single switching layer: either optical and wavelength (lightpath) oriented or IP/MPLS and packet (Label

Switched Path, LSP) oriented. Thus, both optical and IP/MPLS layers independently deploy their own fault recovery methods. This results in protection duplications making fault management more difficult and poor resource utilisation. This research provides and evaluates new QoSP routing schemes that consider both IP/MPLS and optical network layers to compute the LSP and alternative LSP subject to the QoS requirements of the traffic. Two network scenarios are considered in this thesis: *static multi-layer network scenario* and *dynamic multi-layer network scenario*. In the *static multi-layer network scenario*, the logical topology where the LSPs are routed is pre-established. Some of the pre-established lightpaths are are assumed to be already protected at the optical layer. Hence, an enhancement of the QoSP routing algorithms for IP/MPLS networks is achieved by avoiding the protection of those lightpaths that are already protected at the optical layer. In order to deploy this proposal, a formalization of the path failure probability and a new definition of link-disjoint path based on Shared Risk Link Group (SRLG) concept are presented. As a novelty, the backup path is proposed to be a Partial Disjoint Path (PDP) since it may overlap the lightpaths of the working path that are already protected at the optical layer. In order to guarantee fast protection, the proposed algorithms also combine segment protection and shared backups, resulting in a suitable fault recovery time and resource consumption. A complete set of simulations verifies the efficiency of the proposed algorithms.

In the *dynamic multi-layer network scenario*, cooperation between each layer, optical and IP/MPLS, is considered. Although effort has been devoted in developing multi-layer routing algorithms that consider all switching layers, protection is not considered amongst them. Thus, in the proposed QoSP routing algorithms, whenever a new LSP request arrives, the decision of setting up new lightpaths, backup lightpaths and backup LSPs is made. Additionally, whenever a LSP is torn-down, the respective lightpaths and backup lightpaths that do not accommodate any other LSP are disconnected. New constraints are added to the network such as the number of Packet Switching Capable (PSC) ports of the routers (optical grooming switches). The presented QoSP routing algorithms are compared to other algorithms that consider either full optical protection or full IP/MPLS protection. The performance of the algorithms is analyzed according to the different metrics such as the number of PSC ports, the number of wavelengths per fibre and the number of hops (lightpaths). A complete set of simulations proves the efficiency of the proposed algorithms.

This thesis also presents reliability differentiation based on the traffic classification and hence the QoS requirements. Note that when a failure occurs not all the applications affected by the failure require the same level of reliability. Some applications are more stringent about their QoS requirements than others. Moreover, in many cases improving the fault recovery involves very expensive mechanisms in terms of resource consumption, which cannot be deployed throughout the whole network. Thus, new QoSP routing algorithms that take into account the presented traffic classification are presented and evaluated under the static and dynamic multi-layer network scenarios.

CHAPTER I

INTRODUCTION

The use of Wavelength Division Multiplexing (WDM) in optical core networks combined with IP/Multi-Protocol Label Switching (MPLS) for offering traffic-engineering capabilities has become a popular architecture amongst many Internet Service Providers (ISPs) [3]. Due to their high capacity and flexibility, WDM-based optical networks are an appropriate choice for the next-generation optical Internet networks to transport high-speed IP traffic. The aim of this first chapter is to highlight the role that survivability plays in such an architecture. The first section presents a brief overview of next-generation optical internet drivers and technologies. In particular, WDM, IP, MPLS and GMPLS concepts are described; keeping in mind the requirements of the next generation internet which will be supported by an optical fibre backbone. Next, the concept of network survivability is defined accompanied by a discussion about the causes of network impairments. The discussion then focusses upon single link failures; identified as the most common form of network failure. Such failures can result in potentially huge data losses as the effects propagate up and through the network. Most networks deploy a multi-layer architecture and the effects of failure in these networks can be profound; single link failures may result in multiple failures in the upper layers. A key element in identifying, quantifying and controlling the effects of such failures is the concept of a Shared Risk Link Group (SRLG). These SRLG form the basis of many of the ideas elaborated in this thesis and, thus, this chapter concludes with a discussion of this important component.

1.1 Next Generation Optical Internet

In the early 1980s, a revolution in telecommunications networks began with the use of fibre-optic cable. Since then, high cost savings and increased network quality led to many advances in the technologies required for optical networks. However, as optical-fibre deployment increased, no standards existed to mandate how network elements should format the optical signal. The need for optical standards led to the creation of the Synchronous Optical NETwork/Synchronous Digital Hierarchy (SONET/SDH). SONET/SDH provides a guaranteed level of performance and reliability for voice calls and leased lines, the predominant traffic types prior to 1995. Since 1995, however, there has been a dramatic increase in data traffic, primarily due to the explosive growth of the Internet. Additionally, due to the demand of more services and different types of data traffic, Internet Service Providers (ISPs) aim to carry a large volume of traffic in a cost-effective manner. To provide full end-to-end connectivity, a new paradigm was necessary in order to meet the high-capacity required and satisfy different needs such as different traffic types. Optical networks provide the required bandwidth and flexibility to enable end-to-end wavelength services [4]. Wavelength Division Multiplexing (WDM), ushered in a new era in optical networks as they created additional capacity on existing fibres. The SONET/SDH, defined network elements and architectures provide the basis of these WDM optical networks. However, unlike SONET/SDH, instead of using a predefined bit-rate and frame structure, a WDM-based optical network relies on individual wavelengths. The components of the optical networks are defined according to how wavelengths are transmitted, groomed, or implemented in the network.

1.1.1 Multi-layer Architecture

An optical network when viewed using a layered approach, requires the addition of an optical plane or layer. To help define network functionality, networks are divided into several different physical or virtual layers:

- *Internet Protocol* (IP) for carrying applications and services,
- *Asynchronous Transfer Mode* (ATM) traffic engineering,
- *Synchronous Optical Network/Synchronous Digital Hierarchy* (SONET/SDH) for transport,
- *Wavelength-Division Multiplexing* (WDM) for physical layer in order to realize the capacity in the optical networks.

WDM-based optical networks are becoming an appropriate choice for the next-generation optical Internet networks to transport high-speed IP traffic. A key advantage of WDM

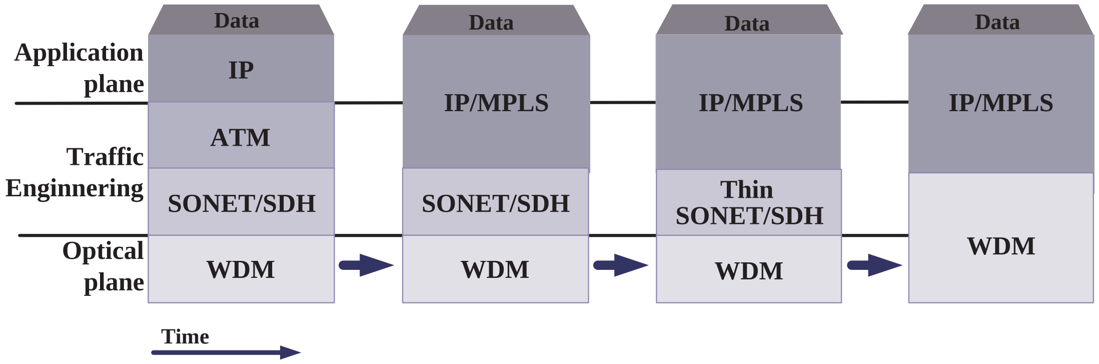


Figure 1: Mutli-layer network architecture evolution.

is that it offers multi-protocol support allowing multiple independent networks protocols to coexist on the same fibre [5]. SONET/SDH and ATM networks have been widely deployed in the transport networks. SONET/SDH systems have several attractive features such as high-speed transmission and network survivability. ATM networks have several attractive features such as flexible bandwidth allocation and Quality of Service (QoS) support. Therefore, ATM and/or SONET/SDH layers can be used between the IP layer and the WDM optical layer for transporting IP packets. A major drawback of this multi-layer approach is that it suffers from increased control and management overhead [6]. The higher data rates associated with direct optical transport offer the potential for bypassing the SONET/SDH and ATM layers. In order to do this, their associated functions must migrate to the WDM optical layer and, in particular, Optical Cross-Connects (OXC). Additionally, over the last few years, under the umbrella of Multi-Protocol Label Switching (MPLS) [7], IP routing has evolved to include new functionality. Ultimately, two layers are considered in the optical network architecture: IP/MPLS and WDM. Figure 1 illustrates the trend towards reducing layers associated with next generation optical Internet. Current research is focused on the designing and implementation of all-optical packet-switched networks [8]. In a long-term scenario, the optical packet switching (OPS) can provide a simple transport platform based on a direct IP over WDM structure which can offer high capacity efficiency, flexibility, and fine granularity [8].

Some recent work has extended MPLS as a control plane that can be also used with new optical technology: Generalized MPLS (GMPLS) [9]. GMPLS relies on a peer model in which all network elements share the same unified control and signaling plane

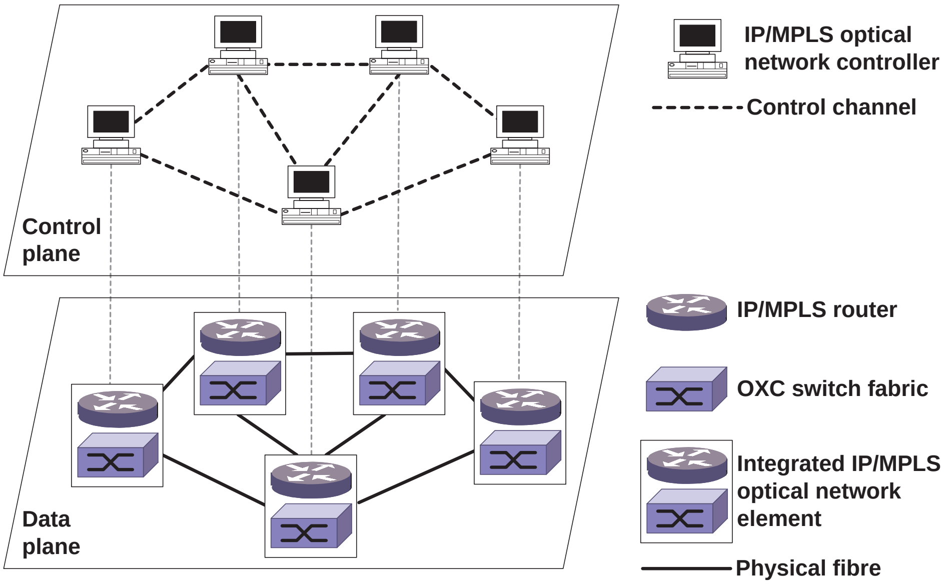


Figure 2: GMPLS-based optical network model [1].

that allows the operation between the IP/MPLS network with the optical network. A common control plane simplifies operations and management, which further reduces the operational costs [6]. GMPLS allows inter-layer communication and coordination. This means that the IP/MPLS layer can communicate with the optical layer to exploit about the underlying fibre topology, lightpath connectivity, as well as protection capability in the optical layer. The GMPLS-based optical network model considered in this research is illustrated in Fig. 2 [1]. Following a brief overview of the drivers, WDM, IP/MPLS and GMPLS concepts are described and placed into context.

1.1.2 Wavelength Division Multiplexing

The use of Wavelength Division Multiplexing (WDM) technology aids support for the rapidly growing demand for data traffic by taking advantage of the huge bandwidth offered by optical fibres. WDM allows multiple optical signals, operating at different wavelengths, to be multiplexed onto a single optical fibre and transported in parallel through the fibre [10]; see Fig. 3. Significant advances in optical component technologies have brought more advanced WDM network elements such as Add-Drop Multiplexer (ADM) and Optical-Cross-Connect (OXC) [5,11]. These devices can selectively process (route, add or drop) different wavelengths. The ADM can add/drop necessary traffic (wavelengths). In many networks, it is necessary to drop some traffic at intermediate points along the route. When a wavelength is ”dropped”, further Internet traffic can be

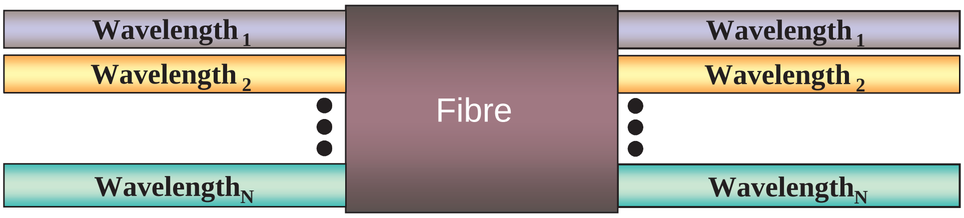


Figure 3: Optical fibre transporting N wavelengths in parallel.

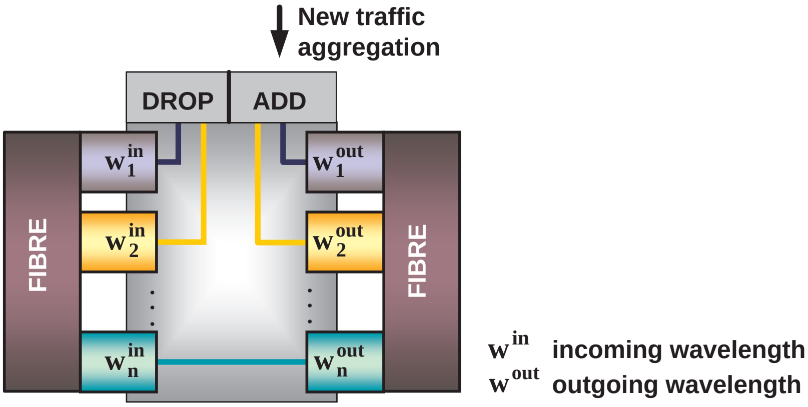


Figure 4: Add-Drop Multiplexer.

aggregated as illustrated in Fig. 4. The Optical Cross-Connect (OXC) is the node responsible for switching/propagating the incoming wavelengths to the respective outgoing wavelengths [10]. The OXC propagates the wavelength from a fibre to another changing or without changing its wavelength. When the OXC cannot assign the incoming wavelength to a different outgoing wavelength, the OXC has wavelength continuity constraint, see Fig. 5a. Otherwise, the OXC provides wavelength conversion; the incoming traffic on one wavelength can be assigned to an outgoing port using a different wavelength (see Fig. 5b). Optical capacity allocation can be enhanced if wavelength conversion is used [12,13]. Although attractive because of the enhanced wavelength utilisation, wavelength conversion is complex in terms of operations and has significant costs associated with it. Therefore, methods to reduce or limit the number of wavelength-conversions or wavelength-conversion nodes are subject to research effort [13, 14]. Moreover, the OXC may utilize optical-electrical conversion at the input port and electrical-optical conversion at the output port, or it may be all-optical. The Optical-to-Electronic-to-Optical (OEO) conversion is currently used because most networking equipment is still electronics based. The optical signals must be converted into electrical to be amplified,

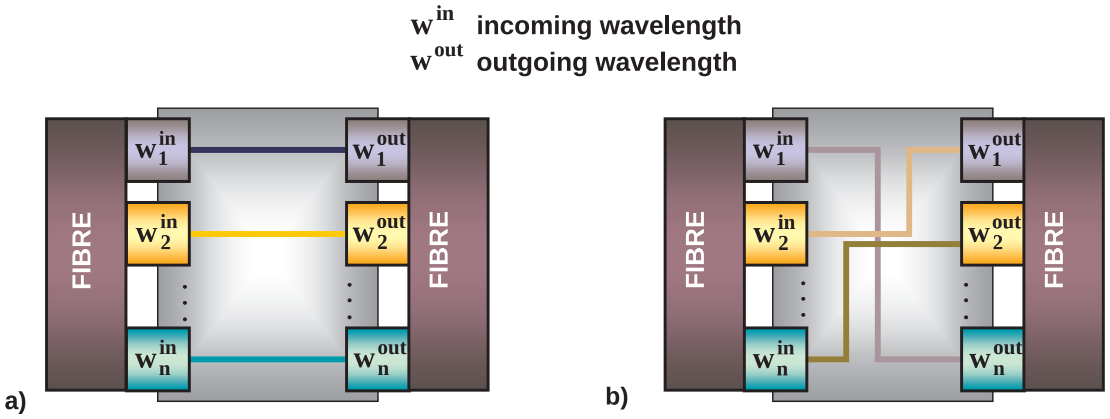


Figure 5: OXC a) with wavelength continuity b) with wavelength conversion.

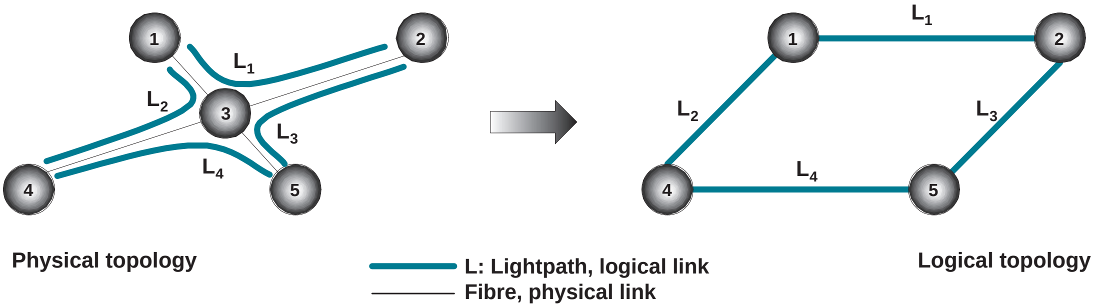
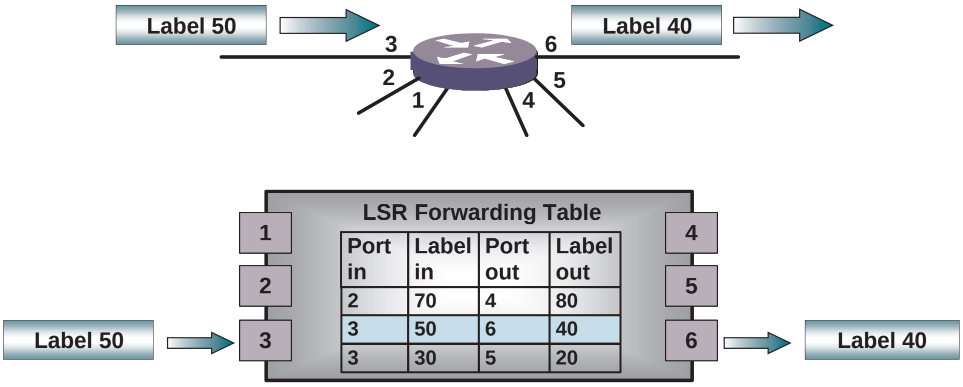


Figure 6: Logical topology example.

regenerated, stored or switched, and then reconverted to optical signals. All optical transmission is faster than its electronic counterpart. Thus, a significant bottleneck in transmission occurs when OEO conversion is used.

Such reconfigurable WDM network elements provide on-demand establishment of high-bandwidth connections, called lightpaths. A lightpath [13] is a wavelength or set of wavelengths that interconnects a node pair over a WDM network. The first case, a single wavelength, occurs when OXCs have wavelength continuity constraint. Such constraint reduces the possibilities of finding a future lightpath between a node pair because specific wavelengths must be available on one end-to-end basis. It is not sufficient just to have free wavelengths. Once a lightpath is established, its entire capacity is allocated to the connection and cannot be used by any other connection. The lightpaths define the logical topology, or virtual topology, where lightpaths now represent direct links between nodes. An example is shown in Fig. 6, where the physical network has two wavelengths per fibre. In this example, four lightpaths (L) are established, defining a logical ring topology.



Port in	Label in	Port out	Label out
2	70	4	80
3	50	6	40
3	30	5	20

Figure 7: LSR forwarding table example.

1.1.3 IP/Multi-Protocol Label Switching

IP has become the predominant network layer protocol in use today. IP is based on best-effort delivery; Internet services cannot support traffic that requires a combination of high capacity and high QoS transport [5]. Moreover, complex and time-consuming route lookups and address matching schemes are used to determine the next hop for a received IP packet; primarily by examining the destination address in the header of the IP packet [15]. Hence, each router makes an autonomous decision about how to forward an IP packet, and forwarding proceeds in a connectionless way at every hop. In order to overcome these issues, Multi-Protocol Label Switching (MPLS) [16,17] introduces a new forwarding paradigm for IP networks. MPLS has simplified routing by adding a short fixed length label to IP packets and any forwarding decisions are based on this label. Variable length IP packets arriving at a Label Switching Router (LSR) are encapsulated with labels when the IP packets first arrive to an MPLS routing domain. The LSR Edge Router (LER) looks at the information in the IP header and assigns an appropriate label. This label selection can be based on QoS and routing considerations, not just on the destination address in the IP header. Next, all the subsequent routers in the MPLS domain will forward the packet based on the label, instead of the IP header. When the packet leaves the MPLS domain, the LER removes the label. The value of the label usually changes at each LSR in the path according to the forwarding tables. Figure 7 shows an example where a packet with the label 50 arrives using the incoming port 3. At the LSR, the label is removed and a new label, with a value of 40, is inserted and then forwarded in the outgoing port 6. A connection setup protocol creates the sequence of *out-port, next-label* entries along a desired path so that a logical circuit is established between any source/ingress and destination/egress node, called Label Switched Paths (LSP). In this way, MPLS introduces the notion of connection-oriented forwarding in an IP network. The elements of the MPLS domain are shown in Fig. 8. The label stacking feature of MPLS also allows aggregation of many small granularity LSPs into one high capacity LSP over a common segment of their routings. This is achieved by adding a new label that switches the flow into an established LSP. Figure 9 shows an example of label stacking.

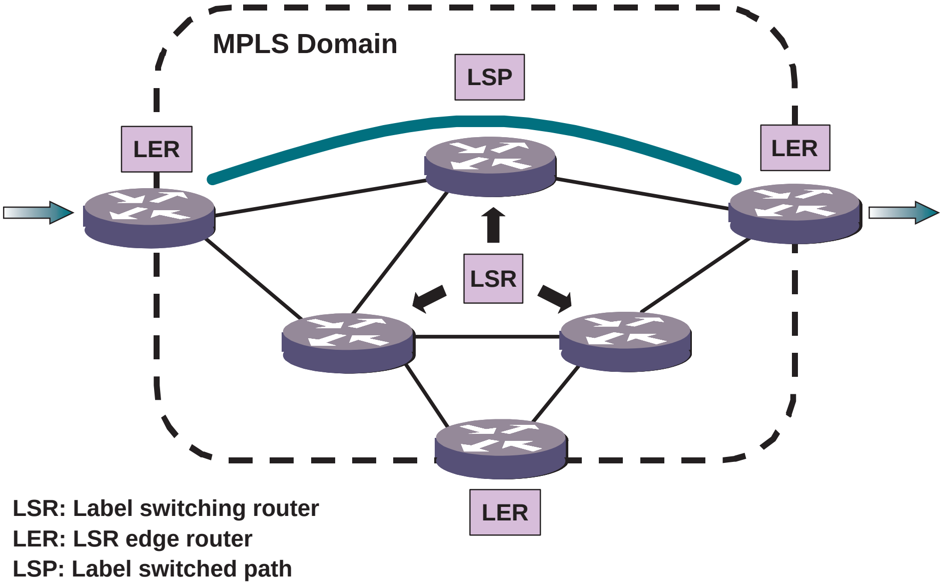


Figure 8: MPLS domain.

MPLS further complements IP technology by introducing a new set of MPLS control procedures partially based on existing IP routing protocols with extensions. When a LSR swaps a label of an incoming packet and forwards it to its downstream LSR, a method is necessary in order to know what label value its downstream LSR is expecting. Several signaling protocols can be used for the distribution of labels between LSRs (LSP creation):

- *Label Distribution Protocol* (LDP) [18]. The LDP defines a set of procedures and messages allowing LSRs to establish LSPs through a network by mapping network layer routing information directly to data-link layer switched paths. The LSR uses this protocol to establish LSPs throughout the network. The drawback of the LDP is that the LDP by itself cannot meet QoS needs.

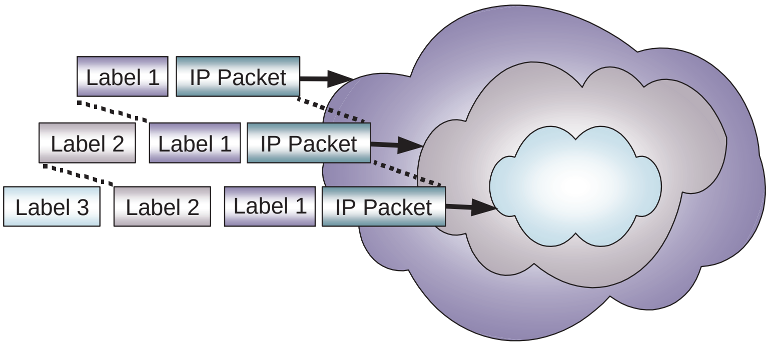


Figure 9: Label stacking.

- *Constraint-based Routed-Label Distribution Protocol* (CR-LDP) [19]. The CR-LDP extends the capabilities of LDP, such as setting up paths beyond what is available for the routing protocol. Constraint-based routing is used to meet Traffic Engineering requirements. In the CR-LDP, an LSP can be established based on explicit route constraints, QoS constraints, etc.

- *Reservation Protocol Traffic Extension* (RSVP-TE) [20]. The RSVP-TE protocol is an addition to the RSVP [21] for establishing LSPs. Like CR-LDP, RSVP-TE also establishes point-to-point LSPs that meet QoS requirements. It supports the instantiation of explicitly routed LSPs without resource reservations. RSVP-TE also supports rerouting, preemption and loop detection. RSVP-TE includes the ability to control all optical networks. RSVP-TE includes the ability to signal optical wavelengths and Shared Risk Link Groups (SRLGs), as well as capacity and other link characteristics.

Applications of MPLS include Traffic Engineering (TE) and Quality of Service (QoS) for different types of services, among others. More detailed information relating to MPLS can be found in [7].

1.1.4 Generalized Multi-Protocol Label Switching

The integration of the IP/MPLS and WDM is facilitated by the development of Generalized MPLS (GMPLS) [9, 22]. GMPLS differs from traditional MPLS since it supports multiple types of switching, i.e. supporting Time Division Multiplexing (TDM), lambda, and fibre (port) switching. Unlike MPLS, the GMPLS architecture requires the control plane and the data plane to be separated. Thus, existing signaling and routing protocols used under the MPLS framework must be modified to make them suitable for non-packet networks. These are being standardized by the Internet Engineering Task Force (IETF), and the result of this process can be summarized as follows [6]:

- Establishment of a new Link Management Protocol (LMP) [23] designed to address issues related to link management in optical networks.
- Enhancements to the Open Shortest Path First/Intermediate System to Intermediate System (OSPF/IS-IS) routing protocols [24] to advertise availability of optical resources in the network, e.g. generalized representation of various link types, bandwidth on wavelengths, link protection type, fibre identifiers.
- Enhancements to the Resource Reservation Protocol (RSVP)/Constraint-Based Routing Label-Distributed Protocol (CR-LDP) signaling protocols [20] for traffic engineering purposes that allow a LSP to be explicitly specified across the optical core.
- Scalability enhancements such as hierarchical LSP formation, link bundling, and unnumbered links.

GMPLS relies on a peer model in which all network elements share the same unified control and signaling plane providing efficient management and use of the network resources. The topology perceived by the network nodes is the one where physical fibre links and logical links (lightpaths) coexist [25]. The peer model supports dynamic routing that can either use only the existing lightpaths or create more lightpaths when it is considered necessary. Suppose the path computation is triggered by the need to route a new LSP in a GMPLS environment. In this case, the signaling protocol will establish a lightpath between two edge routers. This lightpath is, in essence, a tunnel across the optical network and may have capacity that is much larger than the capacity required to support the first LSP. Thus, it is essential that other routers in the network realize the availability of excess capacity within the lightpath so that subsequent LSPs between the routers can use it rather than instantiating a new lightpath. The lightpath may therefore be advertised as a virtual link in the topology in order to address this

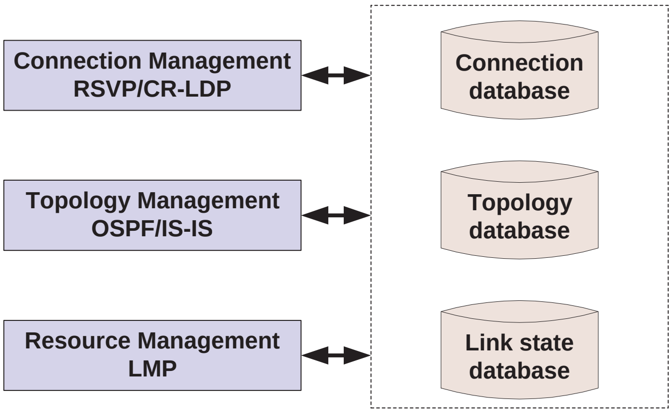


Figure 10: GMPLS controller.

issue. The GMPLS unified control plane creates an opportunity to share topology and resource information across multiple network layers. Each control unit consists of three main functional modules: connection management, topology management, and resource management, as shown in Fig. 10. The connection management module takes care of connection establishment, connection teardown, and connection recovery during failures. In terms of GMPLS implementation, RSVP or CR-LDP signaling protocol could be mapped to this module. The topology management module relies on a routing protocol such as OSPF or IS-IS to perform route discovery and neighbor discovery. Lastly the resource management module monitors the link status and performs resource discovery. The Link Management Protocol (LMP) is suitable for these tasks. Furthermore, three databases are managed at each controller: the connection database, the topology database and the link state database.

It can be concluded that GMPLS enables interoperability between network layers by providing an abstraction of the end-to-end connectivity. GMPLS also allows inter-layer communication and coordination. In the context of IP/MPLS over optical networks, this means that the IP/MPLS layer can communicate with the optical layer to learn about the underlying fiber topology, lightpath connectivity, as well as protection capability in the optical layer. Because GMPLS allows inter-layer coordination, it makes possible the integrated design of survivable networks and promises an efficient and cost-effective way to provision new connections.

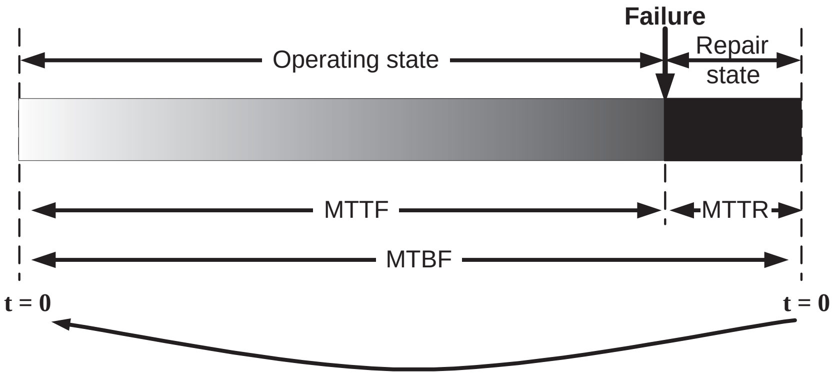


Figure 11: Failure cycle of a repairable system.

1.2 Network Survivability

Survivability is defined by Grover [26] as *the ability of a network to continue to provide service in the event of a failure that might arise.* The main goal of the survivability is to guarantee an acceptable level of reliability during network failures. Next-generation optical backbone networks, as explained in section 1.1, enable increasingly higher volumes of information to be transported. Ensuring a particular level of reliability in this scenario is becoming crucial since a fault results on a large volume of data losses. This section presents the life cycle of network elements and an overview of the most common network failures.

1.2.1 The Life Cycle of a Network Element

Elements of a network, such as link or nodes, follow a succession of repetitive cycles as shown in Fig. 11. Each generally starts at the operating state at time $t = 0$. When a failure occurs, the network element enters the repair state. Once the failure has been repaired, the network again progresses into the operating state, i.e. $t = 0$. The mean time expected before the first failure is the Mean Time To Failure (MTTF) and it corresponds to the operating state. The Mean Time To Repair (MTTR) is the average time spent performing all corrective maintenance repairs (ITU-T E800/4260) [27]. Finally, the Mean Time Between Failures (MTBF) is the MTTF including the time of repair following the last failure, i.e. MTBF = MTTF + MTTR. Data [2] showed that in the network system scenario, the MTTR of physical cable failures is in the range of hours to several weeks. In one example [2], the MTTR was equal to 14 hours with a high variance. For that reason, recovery mechanisms are needed in order to guarantee that most of the affected traffic is diverted appropriately and reaches its destination node in an appropriate time; ranging from milliseconds to minutes.

1.2.2 Network Failures

Network failures can occur as the result of natural disasters (flooding, hurricanes), or as the result of human action (war, terrorism, digging activities) or even by unintentional failures in software or control systems. Bhandari [28] classified the major network failures as follows:

- *Node failure*: due to equipment breakdown or equipment damage resulting from an event such as an accidental fire, flood, or earthquake; as a result, all or some of the communication links terminating on the affected node may fail.
- *Link failure*: due to an inadvertent fibre cable cut. The fibre cable carrying traffic from one telecommunication office to another is buried approximately 3 feet underground in a conduit. Due to ubiquitous construction activity as world economies grow rapidly, accidental fibre cuts occur frequently, despite increased network care and maintenance efforts.
- *Software failure*: this type of failure can impact a large portion of the given network, and is, in general, hard to identify and recover from.

1.2.3 Single Failure Classification

Link and node failures, i.e. cable cuts and equipment failures respectively, typically represent the most common failures. Most operators consider two network scenarios [29]:

- *Single-link failure*: when a link between two adjacent nodes fails. As a consequence, no direct information exchange between these two nodes is possible until the fault is repaired (see Fig. 12).
- *Single-node failure*: when a node element fails. The failure of a single node automatically takes all attached links out of service.

The importance of single failures is based on two assumptions [29]:

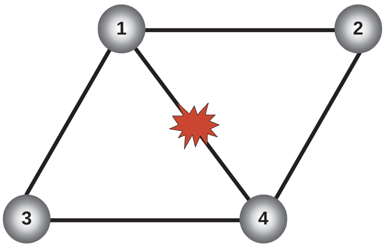


Figure 12: Single-link failure scenario.

- The failure of a link or node in the network is statistically independent of the failure of another link or node in the network, in most cases.
- If the network scale is not too large, the Mean Time To Repair (MTTR) for a single-link or single-node failure is usually much shorter than the Mean Time Between Failures (MTBF). Thereby, the probability that more than one link or node fails at the same time is very low.

This thesis focuses on single link failures, also referred through the document as fibre failures, cable cuts and cable failures. Node failures are beyond the scope of this research. A more detailed analysis of the causes of the fibre optic cable failures was reported by Crawford [2] and it is shown in Fig. 13. From this report, all 160 of the cable failures were single-failure events. Moreover, a pan-European "carrier's carrier" has independently estimated an average of one cable cut every four days occurs within their network [26].

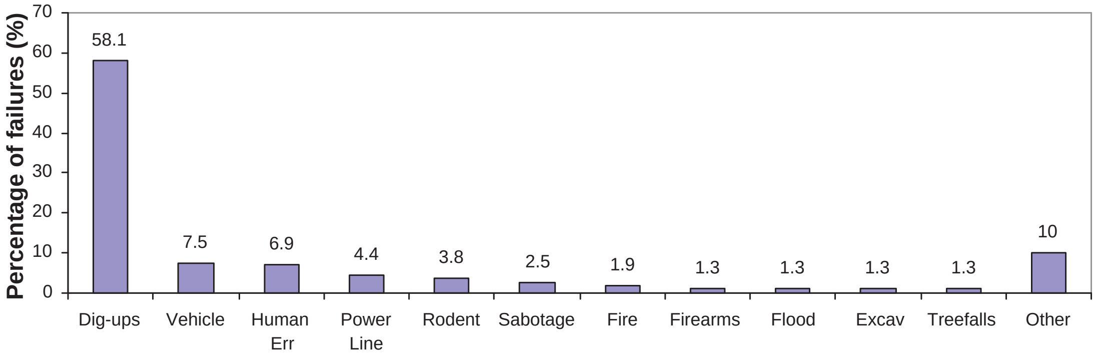


Figure 13: Immediate cause of failures for 160 fibre optic cable cuts [2].

1.3 Multi-layer Network Survivability

In the previous section, the network failures were described and single link failures, i.e. cable cuts, were identified as one of the most common types of failures. However, as shown in section 1.1, next generation networks should have two layers: the physical topology where the lightpaths will be set up, and the logical topology where the LSPs will be set up. In this multi-layer scenario, a cable cut in the lower network layer leads to multiple link failures in the upper network layer.

1.3.1 Multiple Link Failures

In the IP/MPLS over optical network scenario, whenever a single link failure occurs, all the lightpaths, carried by this link, are broken. This is shown in Fig. 14, where the lightpaths L_1 and L_4 fail due to the failure of the physical link 2-5. As a result, multiple failures are detected at the upper layer, since multiple lightpaths are broken, i.e. the logical links of the upper layer. Following the example showed in Fig. 14, a cable cut from node 5 to node 2 provokes two link failures in the logical layer: the logical links 1-5 and 2-5, i.e. the lightpaths L_1 and L_4, respectively.

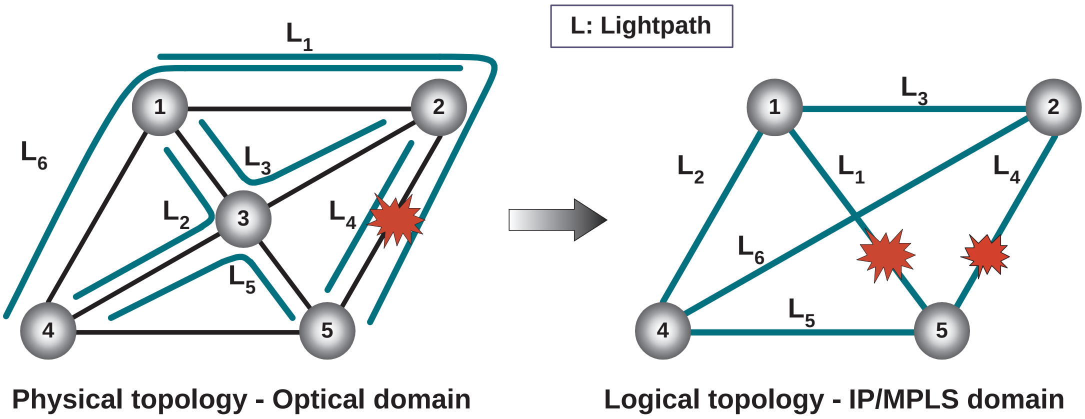

L: Lightpath
L1
L6
L3
L2
L4
L5
1
2
3
4
5
Physical topology - Optical domain
L3
L2
L1
L4
L6
L5
1
2
4
5
Logical topology - IP/MPLS domain

Figure 14: Mapping a single link failure in the physical and logical topology.

1.3.2 Shared Risk Link Group

As shown above, a failure may affect different elements of the network. The Shared Risk Link Group (SRLG) [30] concept has been developed to classify the network elements that may be affected by the same failure. For instance, a single fibre may be one SRLG, indicating that all lightpaths routed by this fibre will be affected if the fibre fails. In Fig. 14, if single fibre failures are only considered, eight SRLGs are identified. Table 1 shows these eight SRLGs and their associated elements, i.e. logical links. Note that:

- One SRLG may have more than one logical link. In this case, the upper layer will perceive multiple failures. For instance, in Table 1, the failure risk of the single fibre 2-5, $f_{2,5}$, is shared with the logical links L_1 and L_4. Thus, the failure of this fibre will result in the failure of both logical links.

- One logical link belongs, at least, to one SRLG. For instance, the logical links L_1, L_2, L_3, L_5 and L_6 in Table 1 belong to two SRLGs. It is important to highlight that the more SRLGs to which a logical link belongs, the more often it may fail.

Table 1: SRLG example based on Fig. 14 and only considering single fibre failures.

Logical link		**Shared Risk Link Groups: Single fibre (f)**								**Total number of SRLG per lightpath**
ID	*Lightpath*	$f_{1,2}$	$f_{1,3}$	$f_{1,4}$	$f_{2,3}$	$f_{2,5}$	$f_{3,4}$	$f_{3,5}$	$f_{4,5}$	
L_1	1-2-5	●				●				2
L_2	1-3-4		●				●			2
L_3	1-3-2		●		●					2
L_4	2-5					●				1
L_5	4-3-5						●	●		2
L_6	4-1-2	●		●						2
Total number of lightpaths per SRLG		2	2	1	1	2	2	1	0	

1.4 Motivation and Layout of the Thesis

The use of WDM-based optical network technology in core network combined with MPLS for offering traffic-engineering capabilities has been selected as a suitable choice by many Internet Service Providers (ISPs). In particular, GMPLS offers the tools for traffic engineering, constraint-based routing and many other features required by future Internet applications. Many of these applications require high reliability and QoS guarantees from the network. However, single fibre failures occur frequently and lead to multiple failures in the upper layers. The amount of time taken to repair them can be significant, causing disruptions in the service of affected applications. Such networks are required to support an increasing number of heterogeneous applications with diverse service requirements. Moreover, some applications cannot tolerate these disruptions. Survivability is becoming crucial in this network scenario for ensuring the level of reliability required by the traffic. Recovery mechanisms are needed in order to guarantee that the affected traffic reaches the destination node in an appropriate time. As will be described in Chapter 2, recovery mechanisms compute an alternative path where any traffic affected by a failure may be switched. The selection of the working and alternative paths is a crucial step to offer the required QoS of the traffic services. Some parameters, such as recovery time, could be affected negatively if no suitable routing algorithms are used. Although different QoS routing algorithms exist in the literature, reviewed in Chapter 3, they only consider one network layer.

The contribution of this research is to provide and analyze new QoS with protection (QoSP) routing algorithms that consider both IP/MPLS and optical network layers to compute the working and alternative paths. This thesis is divided in two phases:

Phase 1: Static multi-layer network scenario (Chapter 4). In this phase, the logical topology where the LSPs are routed is given. Some of the lightpaths are assumed to be already protected at the lower layer. An enhancement of the QoSP routing algorithms for IP/MPLS-based networks is achieved by avoiding the protection of those lightpaths that are protected at the optical layer. A new definition for link-disjoint path based on SRLG is made in order to facilitate greater sharing of spare capacity and, consequently, minimize resource consumption.

Phase 2: Dynamic multi-layer network scenario (Chapter 5 and 6). The schemes proposed in Phase 1 result in efficient resource consumption and recovery time under the assumption of a partial protected optical layer. However, no study has been conducted that relates the relative impact; typically because the logical topology is given. In this second part of the thesis, the logical topology will be dynamically set-up. This phase provides 1) an explanation of why the recovery mechanisms cannot be only applied at optical layer, 2) an analysis of the advantages and disadvantages of the recovery mechanisms at each network layer and, finally, 3) new QoSP routing proposals for dynamic multi-layer routing that consider a cooperation between IP/MPLS and optical layers.

1.5 Concluding Remarks

This chapter is intended to lay out a basic groundwork for understanding the importance of survivability on IP/MPLS over optical networks against single link failure. After explaining the principles of each network layer, a brief overview of network survivability was given. The main points to highlight are:

- Survivability is defined as the ability of a network to maintain an acceptable level of reliability during network failures.

- Single link failures are one of the most common network failures.
- The mean time to repair (MTTR) of a single link failure is in the range of hours to several weeks.
- A single link failure can result in a loss of several terabits of data per second in the IP/MPLS over optical network scenario.
- A single link failure leads to multiple failures in the upper network layer.
- Shared Risk Link Group (SRLG) should be used to classify the logical links that may be affected by the same single link failure in a multi-layer network scenario.

Hence, single link failures occur frequently and the amount of time taken to repair them can be significant, causing disruptions in the service of affected applications. Heterogeneous and wide number of applications are transported over the network. Moreover, some applications cannot tolerate these disruptions. Thereby, recovery mechanisms are needed in order to guarantee that the affected traffic reaches the destination node without a degradation of the QoS required by the applications.

CHAPTER II

HOW TO PROVIDE NETWORK SURVIVABILITY?

The use of optical technology has enabled operators to meet the rapidly growing demand for data traffic by taking advantage of the huge capacity that optical fibres can carry. Due to their high capacity and flexibility, optical networks are the right choice for the next-generation optical Internet networks to transport high-speed IP/MPLS traffic. In particular, GMPLS offers the tools for traffic engineering, constraint-based routing and many other features required by future Internet applications. Many of these applications require high reliability and Quality of Service (QoS) guarantees from the network. Due to the huge amount of data losses when a failure occurs, survivability has become an important issue in such an architecture. This chapter describes the basic principles concerning survivability. The first section presents an overview of the two survivability techniques: protection and restoration. Both of them require the computation of an alternative path to where the traffic is switched whenever a failure occurs. The discussion then focusses upon relevant parameters that need to be considered on the path computation. These parameters are subject to the Quality of Resilience (QoR) of the traffic such as delay, packet loss and reliability. Thus, some concepts such as fault recovery time and path failure probability are described and formalised. Finally, a traffic classification is made according to the QoR requirements of the traffic. Note that, when a failure occurs, not all the applications affected by the failure require the same QoR. Thus, for each traffic class, the most suitable survivability technique is proposed.

2.1 How to deal against disruptions

Survivability is the capability of a network to maintain service continuity in the presence of faults within the network. Since network failures, e.g. fibre cuts, cannot be easily avoided, survivability techniques are defined in order to switch the traffic affected by the failure from the working path to an alternative/backup path. A first classification of the survivability techniques is derived from the point that the backup path is established [31]:

- *Restoration.* Upon network failure, backup paths are established on demand and the spare capacity is dynamically allocated. Once the backup path has been set up, traffic is then switched.

- *Protection.* The backup path is pre-established before any failure and spare capacity is reserved at the same time that the request is set up. The selection of the backup path is based on the requirements of the traffic.

In case of failure, protection mechanisms are faster than restoration in recovering the traffic, since these do not need to wait for the establishment and reservation of the alternative path. Moreover, protection mechanisms guarantee 100% availability. However, protection mechanisms require more network resources, since there is a need to pre-allocate spare capacity for pre-establishing backup paths. When providing survivability there are many factors involved, the most important being: resource utilisation, request blocking ratio, recovery time and recovery granularity. The major goal is to achieve maximum survivability with minimum recovery time while maintaining efficient resource utilisation.

2.1.1 Network Recovery Components

When a network failure occurs, traffic affected by the fault should be switched to an alternative path. Recovery methods begin with fault identification and end with link recovery. This process involves various network management components [32]:

1. A method for selecting the working and backup paths: routing algorithms.

2. A method for signaling the setup of the working and backup path is required: CR-LDP or RSVP-TE (see section 1.1.3).

3. Mechanisms for fault detection and notification. These convey information about the occurrence of a fault to the network entity responsible for taking the appropriate corrective action. This can be done by transmitting a Fault Indication Signal (FIS).

4. A switch-over mechanism to move traffic from the working path to the backup path.

To provide explicit protection features, two new type of nodes are necessary: a node responsible for the switch over function once the failure is identified and a node where the working and backup paths are merged. They are defined as Path Source LSR (PSL) and Path Merge LSR (PML) [31] respectively, or the Bridge and Selector nodes in the GMPLS proposal.

2.1.2 Protection Backup Path methods

When considering protection, there are a range of different methods to determine an appropriate backup path [33, 34]. In this section, those most commonly cited in the literature are described.

2.1.2.1 Global backup path method

Here, the source node is the responsible for path recovery when the FIS arrives and an alternative, unconnected backup path for each working path is required. The protection process starts at the source node, irrespective of the location of the failure in the working path. The advantage of this method is that only one backup path per working path is required. Furthermore, it is a centralized protection method, which means only one LSR has to be provided with PSL/Bridge functions. On the other hand, this method has a high cost in terms of recovery time since the FIS is sent to the source node. Furthermore, it implies higher packet losses during the switchover time. Figure 15 illustrates the phases involved to protect the working path using the global backup path method. Only node 1 needs PSL functions, and node 5 needs PML functions. Note that the worse case is when the last link of the working path fails, as shown in the example, because the FIS is sent to the source node resulting in the longest recovery time.

2.1.2.2 Local backup path method

With the local backup path method, the node that detects the failure is the responsible for switching the traffic to the backup path. Hence, the restoration begins closer to the fault offering faster recovery time as well as a significant reduction in the packet loss. However, every node (except the destination node) has to be provided with switchover functions (PSL/Bridge). A PML/Selector also needs to be provided for each node except the source node. Another drawback is the maintenance and creation of multiple backups:

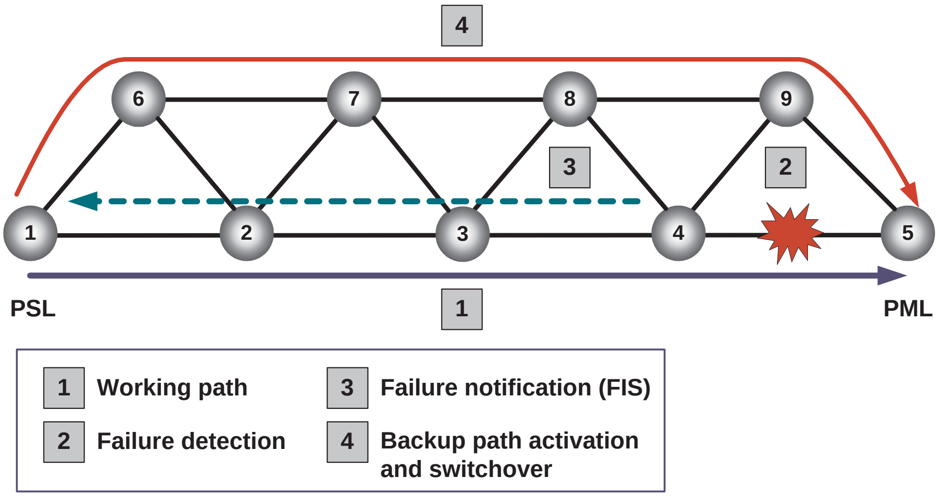

4
6
7
8
9
3
2
1
2
3
4
5
PSL
1
PML
1 Working path
2 Failure detection
3 Failure notification (FIS)
4 Backup path activation and switchover

Figure 15: Global backup path.

one per link. This can lead to low resource utilisation and increased complexity as shown in Fig. 16. For each link a local backup path should be computed. Thus, when a failure occurs, for instance the link 4-5 in the diagram, the upstream node of the affected link (node 4) detects and switches the traffic to the downstream node (node 5). In order to protect whole the path, each intermediate node of the working path must be provided with PSL and PML functions.

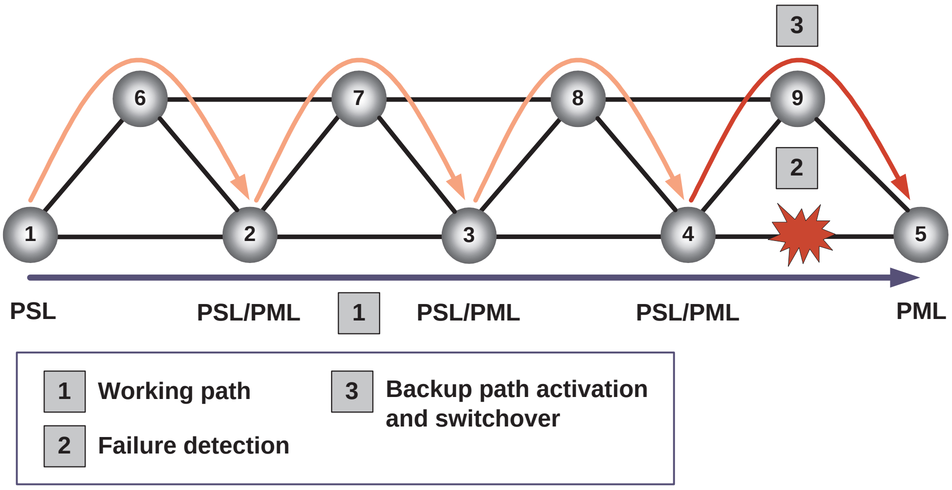


Figure 16: Local backup path.

2.1.2.3 Segment backup path method

An intermediate solution between local and global methods establishes segments to protect the working path. With this method, the number of PSL and PML is lower than the local protection case and it offers faster recovery time than global protection.

Global and local protection may be seen as an extreme case of segment protection. Figure 17 illustrates an example of the segment backup path method where the working path is protected by two segment backup paths.

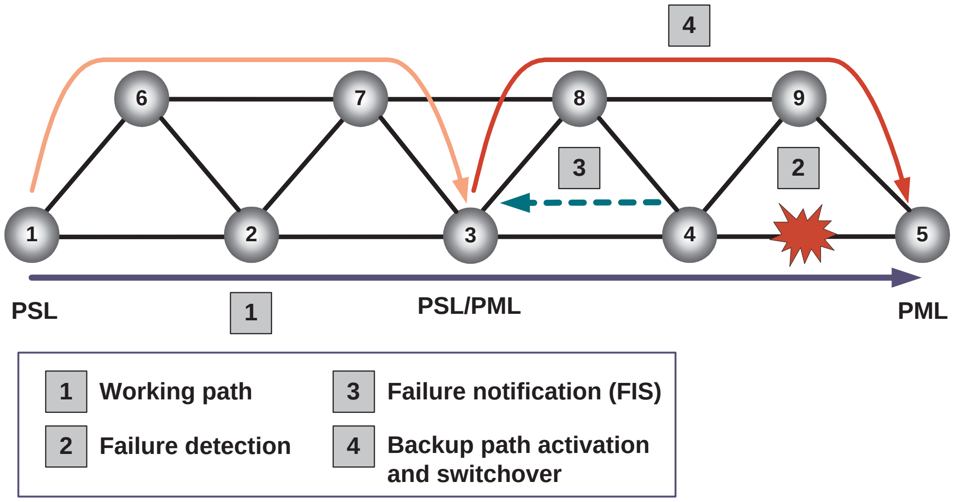

4
6
7
8
9
3
2
1
2
3
4
5
PSL
1
PSL/PML
PML
1 Working path
2 Failure detection
3 Failure notification (FIS)
4 Backup path activation and switchover

Figure 17: Segment backup path.

2.1.2.4 Reverse backup path method

The main feature of this method is to reverse the traffic affected by the failure, back to the source of the working path through a reverse backup path (see Fig. 18). As soon as the failure is detected, the LSR that detects the failure reroutes incoming traffic to the backup path sending it back in the opposite direction, to the source node. This method, like the local repair method, is particularly useful against the loss of sensitive traffic. Another advantage is the simplified fault indication, since the reverse backup transmits the FIS to the source node and the recovery traffic path at the same time.

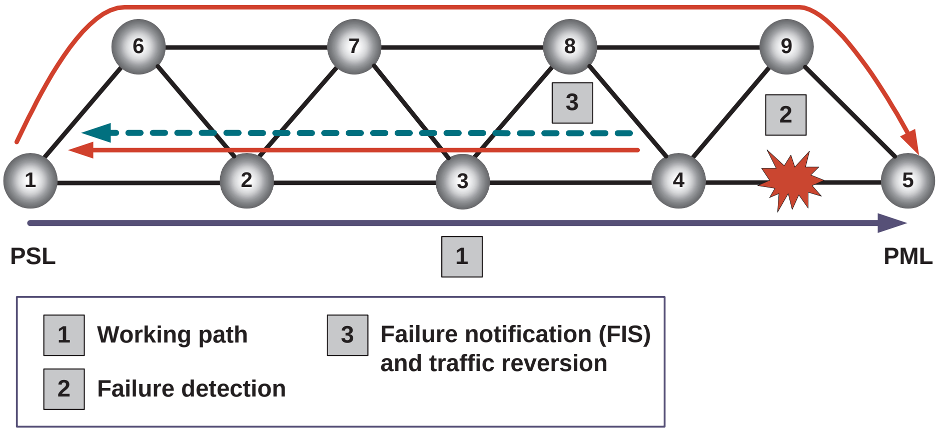


Figure 18: Reverse backup path.

2.1.2.5 p-Cycles

Another fault recovery model is the protection by cycles known as p-cycles [35]. The p-cycles method is based on pre-configuring protection cycles in a mesh network. A p-cycle protects all those links that have their end nodes (source and destination) in the same p-cycle. Consequently, the links belonging to the p-cycle, the links 1-2, 1-3, 2-4 and 3-4 in Fig. 19 are protected. Moreover, the links where their end nodes belong to the p-cycle are also protected and known in the literature as straddling links. In Fig. 19 the link 2-3 is a straddling link. Thus, when a link of the cycle fails, it is protected by the remaining links of the p-cycle as shown in Fig. 19a. On the other hand, if a straddling link fails, then it can be protected by the two alternative paths provided by the p-cycle, that is BP_1 and BP_2 in the Fig. 19b. In this framework, protection-switching decisions can be made quickly because they are carried out in the link that fails.

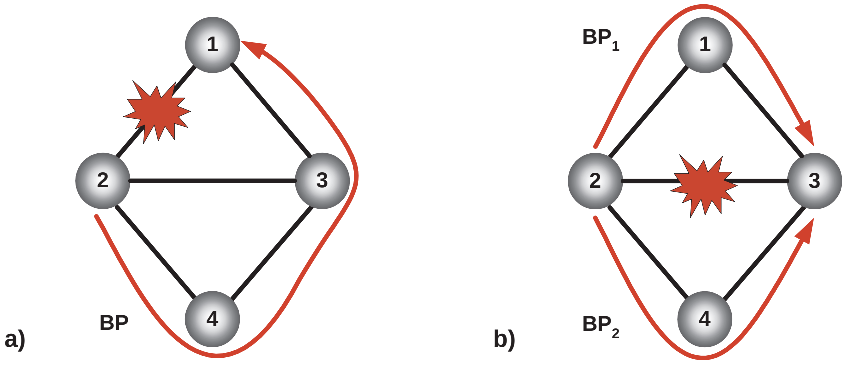


Figure 19: P-cycles method.

2.1.3 Protection Techniques

Protection techniques can be implemented following several architectures [36]: $1+1$, $1:1$, $1:n$, and $m:n$.

- In $1+1$ protection architecture, a backup path is dedicated to each working path. In normal operation mode, identical traffic is transmitted simultaneously on both working and backup paths. At the end of the protected path, selection between the working and protection traffic is made based on some predetermined criteria, such as the transmission performance requirements or default indication.
- In $1:1$ protection architecture, a backup path is also dedicated to each working

path. The protected traffic is normally transmitted by the working path. When the working path fails, the protected traffic is switched to the backup path.

- In $1 : n$ protection architecture, a dedicated backup path is shared by n working paths. In this case, not all of the affected traffic may be protected at the same time. In order to guarantee the recovery of all the working paths, they must satisfy the constraint that these working paths are link disjoint and they do not belong to the same SRLG. Thus, ensuring that a single link failure will only affect to one working path.

- The $m : n$ protection architecture is a generalization of the $1 : n$ architecture. Typically $m <= n$, where m dedicated backup paths are shared by n working paths. Similar constraints concerning SRLG can be applied.

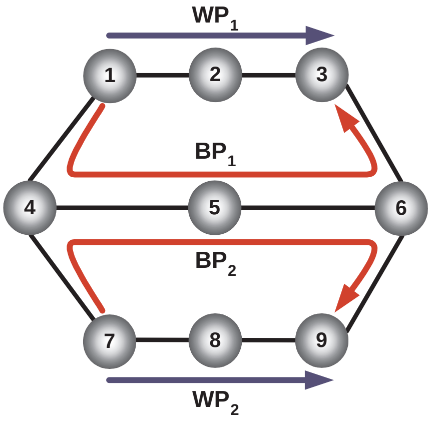
2.1.3.1 Dedicated vs. Shared Protection Study

The reservation of the spare capacity is dedicated for $1 + 1$ and $1 : 1$ protection architectures and is shared for $1 : n$ and $m : n$ protection architectures.

- *Dedicated Protection* ($1 + 1$ or $1 : 1$ protection). At the instant of the working path setup a link-disjoint backup path is reserved and dedicated to it. In $1 : 1$ protection, the backup path could be used to carry other preemptive traffic, and when a link fails in the corresponding working path, the backup path will then be used to carry traffic from the working path.

- *Shared Protection* ($1 : n$, $m : n$ protection). At the time of the working path setup, a link-disjoint backup path is also reserved. However, the spare capacity of the backup path may be shared with other backup paths.

Shared protection saves a large amount of resources by maintaining the same level of protection for single failures. Figure 20 shows the resource consumption of both dedicated and shared protection schemes when two working paths (WP_1 and WP_2) that require 1 unit of capacity are established. In the case of dedicated path protection the spare capacity consumed in each link is equal to the sum of the capacity of each backup path that crosses this link. The total capacity consumed in this example is 8 units. On the other hand, when shared protection is applied the total spare capacity consumed is 6 units. In this case, the links 4-5 and 5-6 are shared by both backup paths (BP_1 and BP_2) because their respective working paths are link disjoints.



Protection technique	Backup paths	Links						Total spare capacity
		1-4	4-7	4-5	5-6	6-3	6-9	
Dedidated Protection	BP_1	1	0	1	1	1	0	8
	BP_2	0	1	1	1	0	1	
Shared Protection	BP_1	1	0	1	1	1	0	6
	BP_2	0	1			0	1	

Figure 20: Dedicated versus shared protection example.

2.1.3.2 Inter-demand and Intra-demand Sharing

Kodialam [37] defined two levels of shared backup: inter-demand sharing and intra-demand sharing. In the case of inter-demand sharing, the spare capacity is shared between different working paths when global/local/segment path protection are used. This corresponds to the case shown before in Fig 20. On the other hand, intra-demand sharing is applicable when local or segment backup methods are used to protect the working path. This case is illustrated in Fig. 21, where the spare capacity used in the link 1-2 is shared by the segment backups BP_1 and BP_2.

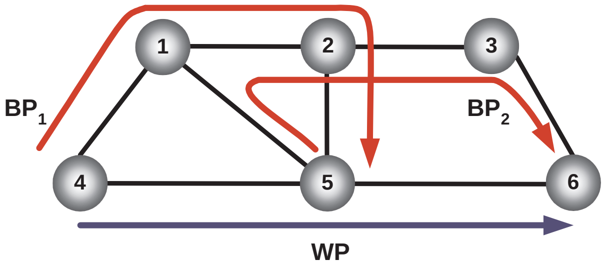


Figure 21: Intra-demand sharing.

2.1.4 Restoration Technique

The restoration technique is also referred as re-routing and 0:1 protection architecture, i.e. no backup path is established prior failure to protect the working path. Thereby, restoration dynamically finds a backup path once a failure has occurred. The resources in the backup path are the currently unassigned (unreserved) resources in the same layer. Traffic preemption may also be used if spare resources are not available to carry the higher-priority protected traffic. The selection of a recovery path may be based on a) preplanned configurations b) network routing policies, and c) current network status (such as network topology and fault information). Signaling is used for establishing the new paths to bypass the fault. Thus, restoration involves a path selection process followed by rerouting of the affected traffic from the working entity to the recovery entity.

In the next-generation optical Internet, a single physical link failure usually brings down a number of logical links (see section 1.3). When IP/MPLS restoration is applied, these logical link failures are detected by IP/MPLS routers, and alternative routes in the logical topology should be found. To facilitate such restoration, the logical topology should remain connected after a failure of a physical link. This can be guaranteed by an appropriate mapping of logical links on the physical topology known as the link-survivable mapping [38]. Figure 22 illustrates the advantage of using link-survivable

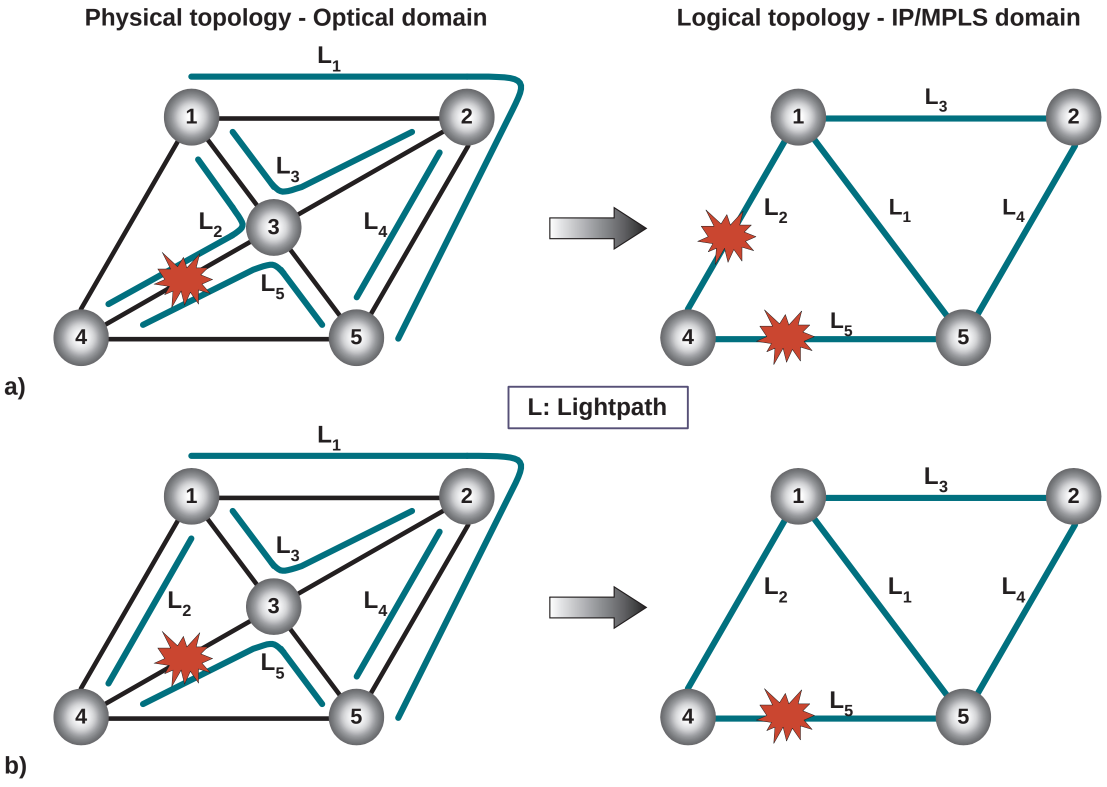

Physical topology - Optical domain
Logical topology - IP/MPLS domain
a)
b)
L: Lightpath
L_1
L_2
L_3
L_4
L_5
1
2
3
4
5

Figure 22: Illustration a) without link-survivable mapping b) with link-survivable mapping.

mapping. On the upper part of the diagram, the logical topology design was based on obtaining a network with the logical links 1-2, 1-4, 1-5, 2-5 and 4-5. On the bottom part of the diagram, the design was based on obtaining the same logical topology but avoiding the isolation of part of the network due to failures. The difference of both logical topologies is how the lightpath L_2 is routed. Thus, in the first design the node 4 is isolated when the physical link 3-4 fails, but in the second design the node 4 remains connected.

2.1.5 Summary of the Backup Path Methods and the Protection Techniques

Table 2 and Table 3 summarize the respective characteristics of each backup path method and protection techniques. In Table 2, a trade-off between resource consumption and recovery time is highlighted. Backup path methods that provide faster recovery time consume more spare capacity than the methods that have long recovery time.

Table 2: Characteristics of the backup path methods.

Method	Spare capacity	Recovery time	Complexity
Global, Reverse	Low	Slow	Requires signaling scheme.
Local	Highest	Fastest	Requires signaling scheme and hardware support. All the nodes require switchover functions (PML and PSL).
Segment	Medium	Medium	Requires signaling scheme and hardware support.
Restoration	None	Slowest	Low - Best effort.

The same is illustrated in Table 3. The protection techniques that have fast recovery time, $1 + 1$ and $1 : 1$, suffer more in terms of higher resource consumption than the protection techniques that aim to share resources.

Table 3: Characteristics of the protection techniques.

Protection techniques		Spare capacity	Recovery time	Availability	Complexity
Dedicated	$1 + 1$	Highest	Fastest	Highest	Requires hardware support
Dedicated	$1 : 1$	High	Medium	High	Requires signaling scheme
Shared	$1 : n$, $m : n$	Medium, low	Medium	Medium	Requires signaling scheme
Unprotected	$0 : 1$	Lowest	Slowest	Low	Low, Best effort

2.2 Fault Recovery Time Evaluation

A reduction in recovery time is a key aspect to consider in order to reach the level of reliability required by many current traffic services. By reducing the recovery time, traffic delay and packet losses are also reduced.

2.2.1 Failure Recovery Time Phases

When a failure occurs the network must go into a state known as the recovery state. This state is characterised by a series of distinct sequential generic phases that describe operations taking place on the occurrence of each and every fault. The operations are listed below:

- *Phase 1*: Fault detection.
- *Phase 2*: Fault localization and isolation.
- *Phase 3*: Fault notification.
- *Phase 4*: Recovery (Protection/Restoration).
- *Phase 5*: Reversion (normalization).

Fault detection, localization and notification phases combined refer to fault management. The term *recovery mechanism* is used to cover both protection and restoration mechanisms. Reversion, also known as normalization process, is defined as the mechanism by which traffic currently carried on a recovery backup path is switched to the former primary path. The reversion process is not a topic that is dealt with directly in this thesis.

As presented in Section 2.1, alternative/backup paths are required in order to recover from network failures. These backup paths can be either computed, established and allocated prior the failure or a posteriori. Given such possibilities, a range of different subphases in an a-priori or a-posteriori manner exists within the recovery process; these are illustrated in Fig. 23 and described in Table 4.

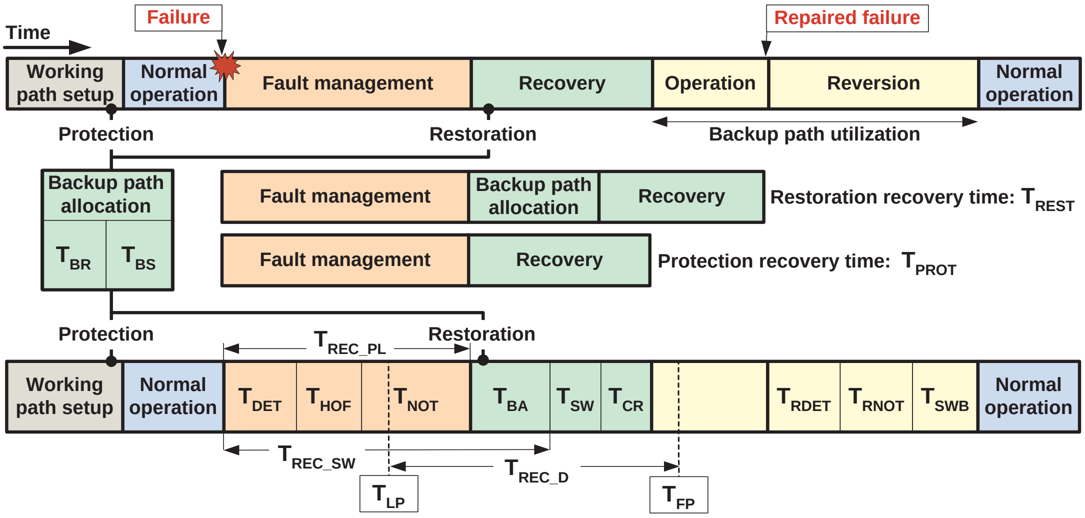


Figure 23: Failure recovery time phases.

Table 4: Description of the recovery time phases.

	Acronym	**Name**	**Description**
	T_{REST}	Restoration recovery time	Total recovery time for restoration schemes
	T_{PROT}	Protection recovery time	Total recovery time for protection schemes
	T_{REC_PL}	Time with packet losses	Time in which there are packet losses
	T_{REC_SW}	Recovery time before switchover	Time required before starting the switchover of traffic
	T_{REC_D}	Delay associated to the recovery time	Delay between the last packet from the working path and the first packet from the backup path
Fault management	T_{DET}	Fault detection time	Time to detect the fault
	T_{HOF}	Hold-off time	Time to allow the lower layers to recover the failure
	T_{NOT}	Fault notification time	Time to inform to the node responsible of the switchover that a failure has occurred
Recovery	T_{BR}	Backup routing time*	Time for new backup creation, routing (T_{BR}) and signaling (T_{BS})
	T_{BS}	Backup signaling time*	Time required to activate the backup path before the switchover
	T_{BA}	Backup activation	
	T_{SW}	Switch over time	Time to switch the traffic from a working path to the backup path
	T_{CR}	Recovery completion time	Time to complete the fault recovery, i.e. the time it takes the first packet to arrive from the backup path to the merging node (LMR)
Reversion	T_{RDET}	Initial path recovery detection time	Time to detect that the working path has been repair
	T_{RNOT}	Initial path recovery notification time	Time to notify about the working path recovery
	T_{SWB}	Switchback time	Time taken to switch the traffic from the backup path back to the working path

* *Only when restoration technique is applied.*

The restoration recovery time differs from the protection recovery time, because restoration has to compute, calculate and route the backup path. Hence, when a failure occurs more recovery phases are incurred within restoration schemes resulting larger delays. The restoration recovery time T_{REST} can be formulated as follows:

$$T_{REST} = T_{DET} + T_{HOF} + T_{NOT} + T_{BR} + T_{BS} + T_{BA} + T_{SW} + T_{CR} \tag{1}$$

See Table 4 for the description of the T_{REST} components.

Attentively, the protection recovery time T_{PROT} does not include the backup routing and backup signaling since the backup path was pre-established:

$$T_{PROT} = T_{DET} + T_{HOF} + T_{NOT} + T_{BA} + T_{SW} + T_{CR} \tag{2}$$

When a failure occurs, packets are lost until the traffic is switched over to the backup path. This time is referred in Fig. 23 as T_{REC_PL} and can be evaluated as follows:

$$T_{REC\_PL} = T_{DET} + T_{HOF} + T_{NOT} \tag{3}$$

The packet loss is a function of T_{REC_PL} and the number of packets directly corrupted by the failure. Generally, losses cannot be totally avoided by most of the protection mechanisms exceptions being $1+1$ protection. However, there are some proposed mechanisms [39] that mitigate this problem by applying tagging and buffering techniques. Without recovery mechanisms such as protection and restoration, traffic flows can take a significant time to respond to loss of connectivity caused by failures with the network. The use of recovery mechanisms, such as protection/restoration, can reduce response times from the order of seconds to milliseconds. Longer the recovery time, the more traffic (packets) is lost.

In the case of protection techniques, the time required to start the switchover is expressed as:

$$T_{REC\_SW} = T_{DET} + T_{HOF} + T_{NOT} + T_{BA} \tag{4}$$

2.2.2 How to Reduce the Recovery Time

The recovery time depends on the failure recovery phases. Table 5 sums up the recovery time phases with their dependencies. Each phase is then described in depth in order to identify those that can be reduced when recovery mechanisms are applied.

Table 5: Recovery time reduction.

Acronym	Name	Time reduction
T_{DET}	Fault detection time	Depends on the monitoring technique to detect the failure.
T_{HOF}	Hold-off time	Depends on the strategy used: hold-off timer or recovery token signal.
T_{NOT}	Fault notification time	Depends on the failure notification delay and the notification method used.
$T_{BR} + T_{BS}$	Backup creation	Depends on the routing and signaling method applied.
T_{BA}	Backup activation	Depends on the backup path distance and signaling process.
T_{SW}	Switchover time	Depends on the node technology.
T_{CR}	Complete recovery time	Depends on the backup distance.

2.2.3 Fault Detection Time Reduction

The fault detection time, T_{DET}, cannot be easily modified. In some system nodes the lower layers report failure detection by means of alarm indications. In other cases the failure detection process is carried out using monitoring techniques [40]. When monitoring techniques are applied, the monitoring rate can be increased in order to report faster detection. However, this can cause scalability problems [40]. A hello protocol was proposed by [40] similar to the Open Shortest Path First (OSPF) [41] in order to detect the failures that are not reported by lower layers. However, timers in routing protocols are typically set to relatively large values when compared to what is required for a recovery mechanism. This hello protocol provides a mechanism which is complementary to existing mechanisms such as physical layer fault detection through liveness messages exchanged between neighbouring nodes. Each node sends a liveness message periodically to its neighbours. A liveness message carries the identification (ID) of the node and the IDs of its neighbours discovered through the liveness messages sent by its neighbours. A node can learn if a bi-directional link is working properly if it sees its own ID in the liveness message sent by the node at the other end of the link.

2.2.4 Hold Off Time Reduction

When a failure occurs, typically a fibre cut, both IP/MPLS and optical layer will detect the failure. In order to avoid the activation of the recovery mechanisms of both layers at the same time, two strategies exist [29]: hold-off timer and recovery token signal. In the *hold-off timer* case, the optical layer first tries to recover from the failure. After the expiration of the holt-off timer, if the optical recovery has succeeded, no recovery actions are needed in IP/MPLS; otherwise, recovery actions in IP/MPLS should be initiated. The hold-off time must be long enough to guarantee that the optical recovery actions have finished. In the *recovery token signal* case, the optical layer explicitly sends the recovery token to the IP/MPLS layer when it cannot recover all or part of the traffic. When the IP/MPLS layer receives the token, the recovery actions are initiated. The advantage of the recovery token signal scheme is that it reduces any added delay required to recover the failure at IP/MPLS layer when the optical layer is unable to do it.

2.2.5 Fault Notification Time Reduction

One of the major considerations in a path recovery mechanism is the control of delay incurred by the failure notification messages since such delays may cause packet loss. The Internet Engineering Task Force CCAMP (IETF-CCAMP) working group has dedicated significant efforts to formalize and minimize the recovery time. One of the most challenging components to minimize recovery time within the recovery cycle is the fault notification time. There are several options that can be used to inform about the fault, including GMPLS-based signaling and flooding. In GMPLS-based signaling, there is generally one fault notification message per disrupted LSP and it is *per-LSP* based. On the other hand, GMPLS-based flooding is *per-failure* based. As well as depending upon the notification strategy, the failure notification time is also a function of the time needed to propagate the fault indication signal (FIS) and the distance between the node detecting the failure and the node responsible for the switchover.

2.2.5.1 Per-failure vs. Per-LSP Notification Strategy

The main difference between *per-failure* and *per-LSP* notification is the number of notification mechanisms that must be executed simultaneously. *Per-failure* fault notification allows one mechanism to notify all relevant nodes of the fault. This is the case of flooding techniques where the node that detects the failure floods the network with information about the fault. On the other hand, *per-LSP* notification requires activating as many mechanisms as the number of LSPs affected by the failure. In an optical network carrying possibly 100's of wavelengths per fibre, *per-LSP* notification can be taxing on the hardware and resource-intensive. An implementation of *per-LSP* failure notification uses the control plane signaling by sending RSVP-TE notify messages. Signaling and flooding techniques are also used to activate the backup path. In the analysis of the backup path activation reduction, subsection 2.2.7, more details pertaining to these techniques are given.

2.2.5.2 Fault Notification Distance

The fault notification distance depends on the fault recovery method applied. The notification distance $D_{p,u}$ is usually defined as the number of links between the upstream node u of the failed link and the upstream PSL p, which is responsible of the switchover (see Fig. 24). However, when $1+1$ protection technique (see Section 2.1.3) is used and the destination node d is not able to detect the failure, the downstream node f of the failed link must notify the failure to d, who will then execute the selection of the backup path as the active path. Hence, the notification distance for $1+1$ protection technique is equal to $D_{f,d}$. Note that, in the case of global and reverse backup path methods, p is always the source node s of the working path, i.e. $D_{s,u}$. On the other hand, when local backup paths are used, then p is also the node that detects the failure, i.e. $p = u$. In this case, the notification distance is always zero, i.e. $D_{p,u} = D_{p,p} = D_{u,u} = 0$.

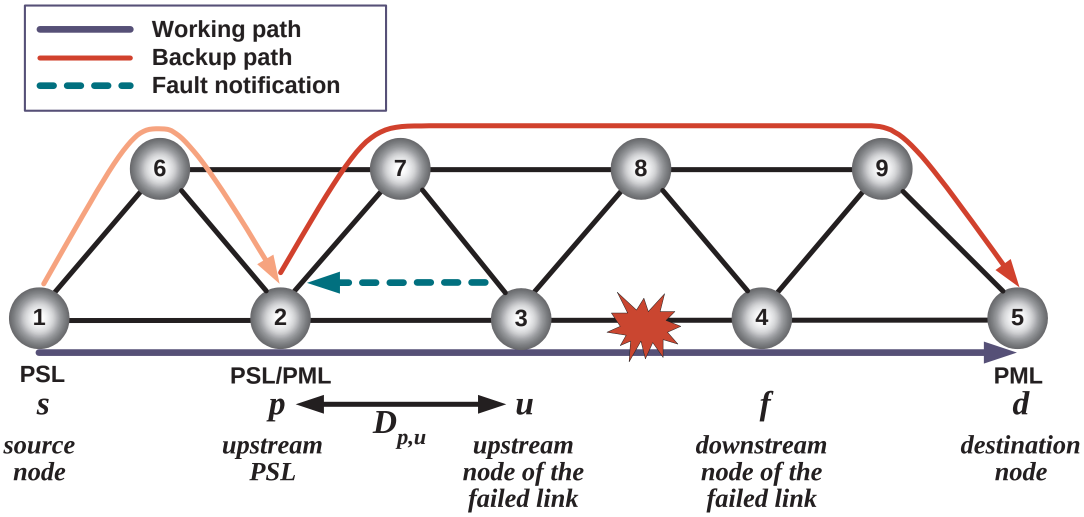


Figure 24: Fault notification distance.

Knowing the fault notification distance is not sufficient to compute the fault notification time. Other factors affecting the notification time can be introduced [42]: the time needed to traverse each link and the delays incurred at the nodes. These factors are analyzed in the following sections.

2.2.5.3 Link Delay

The time to traverse each link is the sum of the transmission time and the link propagation time. The link propagation time, T_{PROP}, corresponds to the latency in the propagation of the packets along the link and is expressed as:

$$T_{PROP} = L_{LINK} \cdot L_{PROP} \tag{5}$$

where L_{LINK} is the physical length of the link and L_{PROP} is the light propagation speed. The L_{PROP} in the fibre is approximately $\frac{2}{3}$ of its speed in free space, i.e. around 200.000 km/s.

The transmission time, T_{TRANS} is computed based on the link capacity as follows:

$$T_{TRANS} = \frac{P_{SIZE}}{S_{LINK}} \tag{6}$$

Where P_{SIZE} is the packet size and S_{LINK} is the link speed. Although the transmission time cannot really be ignored in calculating delays, it cannot be reduced because it is a function of packet size and bit rate. The link propagation time is basically determined by the physical distance traversed, thus, it cannot be reduced beyond a point.

2.2.5.4 Node Delay

From the node point of view, two delays are important: the queuing/buffer delay and the node processing time. The buffer/queuing delay, T_Q, is a function of network load, network topology, position of the fault and fault notification scheme used. Using priority queuing for fault notification messages will ensure that the queuing delay will be bounded. In the case of flooding for fault notification, $T_Q \approx 0$. In the case of per-LSP fault notification, as in the case of using a signaling protocol, the maximum queuing delay in a node n depends on the number of LSPs affected by the failure. This explains numerically basis for the choice of flooding rather than use signaling protocols for fault notification.

The node processing time, T_{PROC}, is considered in the literature [43] as few tenths of a millisecond in the case of a Reservation Protocol (RSVP) object. This value is smaller in the case of a Link Management Protocol (LMP) message requesting the activation of an LSP path.

2.2.5.5 Fault Notification Time Evaluation

In summary, the delay incurred at the recovery process by transmitting a packet from a node u to a node v can be expressed as follows:

$$T_{NOT}^{LSP_j} = \sum_{i=u}^{v} \left(T_{PROP_i} + T_{TRANS_i} + T_{PROC_{n_i}} + T_{Q^j_{n_i}} \right) \tag{7}$$

Where T_{PROP_i} is the propagation time at link i, T_{TRANS_i} is the transmission time at link i, $T_{PROC_{n_i}}$ is the node processing time at node n of link i and $T_{Q^j_n}$ is the queuing time of the fault indication message j in node n. A detailed formulation and analysis of the queuing time and node processing time, are beyond the scope of this work. The presented formulation only gives an approximated boundary for evaluating the recovery time, enough to achieve the objectives proposed in this thesis.

2.2.5.6 Reducing the fault notification time

The fault notification time, equation (7), is a function of the following parameters:

- Link propagation time.
- Link transmission time.
- Node processing time.
- Node queuing time.

If, possible, reduction in any of these parameters will result in a reduction of recovery delay. As described, link transmission time cannot be changed as it is a simple function of packet size and bit rate. The node queuing delay depends on the notification strategy used, position of the fault, network load and network topology; this delay is beyond the scope of this work. The node processing delay may be reduced by decreasing the number of hops between the node that detects the failure and the node responsible of the switchover. The link propagation time depends on the link length (physical length) and transport media. The longer the total physical link length is between the node that detects the failure and the node responsible of switching the traffic; the longer the total propagation delay. Note that the shortest route in terms of number of hops is not always the fastest. Under the assumption that the rapid technological evolution should be able to improve the node technology reducing node delays, only the link propagation time is taken into account in this thesis. Since the link propagation time depends on the link length, the minimization of the fault notification distance (total link length) is only taken into account.

2.2.6 Backup Creation Time Reduction

The time taken to create the backup path depends on the moment that the backup path is established and the spare capacity is allocated. In the case of on-demand backup paths, the backup path creation time depends on the routing and signaling methods used since the backup path is computed after the fault is detected. On the other hand, if the backup path is pre-established and pre-allocated prior the occurrence of the fault, i.e. using protection techniques, this delay is avoided. When the level of reliability, in terms of recovery time, is required to be high, protection techniques should be used. In IP/MPLS, a backup path can also be pre-established without allocating resources, i.e. spare capacity. This technique is known as *fast restoration*. However, in network scenarios with high traffic loads and no packet prioritization techniques, not reserving capacity could result in longer activation times [32, 44].

2.2.7 Backup Activation Time Reduction

Nodes on a backup path are aware that they are protecting against the failure of a particular resource. Thus, when nodes are notified of the failure, they activate the backup path by performing any required hardware configuration (for example, moving mirrors in the case of a MEMS-based switching fabric) [45]. The notification mechanisms used to activate the backup path are the same that are used to notify the failure to the node that is responsible of the switchover. Therefore, GMPLS-based signaling and flooding methods may be used.

2.2.7.1 GMPLS-based Signaling Method

In signaling-based techniques, when a failure occurs, the detecting node sends a fault indication message (FIS) for each LSP affected by the failure. The PSL of each affected LSP, then, sends data on the backup path when it receives the failure acknowledgment message from the node responsible for traffic merging (PML).

2.2.7.2 GMPLS-based Flooding Method

In the case of flooding techniques, the detecting node sends a message with information about the fault to all network nodes. The recovery nodes affected by the failure take the necessary actions. The advantage of flooding with respect to signaling-based techniques is the minimum notification/activation delay, due to the following:

- *Buffer delay*: The queuing time is zero.
- *Node processing delay*: Minimum time for processing the notification messages.
- *Minimum path delay*: In flooding techniques notification messages follow the minimum reverse path (in terms of delay). Messages are dispatched in all directions to all nodes in the networks, this guarantees the minimum delay.

Therefore, all nodes are informed about the failure and this can improve the evaluation of future routing requests, avoiding signaling failures. A case where flooding may be inefficient is when two nodes x and y are interconnected via multiple links $l_1, l_2, ..., l_n$. If x receives a new FIS from another neighbor node z, it will flood n copies of the same FIS to its neighbor node y, one per each link $l_1, ..., l_n$ though y will just need one copy of the FIS. Hence, some capacity will be unnecessarily consumed and this will also consume CPU on both x and y nodes. Finally, y will acknowledge the reception of the FIS retransmitting the FIS over all the $n - 1$ other links back to x. Modifications of the flooding procedure from a per-link basis to a per-neighbor basis have been proposed [29].

2.2.8 Switchover Time Reduction

Switchover is the process of switching the traffic from the working path through which the traffic is flowing, to the alternative/backup path. This phase starts after the responsible entities, PSL and PML nodes, are notified of the failure and the backup path is activated. This operation critically depends on the node technology and, therefore, cannot be easily reduced.

2.3 Failure Probability Evaluation

The network reliability expressed in terms of failure probabilities is an important issue and has been also considered in the literature [32,46]. By reducing the failure probability of a working path, the reliability of the traffic that is transmitted through it is increased. Quantifying reduction that can be offered requires taking into account explicit link failure probabilities. Note that, in the multi-layer scenario IP/MPLS optical networks, two kind of links exist: physical links and logical links (lightpaths).

2.3.1 Physical Link Failure Probability

The calculation of the failure probabilities of optical fibres is a key topic within survivable networks [47]. It is possible to estimate the probability of failure of individual links based upon historical and physical data available. The calculation can be approximated based on known probabilities regarding certain aspects of transmission technology, for instance, the type of physical link, the node characteristics, the geographical distribution of the network links, etc. In addition, there may be failures due to accidents, natural catastrophes caused by the temporary manufacturing process mistakes and other reasons which are statistically infrequent but which can be obtained from the tests. These factors determine the actual value of the link failure probability of a physical link (i, j), F^P_{ij}, at a given moment. In this work, it is assumed that all the physical link failure probabilities of a given network are known and they are also independent of each other. Moreover, the $F^P_{ij} \ll 1$ for all physical links.

2.3.2 Path Failure Probability

Based on the knowledge of the physical link failure probabilities of the network, the Path Failure Probability, PFP, can be computed [44, 48]. The failure of a path results from the failure of any of the physical links that the path crosses. The PFP is:

$$PFP = 1 - \prod_{(i,j) \in \mathcal{P}} (1 - F^P_{ij}) \tag{8}$$

Where $\mathcal{P}$ is the set of the physical links of the path. This product is simplified based on the assumption $F^P_{ij} \ll 1$, obtaining:

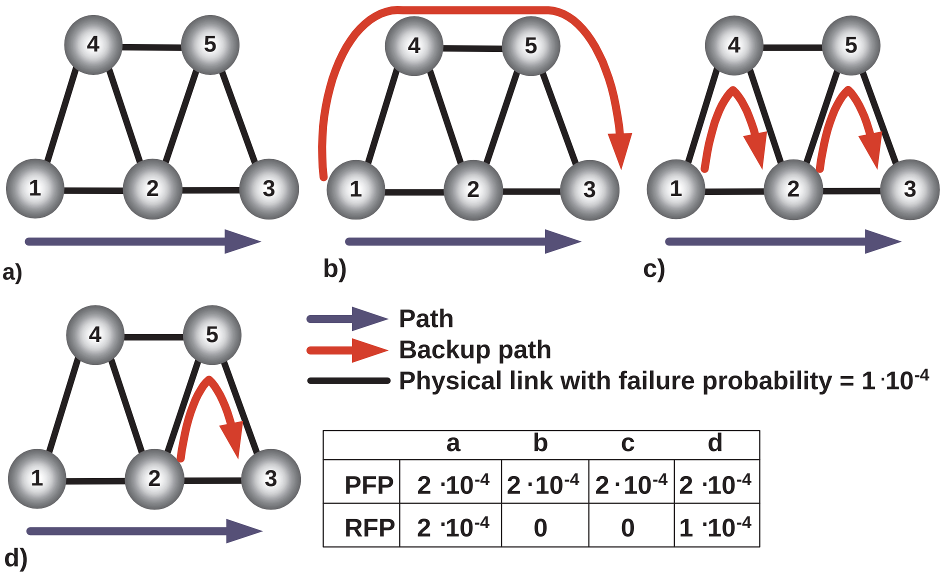
$$PFP = \sum_{(i,j) \in \mathcal{P}} F^P_{ij} \tag{9}$$

where the PFP is approximated by the sum of individual failure probabilities of the physical links that the path crosses. These approach, equation (9), overestimates the value of PFP, it is a conservative approach.

2.3.3 Residual Failure Probability

The effects of PFP is mitigated when the working path is fully or partially protected by a backup path. With *full protection* all the links of the working path are protected by the backup path. On the other hand, with *partial protection* not all the links of the working path are protected. The Residual Failure Probability, RFP, is defined as the sum of the failure probabilities of the unprotected links. In the simple network scenario shown in Fig. 25, a working path is established crossing two links with different failure probabilities. According to the backup path routing policy used, three possibilities exist:

- $RFP = PFP$. The residual failure probability is equal to the path failure probability when the working path is unprotected, i.e. no backup path is established to protect the links of the working path (see Fig. 25a).
- $RFP = 0$. The residual failure probability is 0 when whole the links of the working path are protected with protection mechanisms (see Fig. 25b and 25c).
- $0 < RFP < PFP$. The residual failure probability is also reduced but no zero when some (not all) of the links of the working path are unprotected. In the example shown in Fig. 25d, the RFP is $1 \cdot 10^{-4}$ because the physical link $1-2$ is unprotected.



	a	b	c	d
PFP	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$
RFP	$2 \cdot 10^{-4}$	0	0	$1 \cdot 10^{-4}$

Figure 25: Residual failure probability when a) no protection, b) global protection c) local full protection d) local partial protection is applied.

2.3.4 Lightpath Failure Probability

In the optical domain, the lightpaths connect node pairs by allocating a free wavelength on all of the physical links (fibres) that they cross. Hence, the lightpath failure probability of a logical link (u, v), F^L_{uv}, depends on the physical link failure probability and the protection mechanism used to protect it. This probability can be expressed in terms of residual failure probabilities as follows:

$$F^L_{uv} = RFP_{uv} \tag{10}$$

Note that in the optical domain, either the whole lightpath is protected or none of its physical links are protected. Therefore, the RFP of a lightpath (u, v), RFP_{uv}, is evaluated as:

$$RFP_{uv} = \begin{cases} 0 & \text{if all the physical links of lightpath } (u,v) \text{ are protected} \\ PFP_{uv} & \text{if any of the physical links of lightpath } (u,v) \text{ are not protected} \end{cases} \tag{11}$$

2.3.5 LSP Failure Probability

In the IP/MPLS domain, the LSPs connect node pairs by allocating capacity on all the lightpaths that they cross. The LSP failure probability is a function of the residual failure probabilities of the lightpaths that the LSP crosses:

$$F_{LSP} = \sum_{(u,v) \in LSP} RFP_{uv} \tag{12}$$

Table 6 summarizes the failure probability evaluation of each element of the network.

Table 6: Failure Probability Evaluation.

Domain	Link Failure Probability			Path Failure Probability	
Optical	Physical link	Technology, Tests, Statistics	F^P_{ij}	Lightpath	Eq.9
IP/MPLS	Lightpath	Eq.11	F^L_{uv}	LSP	Eq.12

2.4 *Differentiated Quality of Service with Protection*

Survivability techniques are now key factors improving and satisfying the increasing requirements of Quality of Service with Protection (QoSP). Although not all applications require the same level of reliability, networks do not currently offer a large set of differentiated recovery methods. Besides, some applications are more stringent about their QoSP requirements than others. Moreover, in many cases, improving fault recovery involves, in terms of resource consumption, very expensive mechanisms, such as $1 + 1$ protection, which cannot always be deployed throughout the whole network. The differentiation of recovery methods as the part of service level agreements may be a very convenient means whereby operators can increase revenues for minimal incremental cost [49].

2.4.1 Factors Involved in the Traffic Classification

In the previously proposed frameworks introducing traffic differentiation, a range of different factors are used to distinguish the QoSP for traffic classes:

1. *Reliability of Service Classes* (RoS) [50]. This framework differentiates between classes without distinction between the backup path methods used. Hence, RoS framework only selects the survivable technique used to protect the traffic: protection, restoration or unprotected.

2. *Resilience Classes* (RC) [51]. RC framework is similar to RoS framework because the traffic differentiation is also based on the survivability technique used. Additionally, in the former the emphasis is further placed on some quality aspects like recovery time.

3. *Differentiated Reliability Connections* (DiR) [52]. This framework is based on the maximum failure probability assessment for each connection. The connections are differentiated according to their permissible maximum failure probability.

4. *Quality of Service Protection* (QoSP) [48]. This framework takes into account different backup path methods (local, global and reverse). For each traffic class, the most suitable backup path method is selected according to the QoSP requirements in terms of recovery time, traffic loss and resource consumption.

5. *Quality of Protection, Sharing Focused* (QoP) [53]. This framework theoretically takes into consideration sharing of resources, i.e. the amount of shared nodes and links of backup paths.

From these frameworks, the factors that may influence the QoSP and, thus, considered in this thesis are:

1. *Level of reliability.* Each traffic class is classified according to its reliability requirement (protection, restoration or unprotected) and, accordingly, the most suitable backup path method will be also applied.
2. *Fault recovery time.* The recovery time requirements of each traffic class also form part of the traffic classification.
3. *Level of protected links.* Failure probabilities are considered. For those connections with no requirements or limited failure probability, partial protection is applied. Otherwise, full protection is used, i.e. all the links are protected by a backup path.
4. *Survivability technique.* Global, local and segment backup path methods are also considered according to the QoSP traffic requirements.
5. *Protection architecture.* According to each QoSP traffic requirements, either shared or dedicated capacity allocation is selected.

2.4.2 Traffic Services Classification

In this work, following the Differentiated Service draft from the IETF [54,55], four traffic services are considered. The Expedited Forwarding (EF) class is defined to transport real-time traffic, two Assured Forwarding (AF1 and AF2) classes are used by traffic with two different types for losses and, finally, the Best Effort (BE) class for traffic with no QoS requirements.

In this work, the traffic protection requirements are characterised by using each traffic service as shown in Table 7. Four levels of reliability are considered according to the requirements of the traffic services:

- *Null Reliability* (NR). Protection is not required for this traffic class.
- *Low Reliability* (LR). Protection is offered if there are sufficient network resources. Partial protection is applied for LR traffic. With partial protection only those links that can reach a backup path with sufficient capacity are protected. Fast protection is not required, thus global, segment and local protection methods may be used.

- *Medium Reliability* (MR). Full protection is required, i.e. all the links must be protected. In order to reduce the fault recovery time, only segment and local methods may be used.

- *High Reliability* (HR). Fast recovery is required to protect all the links. Consequently, only local protection method is used to protect HR traffic. Moreover, dedicated capacity allocation is also used in order to reduce the recovery time.

Table 7: Traffic Services Classification.

QoSP Factors	**Traffic Class**			
	Best Effort **(BE)**	Assured Forwarding		Expedited Fwd. **(EF)**
		(AF2)	**(AF1)**	
Level of reliability	Null Reliability **(NR)**	Low Reliability **(LR)**	Medium Reliability **(MR)**	High Reliability **(HR)**
Level of protected links	None	Partial	Full	Full
Fault recovery time	High	Medium, Slow	Fast	Very Fast ($\simeq 0$)
Survivability technique	Restoration	Protection: global, segment, local backups	Protection: segment, local backups	Protection: local backups
Protection architecture	None	Shared	Shared	Dedicated

2.5 Concluding Remarks

In this chapter, a first overview of the recovery techniques used to provide survivability, protection and restoration, has been introduced. Both aim to compute an alternative/backup path where the traffic affected by the failure is switched over. Different backup path protection methods have been also presented: global, local, segment, reverse backup path and p-cycles. The strengths and the weaknesses of each backup path method in terms of recovery time and resource consumption have also been identified.

Different applications, traffic classes, coexist into the network with different QoR requirements. In order to guarantee the QoR of the traffic against any single link failure, the backup path method most appropriate to the required QoR must be applied. Thus, a traffic classification was presented in Table 7 (Section 2.4). The advantages and disadvantages of possible backup path methods should be considered when the alternative/backup path is computed. The selection of the working and alternative paths is a crucial step in offering the required QoS and QoR to traffic services. Some parameters, such as recovery time, could be affected negatively if inappropriate routing algorithms are used. The next chapter analyzes QoS routing algorithms and the parameters that these algorithms should take into account in order to ensure an efficient use of the resources.

CHAPTER III

QUALITY OF SERVICE WITH PROTECTION ROUTING ALGORITHMS

As reviewed in Chapter 2, many survivability techniques are proposed that can be broadly classified into protection and restoration techniques. The distinction between these techniques is the timing of spare capacity allocation and the timing of backup route calculations. Protection techniques offer faster recovery time than restoration techniques and provide guaranteed recovery against single link failures. However, the pre-establishment of a link-disjoint path pairs, the working path and the backup path, results in high resource consumption. Considering different protection techniques and including specific objectives in the routing algorithms, the amount of spare capacity used for protection may be reduced and the recovery of all traffic affected by the failure is guaranteed. This chapter presents a literature survey of the evolution of traditional QoS routing algorithms that did not consider failure resilience coupled with current QoS with Protection (QoSP) algorithms. First, the traditional QoS routing algorithms are reviewed and their main objectives are highlighted. In particular, shortest path, load balancing and minimum interference metrics are described. Next, a brief summary of the previous QoSP routing algorithms considering failure resilience are presented. Here, new metrics are introduced into the routing algorithm objectives: available network information and recovery time reduction. The available network information (partial or full) is taken into account in order to make an efficient use of the spare capacity. The reduction of the recovery time is also considered by applying segment backups. Finally, simulations are carried out in order to evaluate the reviewed routing metrics: shortest path, load balancing, minimum interference, network information available and recovery time reduction. These metrics are analyzed in terms of network resources, request rejection ratio and recovery time.

3.1 Traditional QoS Routing Objectives and Algorithms

Traditional QoS routing algorithms use two different objective functions to optimize network performance: the shortest path should be selected for minimizing the length of the path and the least loaded path, i.e. lowest allocated capacity, should be selected for load balancing. There is a third objective, which is the minimization of the number of request rejections. Minimum interference schemes optimize the bandwidth usage and minimize the number of rejected requests [56].

3.1.1 Shortest Path Routing Algorithms: The Dijkstra Algorithm

Finding the shortest path between a given node pair in a graph $G = (V, E)$ is the first goal of most routing algorithms. An efficient and commonly cited algorithm for finding the shortest path is the Dijkstra algorithm [57]. The Dijkstra algorithm basically computes the shortest path from a given node in V to the rest of the nodes in the graph. The shortest path of a particular node pair is computed terminating the algorithm once the shortest path to the destination node is reached. The Dijkstra algorithm can easily be employed in a dynamic environment [41].

Algorithm 1 Dijkstra

```
INPUT
$s$: source node;
$d$: destination node;
$G = (V, E)$: network graph;
ALGORITHM
for all $v \in V - \{s\}$ do
  $Cost(v) = \infty$
  $Pred(v) = s$
end for
$Cost(s) = 0$
$Q \leftarrow s$
while ($d \notin Q$ and $Q \neq \oslash$) do
  $u \leftarrow min\_cost(Q)$
  $Q = Q - \{u\}$
  for all $v \in adjacency(u, G)$ do
    if ($Cost(u) + link\_cost_{uv} < Cost(v)$ then
      $Pred(v) = u$
      $Cost(v) = Cost(u) + link\_cost_{uv}$
      $Q \leftarrow v$
    end if
  end for
end while
```

In the Dijkstra Algorithm (see Alg. 1), $Cost(v)$ is a vector containing the path cost from s to v; $Pred(v)$ contains the v's predecessor node. Q represents the list of adjacent vertices which are not visited yet; and $link\_cost_{uv}$ is, in this case, the physical length of the link (u, v). Function $min\_cost(Q)$ returns the element $u \in Q$ with the lowest $Cost(u)$; and $adjacency(u)$ represents the adjacency list of the vertex u in graph G. Note that, initially, a distance value of ∞ is assigned to all nodes of the graph except the source node s. This initial value should be larger than the shortest path to be computed. Since the shortest path is the goal of this algorithm, the distance can be set equal to any value greater than the longest path between source and destination nodes. So, it can be set to a number greater than the number of links $|E|$ multiply by the length of the longest link $length_{max}$: $|E| \cdot length_{max}$. Its efficiency, in terms of order of complexity, is $O(|V|^2)$. Another shortest path algorithm is the Min Hop Algorithm (MHA). It is based on the Dijkstra algorithm and it selects the path with the least number of hops (links) instead of physical length. Thus, the $link\_cost_{uv}$ of all $(u, v) \in E$ is set to 1.

3.1.2 Load Balancing Routing Algorithms

Dijkstra's algorithm is very efficient for computing shortest path but it does not take into account the current state of the network. One important variable of the network is the amount of available capacity of the links. A link will only accept a new request with given capacity requirements, if at least that much capacity is still available on the link. A Widest-Shortest Path (WSP) algorithm has been proposed [58] where both criteria are mixed. The WSP algorithm first selects the path with minimum hop count from among all feasible paths; and then, if more than one path is eligible, the one with maximum reservable capacity is chosen. The maximum reservable capacity of a path is the minimum of the available capacity of all links along the path. The Shortest-Widest Path (SWP) [58] uses the opposite criterion of the WSP, i.e. first criterion is to select suitable paths with maximum capacity and if more than one is feasible, the one with minimum hop count is selected. In other words, WSP gives higher priority to resource utilisation and SWP gives priority to balancing the network load.

3.1.3 Minimum Interference Routing Algorithms

The aim of minimum interference routing is to find a feasible path for an incoming request that *interferes* as little as possible with the future requests. The term *interference* in this domain refers to the fact that the path selected will try to avoid those links that may be *critical* for satisfying new future requests. These algorithms assume neither the knowledge of future requests nor any statistical traffic profile. Figure 26 shows an example of this *interference* effect. Consider that all the network links are unidirectional and have a capacity of 100 units. Lets suppose that three requests arrive at the following order $WP_1/WP_2/WP_3$ requesting 50 units of capacity between the node pair $(s_1, d_1)/(s_2, d_2)/(s_3, d_3)$ respectively. If the MHA is used, the selected route for the request WP_1 is: $s_1 - 4 - 8 - d_1$. For the WP_2, the selected route is: $s_2 - 4 - 8 - d_2$. After setting up WP_1 and WP_2 the remaining capacity of link $4 - 8$ is 0. Thus, this combination produces the rejection of the request WP_3. In this example it would be better to choose the route $s_1 - 5 - 6 - 7 - d_1$ for establishing the request WP_1 even though the path is longer. In this way, the rejection of WP_3 would be avoided.

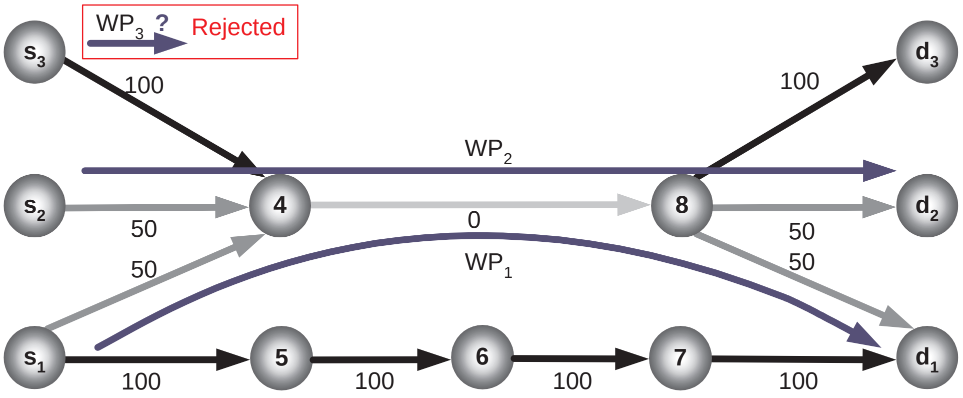


Figure 26: Minimum Interference. Illustrative example.

The first minimum interference proposal introduced by Kodialam [56] was the Minimum Interference Routing Algorithm (MIRA). In MIRA, the identification of *critical* links is based on a pre-process phase of maximum minimum (maxmin) flow computation to generate a weighted graph. The link weight w_{uv} is an increasing function of link *criticality*. Once the weighted graph is obtained, links with residual capacity inferior to the required capacity, *critical links*, are removed from the graph G. Finally, Dijkstra's algorithm is used to select the path but, instead of adding the *length* of the link to the

Cost, the link weight w is used.

MIRA was proposed for MPLS-based network scenarios. Minimum interference routing has been also extended to include the establishment of lightpaths as well as routing in the logical topology in the Maximum Open Capacity Routing Algorithm (MOCA) [59]. However, these algorithms suffer complex computation with large calculation times, to overcome this drawback, new algorithms have been proposed. A first proposal without maximum-flow calculations was presented by Iliadis in Simple MIRA (SMIRA) [60]. SMIRA uses a new procedure for obtaining the set of critical links without maxmin-flow computation, called k-widest-shortest-path under bottleneck elimination. This procedure identifies a set of critical paths by using a WSP algorithm. Another similar procedure, in terms of not using maxmin-flow calculations to obtain the critical links, is the Wang, Su and Chen's (WSC) algorithm [61]. These algorithms are also proposed for MPLS-based network scenarios. Two enhanced proposals of SMIRA and WSC were presented in Integrated SMIRA (SMIRA-I) [62] and the Light Minimum Interference Routing (LMIR) algorithm [63]. Light Minimum Interference Routing (LMIR) is one of the most recent proposals based on the Dijkstra algorithm. LMIR attempts to find the paths with the lowest capacity among all source and destination node pairs in order to determine the critical links. The major advantage of LMIR is that it does not involve the execution of maximum flow algorithm. A new proposal in the wireless-optical dynamic network scenario was introduced in SMIRA-I. SMIRA-I extends SMIRA to compute the critical weight for each actual link and potential link (dynamic scenario).

3.2 QoSP Routing Schemes

A crucial aspect in the development of a fault management system is the selection of backup paths (BPs). Although routing algorithms reviewed in the above section (MHA, WSP, MIRA, LMIR, etc.) can be used to compute the BP, they do not include any specific objective to actually improve the protection level, such as an efficient use of the spare capacity or the minimization of the fault recovery time. In the recent literature, there has been much interest in setting up QoS paths that are resilient to faults. The drawback of using QoS with Protection (QoSP) routing algorithms is the resource consumption. Depending on the recovery method applied, the amount of resources consumed changes, thus, the number of rejected requests may increase.

3.2.1 Shortest Link Disjoint Path Algorithm

When a number of paths exist between a given pair of nodes in the graph, not all paths may be independent each other because they may share nodes and/or links. A pair of paths are link (node) disjoint if they have no links (nodes) in common. In this thesis, node disjoint paths are not considered; single link failures are only taken into account. An intuitive method to determine two shortest link-disjoint paths between a source and destination node pair consists of two steps. The two step approach algorithm finds a pair of paths by first finding the shortest path, and then finding the shortest path in the same graph, but with the links of the first shortest path deleted from the graph [28]. However, the disjoint path pair found using this approach may not be optimal. According to the example shown in Fig. 27, the shortest disjoint path pair for the node pair (s, d) obtained using the two step approach is: $s - 3 - 4 - d$ and $s - 1 - 2 - d$. The cost of both paths is $3 + 6 = 9$, because the second shortest path has a significantly larger length. Thus, the optimal solution is: $s - 3 - 6 - d$ and $s - 1 - 4 - d$ with a total cost of $4 + 4 = 8$.

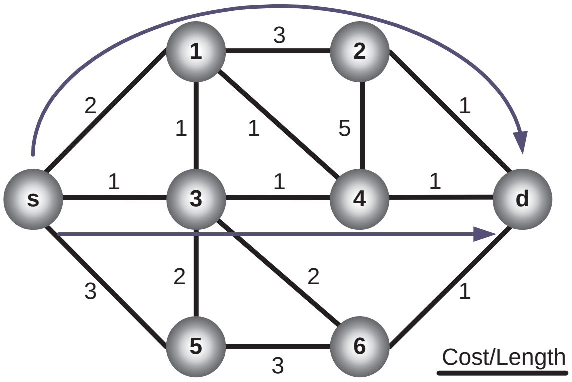


Figure 27: Two steps approach suboptimal solution.

The two steps approach algorithms have what is called the *trap topology problem* [64]. These algorithms can fail to find pairs of link disjoint paths, when they actually exist. An example is shown in Fig. 28. In this network scenario, the shortest path $s-2-3-d$ is selected between the node pair (s, d). Consequently, there is not any link disjoint path of this path selected, despite a link disjoint path pair exists for this node pair: $s-1-3-d$ and $s - 2 - 4 - d$. This drawback is overcome by the Suurballe algorithm [28, 65]. The Suurballe algorithm may be stated as follows:

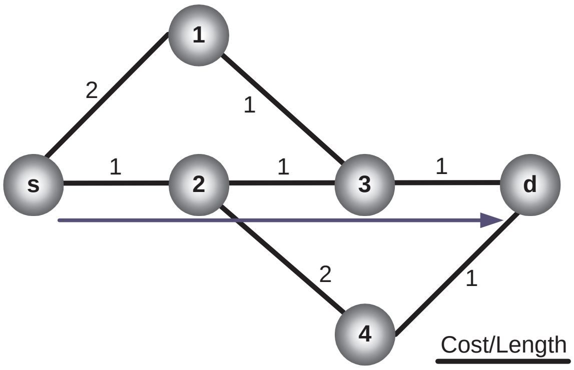


Figure 28: Trap topology problem of the two steps approach.

1. For the given graph $G = (V, E)$, perform the following transformation:

$$length'_{ij} = length_{ij} + D(i) - D(j) \qquad \forall (i,j) \in E$$

   Where $length_{ij}$ is the length of the link (i, j) and $D(k)$ denotes the shortest path distance of node k from the source node s. The Dijkstra algorithm is used to find the shortest distance $D(k) \quad \forall k \in V$.

2. Remove the forward arcs along the shortest path (s, d). Set the length of the reverse arcs along the shortest path (s, d) to 0.

3. Repeat Dijkstra's algorithm to find the shortest path from (s, d) in the modified graph G'.

4. Erase the common links (if any) of the two paths found from s to d. Merging the two paths, the desired link-disjoint path pair results.

An step by step example is shown in Fig. 29. In this case, the link disjoint path pair is found.

Although the algorithm proposed by Suurballe is optimal and has polynomial computational complexity, it does not take into account other metrics that can affect the establishment of new requests: minimum interference and the spare capacity. When these metrics are considered, two-step routing algorithms are more suitable than one-step because the backup path is computed based on the characteristics of the working path [32].

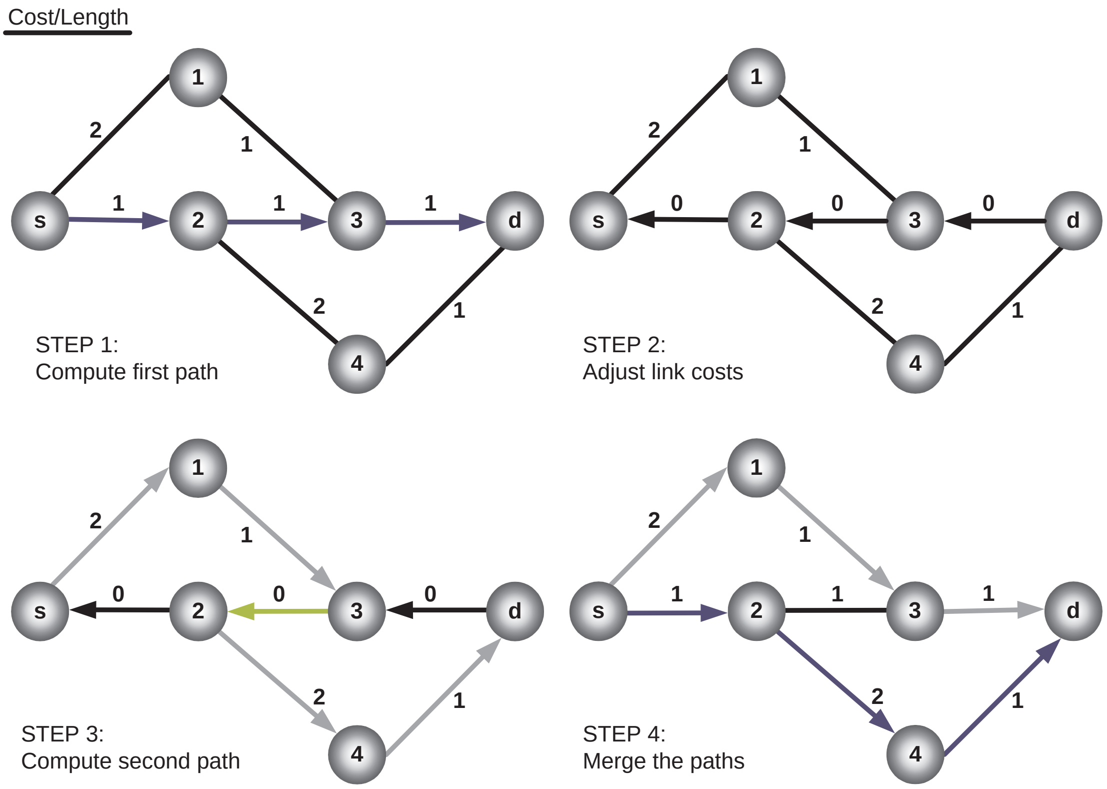


Figure 29: Computing link disjoint paths using Suurballe's algorithm.

3.2.2 Minimum Interference Restorable Routing

Due to the added complexity, key QoS routing proposals, which use minimum interference concept, do not consider protection in their main objectives (see section 3.1.3 for more details). Some preliminary proposals which do consider protection in their objectives, have a high computational cost and resource consumption, e.g. the Minimum Interference Restorable Routing (MIRR) presented by Kar *et al.* [66] where a $1+1$ protection technique is used. Although $1 + 1$ protection achieves good results in reducing the recovery time, it has significant resource consumption. Usually, the fastest protection techniques, such as 1+1 and 1:1, use a large amount of spare restoration capacity. Hence, they are not cost effective for most customer applications. Significant reductions in spare capacity can be achieved by sharing this capacity, $1 : n$ or $m : n$ protection techniques, as detailed in next section.

3.2.3 Shared Backup Routing Algorithms

The accuracy and performance of the shared backup routing algorithms are based on the available network information. In this section, the link state information considered by the shared backup routing algorithms is reviewed. Moreover, routing algorithms based on the available routing network information presented [67–69] by Ho, Li and Kodialam respectively, are reviewed.

3.2.3.1 Routing Information

When the backup path is computed, the shared backup capacity should be taken into account in order to reduce the amount of spare capacity. Therefore, the following link state information is flooded by the routing protocols:

1. *Working capacity* A_{ij}. Total amount of capacity used by working paths.

2. *Spare capacity* S_{ij}. Total amount of capacity reserved. This capacity is not used when the network is in a non-failure condition.

3. *Residual capacity* R_{ij}. Total amount of capacity that is free to be allocated by working or backup paths. Note that, $R_{ij} = C_{ij} - A_{ij} - S_{ij}$, where C_{ij} is the capacity of the link (i, j).

4. *Cost* W_{ij}. The cost of using link i is set by the network operators. A path with a smaller cost is typically preferable. In the case of MHA, the cost is a constant $W_{ij} = 1 \quad \forall(i, j) \in E$.

Such information detailed above is available in current extensions of OSPF/IS-IS for GMPLS [24, 41]. It should be also possible to distribute the spare capacity information using existing OSPF/IS-IS traffic engineering extensions.

Li [68] addressed the issue of how to collect the necessary information to select the backup path in order to minimize the total reserved resources over all network links. For each link, a master node is responsible for maintaining the link status. This node could be the node terminating the bidirectional link having the smaller node *id*. The master node for link (i, j) maintains a local array $Failother_{ij}^{uv}, \quad (u, v) \in E$, where $Failother_{ij}^{uv}$ is the amount of capacity required on link (i, j) to restore all failed working paths if link (u, v) fails. Then, reserving a capacity equal to $S_{ij} = \max_{(u,v)\in E} Failother_{ij}^{uv}$ ensures that there are enough resources reserved to protect against any single link failure. S_{ij} is distributed to all other network nodes using the extended routing protocol. In addition, the master node of link (i, j) along the working path also keeps track of the capacity that has been reserved on other network links to protect against the failure of link (i,j) itself. This information is maintained in another local array, called $Failself^{ij}$, where $Failself_{ij}^{uv}$ stores the capacity required on link (u,v) to restore all LSPs affected if link (i,j) fails. Both $Failself_{ij}$ and $Failother_{ij}$ are updated during the signaling process. During the lifecycle of a working path, the operations include creation and deletion. Signaling extensions and procedures for updating the local data structures during working path creation and deletion have been proposed by Li [68].

The shareable spare capacity is computed according to the routing network information available and the selected working path WP. Hence, in the case of partial information, the shareable capacity is computed based on the maximum working capacity M reserved on each link of the new WP, where $M = \max_{(u,v)\in WP} A_{uv}$. On the other hand, when full information is available in the network, the source node collects the array $T_{ij}, (i,j) \in E$, where T_{ij} is the maximum capacity needed on link (i,j) if any link along WP fails, i.e. $T_{ij} = \max\limits_{(u,v)\in WP} Failother_{ij}^{uv}$. Figure 30 illustrates each category of capacity along a link depending on the network information available.

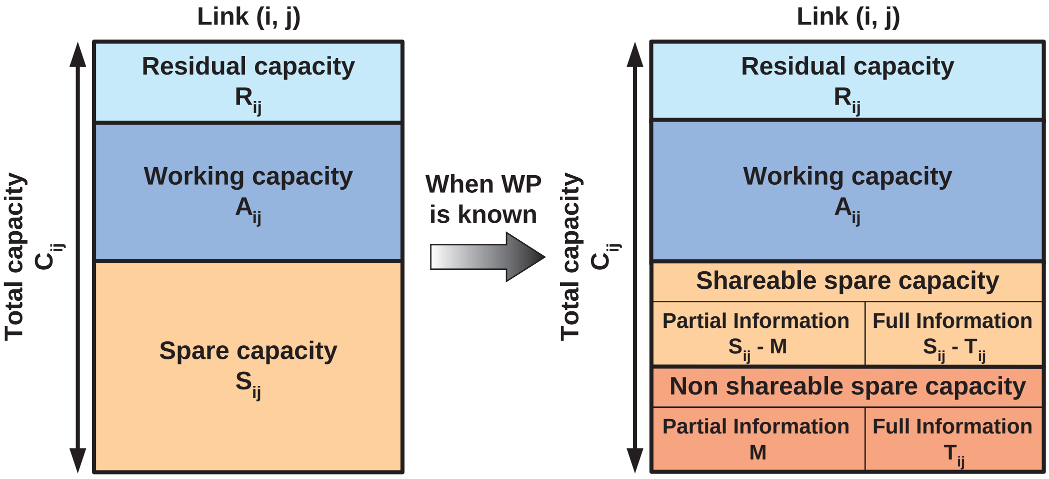


Figure 30: Illustration of categories of capacity along link (i,j).

3.2.3.2 Partial Information Algorithm

A routing algorithm with partial knowledge of the network information was presented by Kodialam [69] and is referred as the Partial Information Routing (PIR) algorithm. The basic idea behind PIR is to weight each link using an estimate of the additional capacity that needs to be reserved if a particular backup path is selected. After the working path

WP is selected, the source node computes the maximum working capacity M over all links along the working path, i.e. $M = \max\limits_{(i,j)\in WP} A_{ij}$. Then the source node assigns a weight to each link in the network:

$$w_{ij} = \begin{cases} \min(b, M + b - S_{ij}) \cdot W_{ij} & \text{if } M + b - S_{ij} > 0 \text{ and } (i,j) \notin WP \\ \varepsilon & \text{if } M + b - S_{ij} \leq 0 \text{ and } (i,j) \notin WP \\ \infty & \text{if } (i,j) \in WP \end{cases} \tag{13}$$

where b is the amount of the required capacity and ε is a small constant. The use of the ε instead of 0 is to guarantee that the minimum hop path will be chosen if multiple paths have $w_{ij} = 0$. The source node then applies a shortest path algorithm to select the backup path BP, that minimizes $\sum\limits_{(i,j)\in BP} w_{ij}$. The goal of the term $\min(b, M + b - S_{ij})$ is to capture the amount of additional spare capacity needed if, upon failure of the working path WP, the connection is routed over a backup path that contains link (i, j). This scheme requires very little additional information compared with the shortest disjoint path algorithms and, in general, chooses the links with the largest reserved spare capacity to avoid increasing S_{ij}. However, the estimator M assumes that when a failure occurs along WP, all the connections that route over the failed link would indeed reroute onto link (i, j) and is, therefore, a pessimistic estimation of the restoration capacity needed due to link failures along WP. In fact, to minimize spare capacity, the backup paths should be spread around to share spare capacity throughout the network. This approach may overestimate the capacity that needs to be reserved in some links.

3.2.3.3 Full Information Routing Algorithm

Given more accurate information about the additional spare capacity that needs to be reserved on each link, a backup path that better minimizes $\sum S_{ij}$ may be selected. To do this requires more information and such reservation schemes are known as Full Information Routing (FIR) algorithms. One such FIR scheme [68] operates by after selecting the working path WP, the source node collects the array $T_{ij}, (i, j) \in E$, where T_{ij} is the maximum capacity needed on link (i, j) if any link along WP fails. This computation is based on the network state before the new restoration connection is routed. The source node then assigns a weight to each link in the network:

$$w_{ij} = \begin{cases} \min(b, T_{ij} + b - S_{ij}) \cdot W_{ij} & \text{if } T_{ij} + b - S_{ij} > 0 \text{ and } (i,j) \notin WP \\ \varepsilon & \text{if } T_{ij} + b - S_{ij} \leq 0 \text{ and } (i,j) \notin WP \\ \infty & \text{if } (i,j) \in WP \end{cases} \tag{14}$$

Then Dijkstra's algorithm is used to select the backup path BP using these weights. The collection of the array T_{ij} is done during signaling exchanges, with flooding link state information [68].

3.2.3.4 Trap Topology Problem

Both information routing algorithms, PIR and FIR, have two distinct steps. Hence, they may also encounter the trap topology problem [64] (see Section 3.2.1). In a trap topology, a link disjoint path pair exists between the source and destination nodes, but it is possible that none of the algorithms can find a backup path. The problem arises because of the two-step nature of these algorithms, where each selects the shortest service path without considering the goal of subsequently selecting a link/node disjoint restoration path. A mincost max-flow algorithm may be used to avoid this dilemma. However, since trap topologies are generally rare in real networks [45], this is not further considered in this thesis.

3.2.4 Segment Backup Routing Algorithms

The algorithms reviewed previously, use the global backup path method to protect the whole working path. Hence, when a failure occurs, the source node is the responsible of switching over the traffic to the backup path. As described in Section 2.5, the global backup path method, due to the fault notification time, is one of the slowest methods. In order to reduce the recovery time, recent proposals are found where the notification distance is considered in order to reduce the recovery time [34, 67, 70]. These proposals take into account segment backup path methods and they focus on the backup path computation once the working path is given. Although, two proposals [34, 70] consider sharing the spare capacity, such sharing is only considered when the backup path is established and not during the backup path computation. Therefore, a reduction of the recovery time is achieved, but efficient use of the spare capacity is not guaranteed.

3.2.4.1 Backtracking Distance

A proposal to limit the recovery time by means of the called *backtracking distance D* has been presented [70]. Here, three specific cases are considered:

1. No backtracking, $D = 0$. Local backup path is used to protect the working path links. This case provides the best recovery time although it requires significant network resources.

2. Limited backtracking, $D = k$. In this case, the backup path can originate at a node on the working path up to k hops away from the node that detects the failure.

3. Unlimited backtracking, $D = \infty$. This case may result in end-to-end (global) backup paths.

Two-step approaches were used to compute the working and backup path. The working path was computed by applying the Widest Shortest Path (WSP), while the paper focused on developing the backup path routing algorithm: the Backtracking Distance Algorithm detailed in Algorithm 2.

Algorithm 2 Backtracking Distance

```
**INPUT**
$s$: source node;
$d$: destination node;
$b$: capacity requested;
$G = (V, E)$: network graph;
**ALGORITHM**
$WP = WSP(G, s, d, b)$
if ($WP$ == NULL) then
  return reject
end if
if ($D = 0$) then
  $G' = modified\_directed\_Steiner(G, WP, s, d, b)$
else if $D = \infty$ then
  $G' = Suurballe\_Algorithm(G, WP, s, d, b)$
else if $D = k$ then
  $G' = modified\_Suurballe\_Algorithm(G, WP, s, d, b, k)$
end if
if ($G'$ == NULL) then
  return reject
end if
$postprocessing(G')$
```

For $D = 0$, the subgraph G', such that there is a backup path with no backtracking if any link in the working path WP fails, is computed. In this case, the approximation directed Steiner tree algorithm [71] is used for this computation. In the case of $D = \infty$, Suurballe's algorithm is used to find the link disjoint path pair (see Section 3.2.1). Finally, for $D = k$, the backup subgraph G' is computed by means of the modified Suurballe's algorithm as shown in Algorithm 3.

Algorithm 3 modified Suurballe

```
**INPUT**
$s$: source node;
$d$: destination node;
$b$: capacity requested;
$G = (V, E)$: network graph;
$WP$: working path;
**ALGORITHM**
$BP = Suurballe(G, WP)$
if ($BP ==$ NULL) then
  return reject
end if
if $maxD(WP, BP) > k$ then
  $G_a = compExtraPaths(G, WP, BP, k)$
  if $G_a ==$ NULL then
    return null
  else
    $G' == BP \cup G_a$
  end if
end if
return $G'$
```

First, a backup path BP for the working path WP is computed. The cost of the computed BP is the lower bound on the cost of any solution for the $D = k$ case. Procedure $maxD$ checks the maximum backtracking distance of the BP and, if it exceeds k, additional paths, G_a, is added to the edges of BP.

3.2.4.2 PROMISE algorithm

A dynamic programming heuristic for a shared segment based protection approach called PROMISE was introduced [34]. This approach has polynomial time complexity. This heuristic tests a subset of all possible segments of the working path WP. Number nodes along the WP from 0 to H, where H is the number of hops along the given WP. In addition, let Θ_m be the *best* way to protect the part of WP from node m to node H (destination) by potentially dividing the WP into multiple working segments and protecting them with one or more BPs. The algorithm starts by first determining the BP for the last hop ($m = H-1$) and ends when the entire WP is protected ($m = 0$). The pseudocode of this dynamic programming based algorithm is presented in Algorithm 4 and denoted as PROMISE. This algorithm picks either the initial solution for Θ_m, or the combination of a previously determined Θ_i and $BP_{m,i}$, whichever results in a lower cost (e.g. bandwidth consumption).

Algorithm 4 PROMISE

```
for $m = H - 1$ to $0$ do
  Initialize $\Theta_m$ to a $BP$ from nodes $m$ to $H$
  for $i = m + 1$ to $H - 1$ do
    Let $BP_{m,i}$ be a $BP$ to protect nodes $m$ to $i$
    $\Theta_m \leftarrow \min(\Theta_m, Combine(\Theta_i, BP_{m,i}))$
  end for
end for
return $\Theta_0$
```

3.2.4.3 Cascaded Diverse Routing

Another heuristic algorithm called Cascaded Diverse Routing (CDR) was introduced [67] to perform survivable routing for shared segment protection. The novelty of this heuristic algorithm is the effort of predefining a set of candidate switching/merging node pairs (PSL and PML respectively, see section 2.1.1) and the adoption of the Iterative Two-Step-Approach (ITSA) algorithm [72]. The CDR algorithm contains the following steps:

1. Selection of the shortest M backup paths from the k-shortest paths in terms of hop count for each node pair.

2. Define a series of PSL-PML pairs along each backup path with a fixed distance D.

3. As a connection request arrives, the ITSA algorithm is invoked upon a set of PSL-PML pairs along a backup path. ITSA is then iteratively performed along each alternate path until a feasible solution is reached.

Steps 1 and 2 can be performed before the connection request arrives (or off-line). The distance D between PSL_i and PML_i in terms of hop count is called the *diameter* of protection domain i.

The cost function used in ITSA for the working path computation is:

$$w_{ij} = \begin{cases} \infty & \text{if } b > R_{ij} \\ b \cdot W_{ij} + \varepsilon & \text{otherwise} \end{cases} \tag{15}$$

For the backup path computation the cost function is expressed as:

$$w_{ij} = \begin{cases} b \cdot W_{ij} \cdot \left(1 - \frac{S_{ij}-T_{ij}}{b}\right) + \varepsilon & \text{if } S_{ij} - T_{ij} + R_{ij} \geq b > S_{ij} - T_{ij} \\ \varepsilon & \text{if } S_{ij} - T_{ij} \geq b \\ \infty & \text{if } S_{ij} - T_{ij} + R_{ij} < b \end{cases} \qquad \forall (i,j) \in E, (i,j) \notin WP \tag{16}$$

3.3 Simulation Model

A discrete-event simulation-based platform has been implemented in Java to model and evaluate the algorithms in distributed mesh topologies. The network topologies adopted in this work are the NSF, European and KL networks. For more details about each network see Appendix A. The values obtained for each network are the mean values obtained after 10 simulation runs (each run over a window of 10.000 LSPs) and for a confidence level of 95%. A confidence interval of at most 1% has been achieved. For clarity, the lower and upper values defined by the confidence interval have not been plotted. The simulations assume that traffic demands between all source and destination nodes are the same.

Algorithms are evaluated in two different simulation basis:

- *Incremental traffic simulation-based.* Requests arrive one by one and request holding time is assumed long enough, thus, for this set of simulations the accepted requests do not leave. With this assumption, the impact of different request holding times on LSPs and backup LSPs cannot be investigated. This is in fact to the advantages of Kodialam's algorithms as they have ignored such impact [73].

- *Limited holding time simulation-based.* Requests for LSP setup follow a Poisson distribution. The holding time of each source and destination node pair is considered to follow an exponential distribution.

According to the traffic classification presented in Section 2.4, fifty percent of the requests have low reliability requirements (LR), 40% medium reliability requirements

(MR) and 10% high reliability requirements (HR). The reliability requirements are only taken into account by those schemes that offer different levels of reliability.

3.3.1 Single and Multi-Layer Simulation Scenarios

3.3.1.1 Single-Layer Simulation-based

For the evaluation of algorithms oriented to a single-layer, i.e. IP/MPLS, the logical topology where LSPs are routed is given and there is not information about protected and unprotected lightpaths. The LSP capacity is uniformly distributed to 10, 20 or 30 Mbps. In the dedicated backup case, the allocated backup capacity is the same that the capacity of the primary LSP. In the shared backup case, the allocated backup capacity depends on the shareable network resources (see Section 3.2.3).

3.3.1.2 Static Multi-Layer Simulation-based

For this network scenario, the logical topology where the LSPs are routed is given. Some of the lightpaths are assumed to be already protected at the lower layer and remain the same during the simulation. Thus, routing algorithms for IP/MPLS-based networks may avoid the reservation of spare capacity to protect lightpaths that are already protected.

3.3.1.3 Dynamic Multi-Layer Simulation-based

Each fiber is assumed to have the same number of wavelengths, w. The transmission speed of each wavelength is set to 10 Gbps. The number of PSC ports, p, is assumed to be the same in each node. The required LSP capacity is set to 500 Mbps unless specifically stated otherwise. When an existing lightpath does not accommodate any LSP, the lightpath is disconnected.

3.3.2 Figures of Merit

The metrics of interest for evaluating the algorithm performance are:

- *Request rejection ratio.* This value corresponds to the ratio of rejected requests over the whole network.

- *Backup links average.* This value represents the average number of backup path links used per working path.

- *Working links average.* This value represents the average number of working path links.

- *Fault notification time.* This value is expressed in terms of fault notification distance: number of hops or physical length (see Section 2.2.5).

- *Restoration overbuild.* This value corresponds to the average of the total spare capacity, S_{ij}, and working capacity, A_{ij}, over the whole network: $\frac{\sum_{(i,j)\in E} S_{ij}}{\sum_{(i,j)\in E} A_{ij}}$.

3.4 Performance Evaluation

In this section, the different objective functions reviewed in this chapter are evaluated using the simulation model presented in Section 3.3 under single-layer simulation scenario for both incremental traffic and, then, for limited holding time cases. First, the performance of the traditional QoS routing algorithms, the Widest Shortest Path (WSP) and the Min Hop Algorithm (MHA) are compared, when used to compute the backup path. Next, the protection techniques presented in 2.1, i.e. global and local backup path methods, are also considered in order to analyze the tradeoff between recovery time and resource consumption. After analyzing the traditional QoS routing algorithms, QoSP routing algorithms are also evaluated and compared to traditional QoS routing algorithms. Full and partial information routing concepts are considered in both global and backup path method under incremental traffic.

3.4.1 Traditional QoS Routing Evaluation

3.4.1.1 QoS Restorable Routing Algorithms

The aim of this simulation is to evaluate the amount of network resources used when traditional QoS routing algorithms, the Widest Shortest Path (WSP) and the Min Hop Algorithm (MHA), are only considered. Dedicated and shared mechanisms are considered for recovery resources analysis, though no accurate network information is available for backup path computation. This simulation also compares and evaluates the local and global backup path methods for recovery time analysis. Thus, four restorable routing schemes are evaluated (see Table 8):

- *Min Hop Algorithm under Dedicated protection* (MMD). This algorithm uses MHA to compute both working and backup paths. The backup capacity is dedicated.

- *Min Hop Algorithm under Shared protection* (MMS).This algorithm uses MHA to compute both working and backup paths. The backup capacity is shared.

- *Widest Shortest Path algorithm under Dedicated protection* (WWD). This algorithm uses WSP to compute both working and backup paths. The backup capacity is dedicated.

- *Widest Shortest Path algorithm under Shared protection* (WWS). This algorithm uses WSP to compute both working and backup paths. The backup capacity is shared.

Table 8: Shared vs. Dedicated Traditional QoS Routing Schemes.

Routing Algorithm	Working Path	Backup Path	Protection Technique
MMD	MHA	MHA	Dedicated
MMS	MHA	MHA	Shared
WWD	WSP	WSP	Dedicated
WWS	WSP	WSP	Shared

The metrics of interest for this simulation are: the request rejection ratio, the backup links average, the working links average and the recovery time.

3.4.1.2 Global vs. Local Protection for Traditional QoS Routing Algorithms under Incremental Traffic

As shown in Figs. 31a and 31b, when global/local backup methods are applied, the schemes that reserve dedicated capacity, MMD and WWD, offer the larger blocking probability due to the amount of spare capacity needed to protect the whole working path. Although shared capacity is not taken into account during the backup path computation, intra-demand and inter-demand sharing capacity is considered at the point (instance) of the backup resource allocation by both MMS and WWS algorithms explaining the best result in terms of request rejection ratio. Moreover, global backup path method outperforms local backup method towards all network topologies. For instance, in the case of NSF, 20% of requests are rejected when up to 500/1500 connections have been fulfilled by dedicated and shared local protection respectively (see Fig. 31b). On the other hand, dedicated and shared global protection offer a request rejection ratio of 0.1 when up to 1000/1500 connections have been treated, respectively as shown in

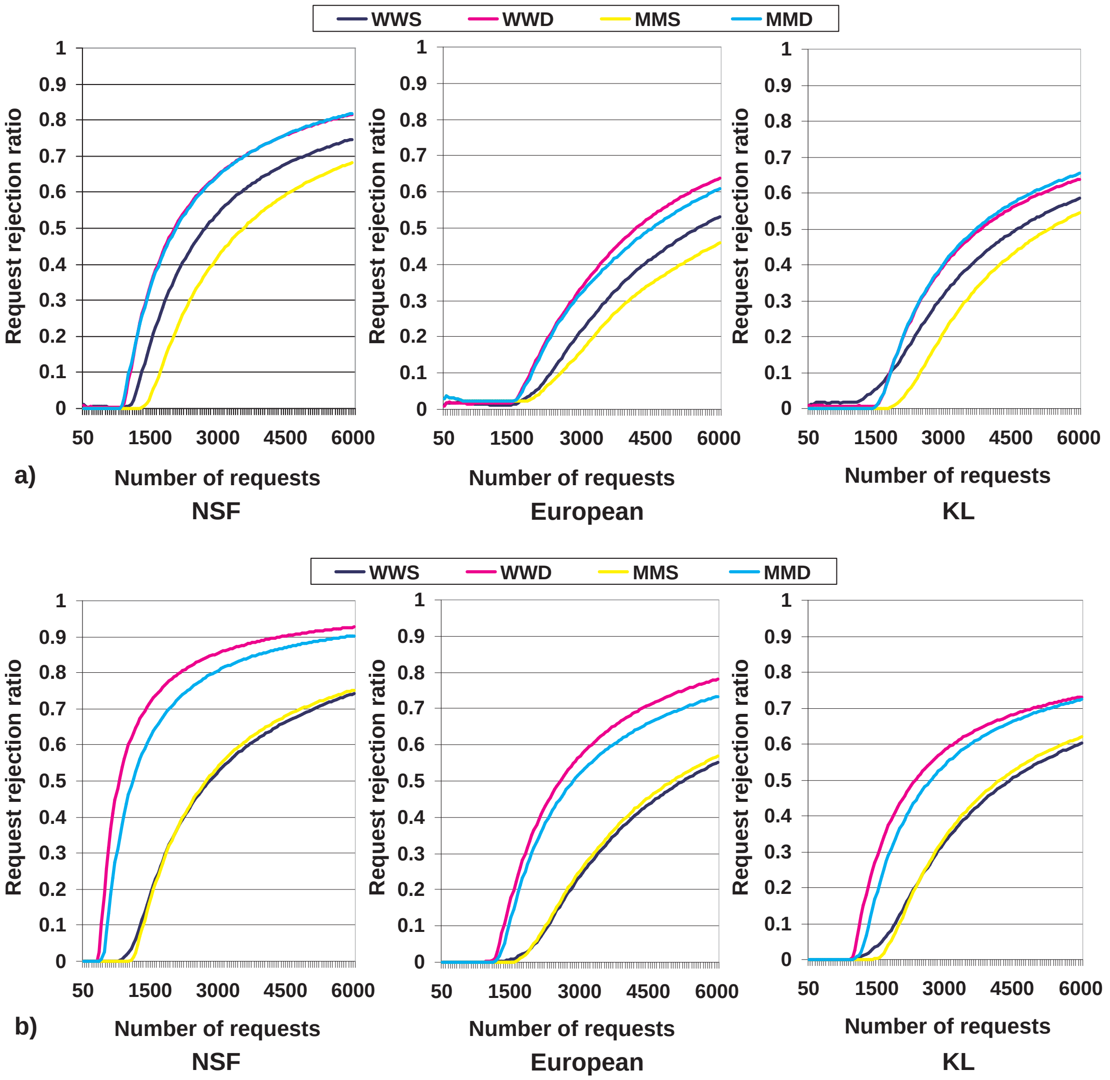


Figure 31: Request rejection ratio when a) global and b) local backup path method is applied for each network topology.

Fig. 31a. Additionally, the MMS outperforms the WWS in Fig. 31a as a consequence of the path computation. Since the MMS selects the min hop path, fewer links must be protected and the number of backup path links is reduced. This is shown by means of the backup link ratio and working link ratio in Figs. 32 and 33 respectively. Note that because of the incremental traffic assumption, results are stabilized for each algorithm after around 1500 and 2000 attempted requests; after this point, most of the incoming requests are rejected.

As shown in Fig. 32 the routing schemes that compute the backup path by means of the WSP, i.e. WWS and WWD, incur large backup path lengths (hops). In the local backup path case, this difference dramatically increases from around 7 hops to around

9 hops in the NSF network topology. Figure 32 also shows the length difference of the backup path in regards to the backup path method used.

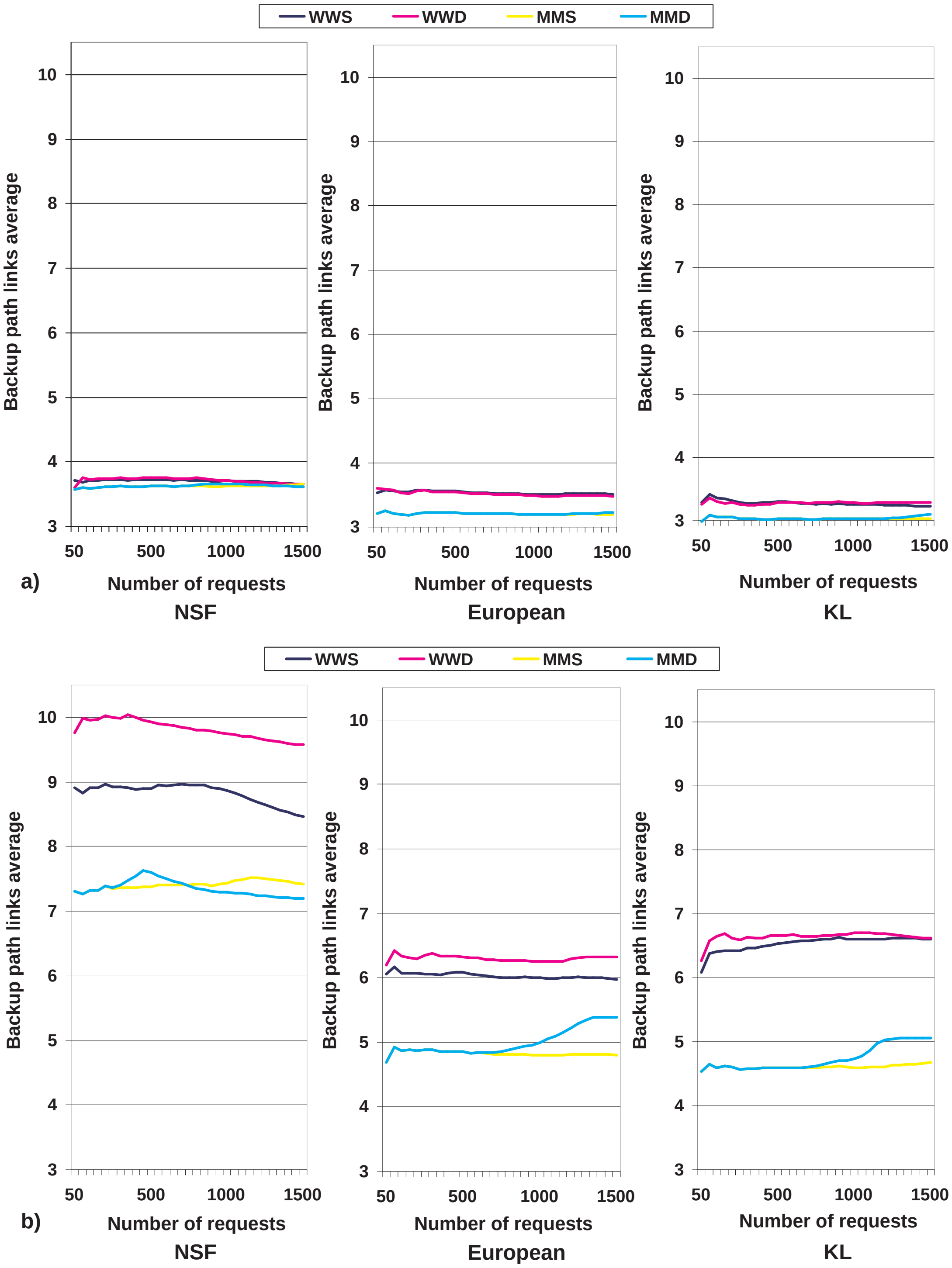


Figure 32: Backup path links average when a) global and b) local backup path method is applied for each network topology.

As expected, the local method is worse in terms of resource consumption because backup paths have a length average of $7 - 10$ hops in the NSF topology and 4-7 in both European and KL topologies. On the other hand, the global scheme yields a backup path length average of $3 - 4$ hops in all the topologies. Figure 33 illustrates the larger working path length when WWS and WWD algorithms are used leading to a higher request rejection ratio as explained previously. Note that the shape of all the curves is slightly going down as the number of attempted requests increases because less network resources are available and only the requests that need low network resources (network links) are accepted, reducing the average of working and backup links.

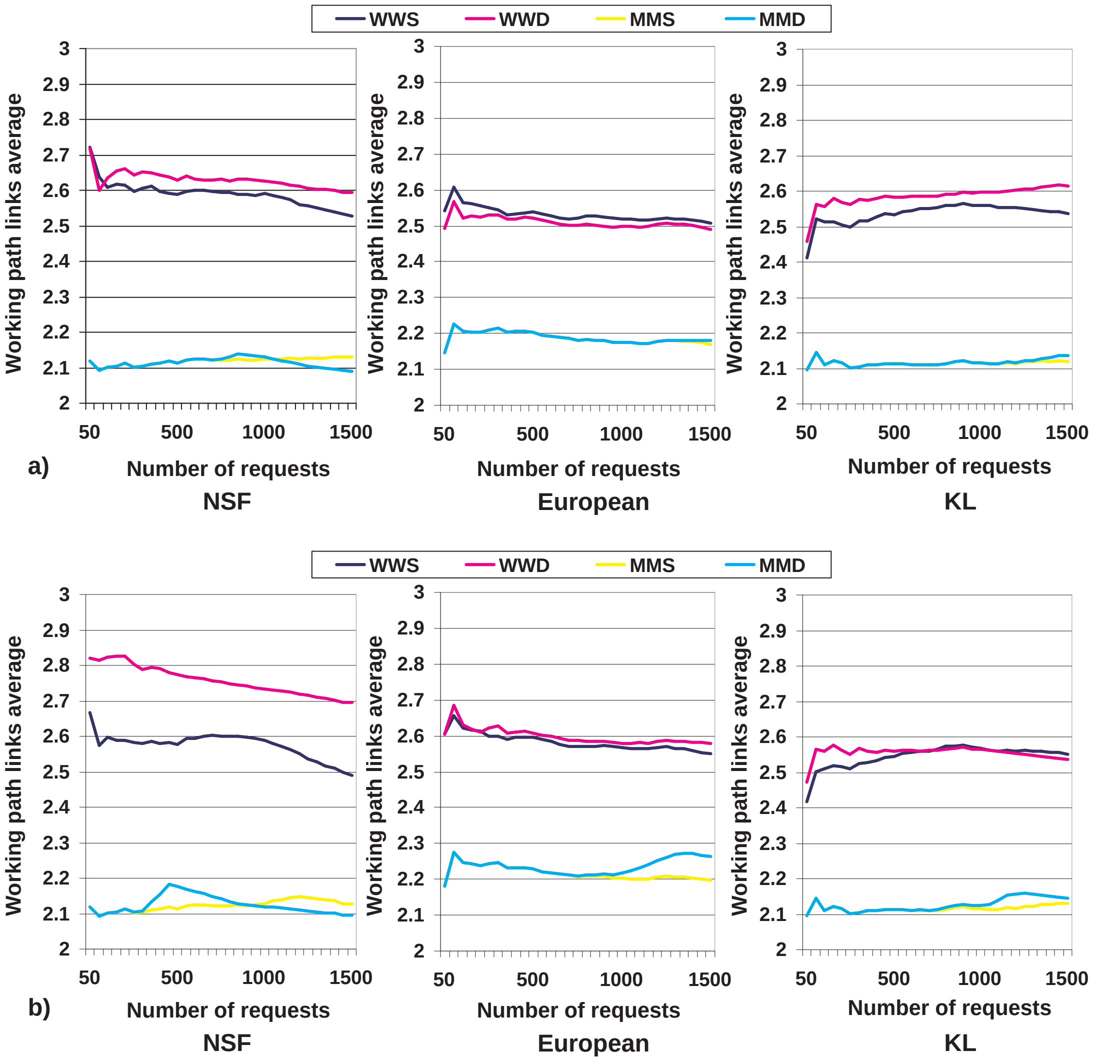


Figure 33: Working path links average when a) global and b) local backup path method is applied for each network topology.

It can be seen that the local backup path method results in high resource consumption. However, it has fast recovery time ($\simeq 0$) as discussed in chapter 2. Although the

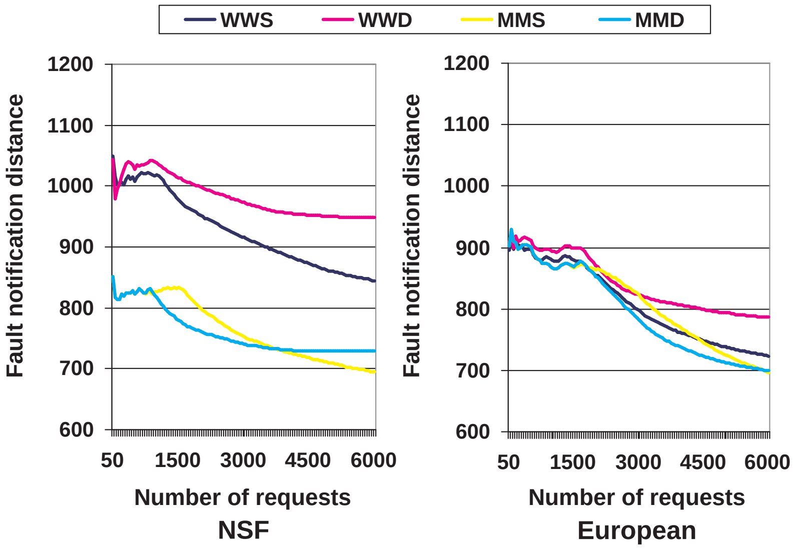


Figure 34: Fault notification time in terms of fault notification distance for NSF (miles) and European (km) topologies.

global approach offers better performance in terms of resource usage, its recovery time in terms of notification distance is high as shown in Fig. 34. As expected, the fault notification distance is lower for MMS and MMD algorithms according to the results shown above. The fault notification distance is slightly decreasing when the number of attempted requests increases because less network resources are available and requests that need low network resources (links) have more probability of being accepted. Thus, the fault notification distance average is reduced. This behavior will be appreciated in the rest of the simulations that consider incremental traffic case.

3.4.1.3 Traditional QoS Routing Algorithms under Limited Holding Time

In this section, WWS and WWD are evaluated in different load network scenarios. Figure 35 shows the request rejection ratio for low, medium and high network loads. As expected, the number of rejected requests increases when the load increases. Moreover, KL network shows similar behavior when WWS and WWD are applied. Thus, the KL topology does not share high level of network resources.

3.4.2 Shared Backup Path Methods Evaluation

3.4.2.1 QoS Restorable Routing Algorithms

The aim of this section is to investigate the effectiveness of QoS routing algorithms that require explicit network information. In these experiments, schemes based upon Full

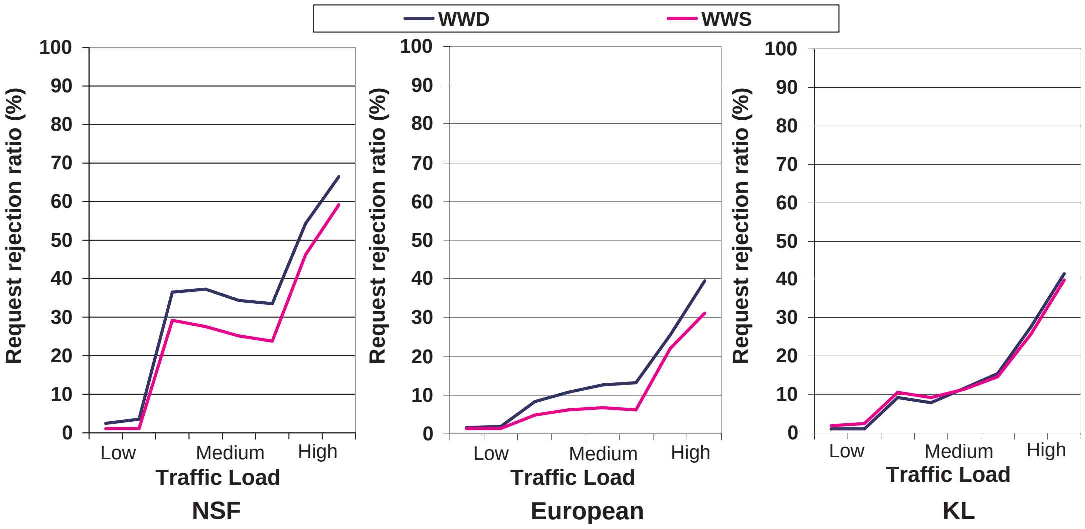


Figure 35: Request rejection ratio for different network loads.

Information Routing (FIR) and Partial Information Routing (PIR) schemes are evaluated through a series of simulations based experiments. Although these algorithms have been already defined as global backup path methods, in the presented experiments these algorithms are also evaluated under local backup path methods for recovery time analysis. Thus, four restorable routing schemes based on the network information available and the recovery scheme applied, are evaluated (see Table 9):

- *Full Information Routing* (FIR). This algorithm uses WSP to compute the working path and the FIR algorithm to compute the global backup path.
- *Partial Information Routing* (PIR). This algorithm uses WSP to compute the working and the PIR algorithm to compute the global backup path.
- *Full Information Routing with Fast Protection* (FIRFP). In this case, the working path is protected by local backup paths.
- *Partial Information with Fast Protection* (PIRFP). In this case, the working path is protected by local backup paths.

For this set of routing schemes, only shared protection is applied. The performance of the *best* traditional QoS routing schemes modelled and presented in Section 3.4.1, i.e. the MMS scheme with global backups, will be also plotted in order to demonstrate the relative enhancement of the new routing schemes.

Table 9: Routing Schemes for Routing Information Evaluation.

Routing Algorithm	Working Path	Backup Path	Backup path method
FIR	WSP	FIR	Global
PIR	WSP	PIR	Global
FIRFP	WSP	FIR	Local
PIRFP	WSP	PIR	Local

The metrics of interest for this simulation are the request rejection ratio, the restoration overbuild and the recovery time.

3.4.2.2 Simulation Results for Routing Network Information Algorithms under Incremental Traffic

As shown in Fig. 36, MMS algorithms offer the largest request rejection ratio compare to FIR and PIR because the shareable spare capacity is not considered when the backup path is computed. However, when fast protection is applied, PIRFP and FIRFP, two different scenarios are presented: 1) when the network load is low (up to 2000 requests) 2) when the network load is medium/high (more than 200 requests). In the first case, both PIRFP and FIRFP routing algorithms yield the poorest request rejection ratio. However, for medium/high loaded network, FIRFP experiences an improvement, reducing its request rejection ratio close to that of FIR and PIR schemes. Thus, it can be concluded that full information routing algorithm with fast protection is efficient in terms of request rejection ratio and recovery time (≃ 0) for medium/high network loads. The drawbacks of FIRFP are 1) the technology required in each node, i.e. all

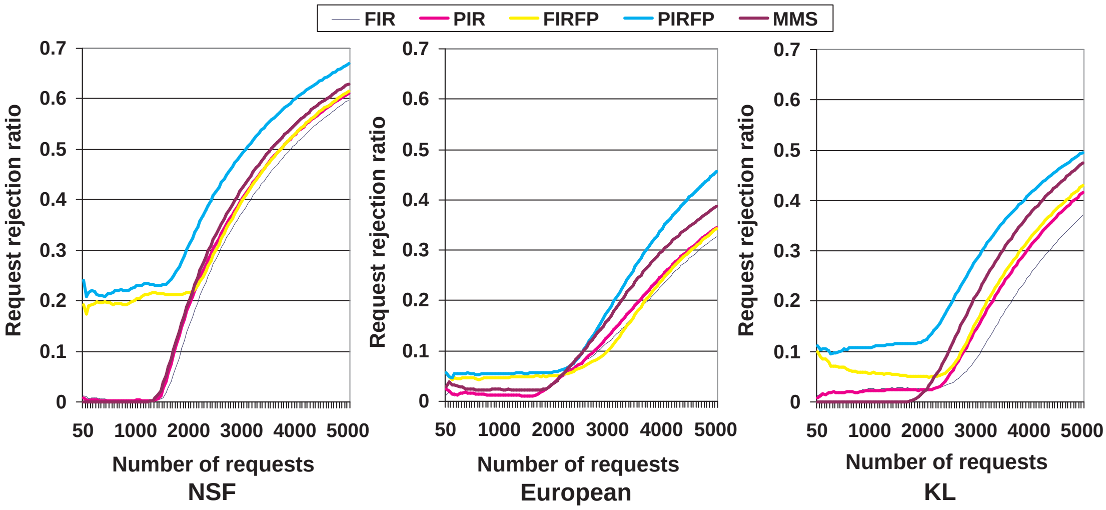


Figure 36: Request rejection ratio for each network topology.

network nodes must have PSL and PML functionalities and b) full network information is required at each node. Note that, there is a slight improvement when FIR algorithm is compared with PIR. Thereby, PIR could be a better choice than FIR, since all routing information is not always available, simplifying the management of the network, i.e. minimizing the flooding actions.

In terms of restoration overbuild, as expected from the request rejection ratio results, the routing algorithm that exhibits the poorest level of sharing is the PIRFP followed by the MMS (see Fig. 37). This demonstrates that, although partial information aims to share the spare capacity, this routing algorithm is only suitable when a global backup path method is used.

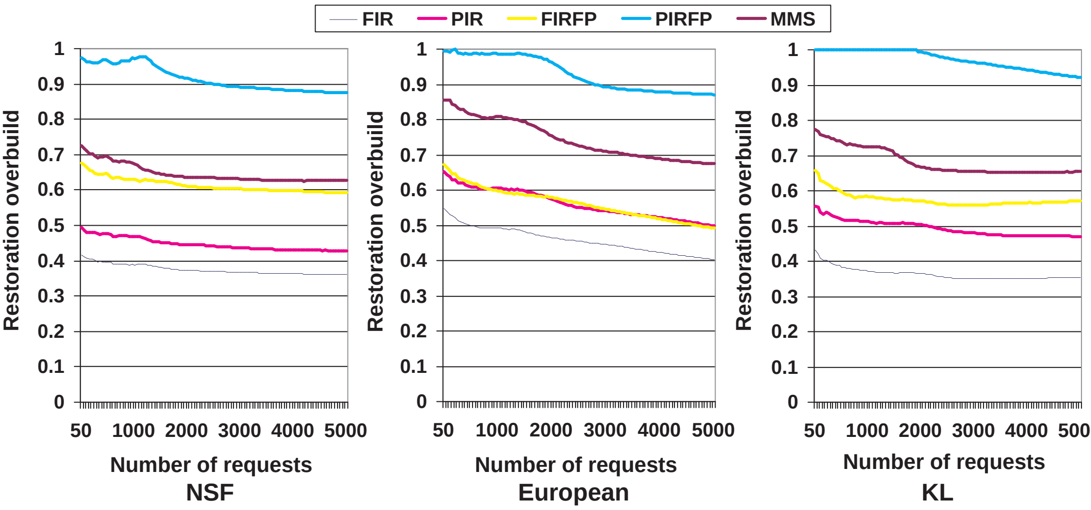


Figure 37: Restoration overbuild for each network topology.

Fig. 38 again shows the trade off that exists between resource consumption (expressed as rejected rejection ratio) and recovery time (expressed as fault notification time). Although FIR and PIR algorithms result in better resource consumption in terms of, they perform worst in terms of recovery time compare to MMS because MMS selects shortest paths.

3.4.2.3 Simulation Results for Routing Network Information Algorithms under Limited Holding Time

As shown in Fig. 39, PIR and FIR algorithms offer the largest request rejection ratio compare to FIRFP and PIRFP for KL and European network because the global backup strategy shares less network resources.

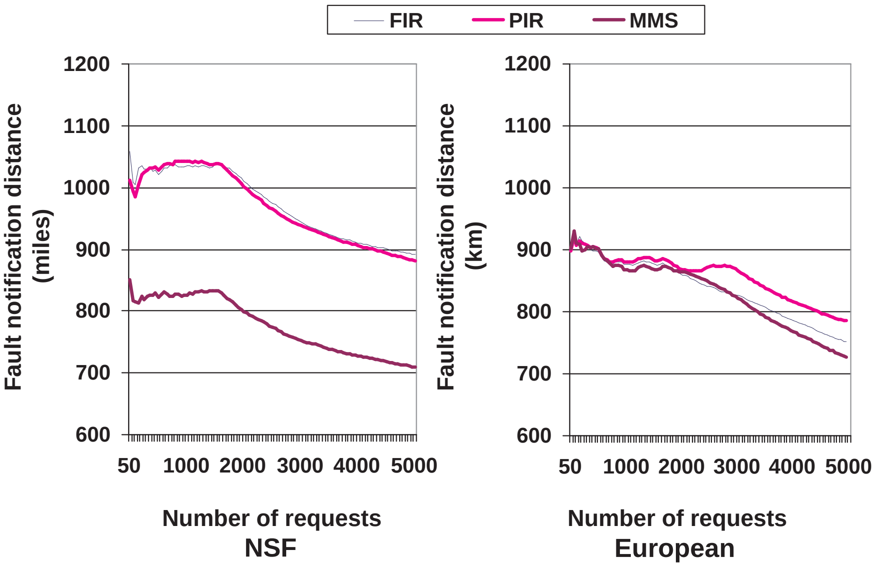


Figure 38: Fault notification distance for each network topology.

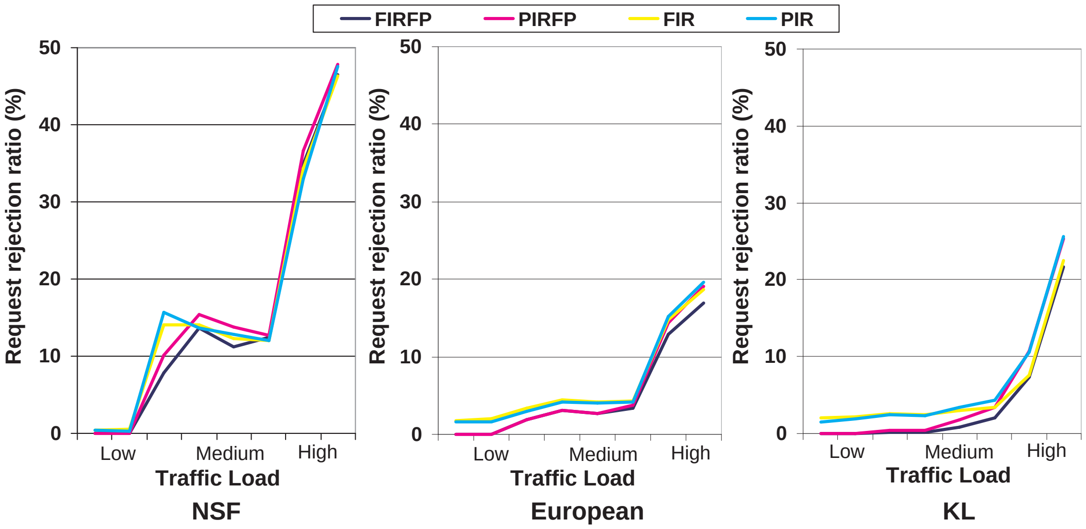


Figure 39: Request rejection ratio for different network loads.

It may be concluded from these simulations that, for low network load, if the objective is to minimize the request rejection ratio then the PIRFP algorithm should be selected to simplify network management. On the other hand, for medium network loads, FIRFP algorithm should be selected since it offers similar request rejection ratio to PIRFP coupled with fast recovery time.

3.5 Concluding Remarks

Shortest link disjoint path algorithms are simple and may be deployed using minimum link state information which can be provided by current OSPF/IS-IS extensions. However, they do not consider capacity sharing as an objective in the backup path selection algorithm and, consequently, poor resource utilisation may result with high levels of spare capacity being reserved. Thus, high spare capacity is reserved. Additionally, although the current minimum interference restorable routing (MIRR) algorithm aims to reduce the block probability, dedicated protection (1+1,1:1) is only proposed. Thus, a large amount of spare capacity is used.

Significant reductions in spare capacity can be achieved by sharing this capacity. The accuracy and performance of the shared backup routing algorithms are based on the available network information to select the backup path. Partial Information Routing (PIR) and Full information routing (FIR) schemes significantly reduce the spare capacity. However, these schemes have been proposed for path protection (global backup path method). Consequently, they suffer from extended recovery times. Segment backup path method has been selected by recent research to guarantee fast protection. However, most of these schemes do not take into account spare capacity in their routing objectives.

All these routing schemes have been evaluated into a single network layer (IP/MPLS). The next chapter proposes new QoSP routing algorithms into the static multi-layer network scenario. Given the logical topology where some logical links are optically protected, the proposed QoSP routing algorithms avoid protection duplications.

CHAPTER IV

QOSP ROUTING IN THE STATIC MULTI-LAYER SCENARIO

In Chapter 3, some of existing QoS and QoSP routing schemes have been reviewed and evaluated according to their routing metrics to compute the working and backup path:

- Shortest path.
- Load balancing.
- Minimum interference.
- Network information available.
- Recovery time reduction.

These algorithms are oriented towards a single network layer, i.e. the IP/MPLS network layer. Thus, no information about the lower layers is taken into account. In this chapter, although the logical topology where the LSPs are routed is given, some links (lightpaths) are assumed to be already protected at the lower layer. Based on this static information, an enhancement of the QoSP routing algorithms for IP/MPLS-based networks may be achieved by avoiding the need to protect those lightpaths already protected at the optical layer. Thereby, static information from the lower layers is included in the metrics of the proposed routing schemes. Avoiding protection duplication is achieved by using link failure probabilities (section 2.3.4). The links protected by lower layers have a link failure probability equal to 0. Consequently, links with failure probability equal to 0 do not need to be protected again. In order to deploy these algorithms, a new definition of link disjoint path pair using Shared Risk Link Groups (SRLG) adapted to this network scenario is introduced. These schemes also encompass shared segment backup computation. The proposed schemes are then evaluated and compared to some of the reviewed QoSP routing algorithms in terms of recovery time, network resources and request rejection ratio. The research contribution on dynamic multi-layer scenario is covered in the next Chapter 5.

4.1 Partial Disjoint Paths

In this section, our novel proposal for fast protection and efficient use of the spare capacity [74] is described. The scheme attempts to enhance the QoSP routing algorithms for IP/MPLS-based networks. Thus, its main objectives are:

- Improving the spare capacity used by using shared backups.
- Reducing the recovery time by applying segment backup methods.
- Assuming knowledge of logical link (lightpaths) being protected by the lower layer (optical layer) in order to avoid protection duplications.

Segment protection and shared backups are combined, resulting in a faster fault recovery time and better resource consumption.

4.1.1 Avoiding Protection Duplications

Although this proposal is IP/MPLS focused and the logical topology through which LSPs are routed is given, the lower layer information is also considered. Some of the logical links are already protected at the lower layer. The objective is to avoid protecting at the IP/MPLS level those logical links that are already optically protected. An example is shown in Fig. 40. The working path between node pair $(3, 2)$ does not need to establish a backup path because link $(3, 2)$ at the IP/MPLS layer, i.e. lightpath L_1 $(3 - 1 - 2)$ at the optical layer, is already protected by the backup lightpath BL_1 $(3 - 4 - 2)$. Thus, the multi-layer fault management is simplified and the resource consumption is reduced.

4.1.2 Fundamentals of Partial Disjoint Path

In this proposal, the logical topology used to route the working and backup paths is considered partially protected, i.e. the optical layer is partially protected. Thus, two types of links coexist at IP/MPLS network scenario:

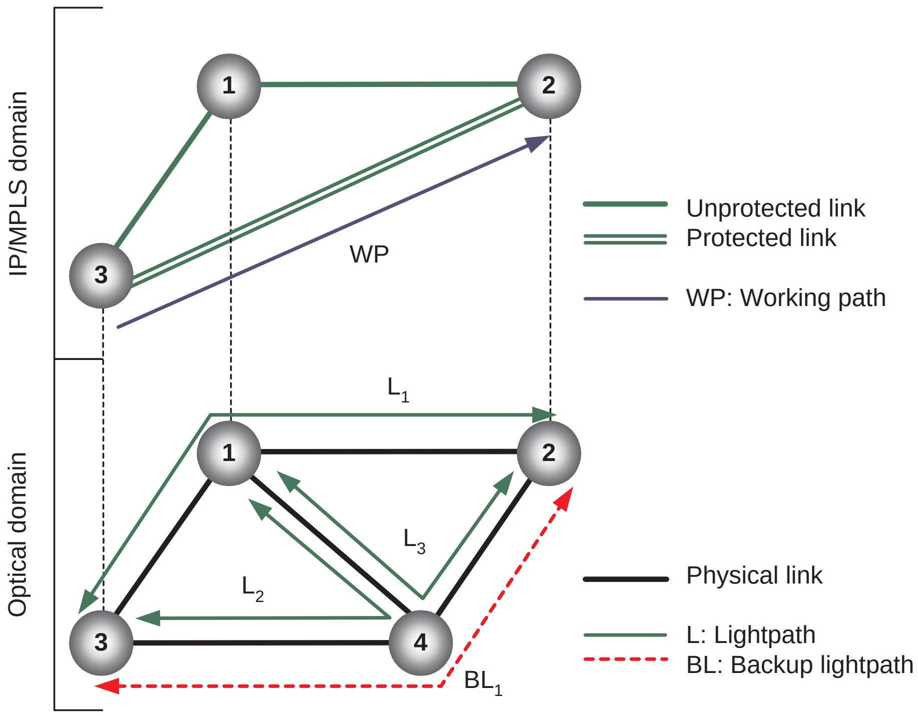


Figure 40: An example of a multi-layer protection scenario.

- *Protected links.* Logical links at the IP/MPLS network that are already protected by the lower layer (optical) recovery mechanisms. These lightpaths do not need extra recovery resources at the IP/MPLS layer.

- *Unprotected links.* Logical links at the IP/MPLS network that are not protected. Some protection mechanism must be offered at the IP/MPLS layer.

Therefore, no extra resources are necessary in the IP/MPLS layer to protect against failure of protected links at the optical layer. Once the working path (WP) is known, the backup path (BP) can be computed. The backup path is proposed to be a Partial Disjoint Path (PDP) since it may overlap the nodes of the WP and the links of the WP already protected at the optical layer. When the PDP overlaps with the WP, more than one backup path, i.e. segment backup paths (SBP), are established. Hence, when a PDP is computed, the optical protected links may:

- Belong to the protected segment path.

- Not belong to the protected segment path.

Both cases are shown in Fig. 41a and 41b respectively. In Fig. 41, two WPs are established sharing link $3-4$ that is already protected at the optical layer. The same PDP is used to protect both the WPs. In the first case a), the computed PDP overlaps WP_A and WP_B. This means that two segment backup paths (SBP_1 and SBP_2) are established between the protected segment paths $s-3$ and $4-d$ since link $3-4$ is already protected. Moreover, the SBP capacity is shared in both cases (Fig. 41a and 41b) since the shared link $3-4$ does not need to be protected at the IP/MPLS layer.

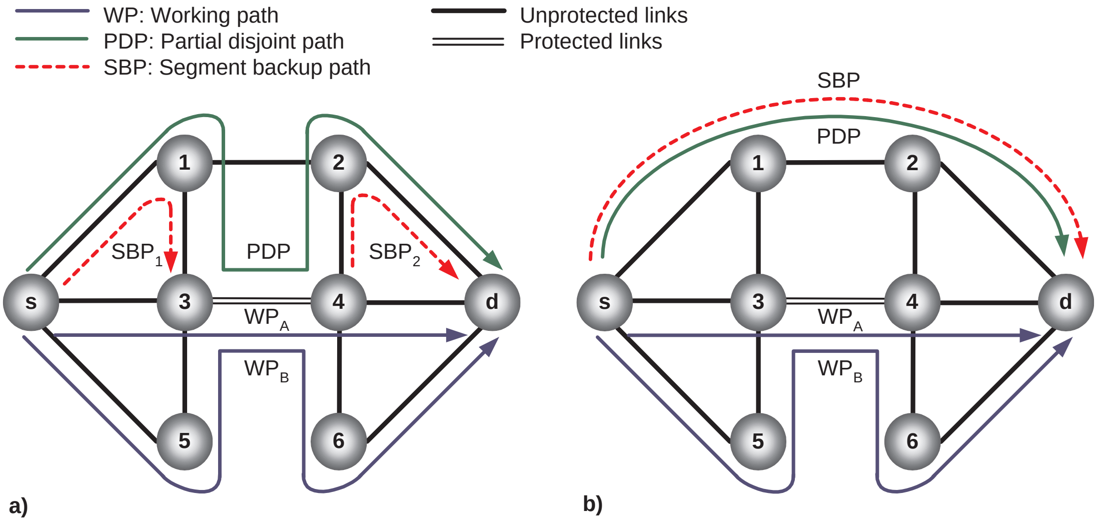


Figure 41: IP/MPLS protection when the partial disjoint path a) overlaps protected links b) does not overlap protected links.

According to the definition of link-disjoint path, based on Shared Risk Link Group (SRLG) [75], sharing the SBP capacity is not possible: *two data paths are link-disjoint if no two links on the two paths belong to the same SRLG.* As shown in Fig. 41b, both the WP_A and the WP_B belong to the same SRLG, since they share link $3-4$, thus backup path capacity cannot be shared. However, this link is already protected at the optical layer, and consequently, the SBP defined at the IP/MPLS layer is not activated against the failure of link $3-4$. Therefore, in this multi-layer network scenario considered in this research, the following modified definition is proposed: "**two data paths are link-disjointed if the links that are *unprotected* do not belong to the same SRLG**".

4.1.3 Partial Disjoint Path Algorithm

A Partial Disjoint Path (PDP) is computed in order to identify the segment backup paths necessary to protect the working path. The variable p_{ij} is defined according to the lightpath failure probability of the logical link (F^L_{ij}, Eq. 10) to identify the unprotected links:

$$p_{ij} = \begin{cases} 1 & \text{if } F_{ij}^{L} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

First, a weight w_{ij} is assigned on each link according to the following equation for full network information available:

$$w_{ij} = \begin{cases} 0 & \text{if } (i,j) \in WP \text{ and } p_{ij} = 1 \\ c_{ij} & \text{if } (i,j) \notin WP \text{ and } R_{ij} + S_{ij} - E \geq b \\ \infty & \text{otherwise} \end{cases} \tag{18}$$

Where E is the maximum capacity necessary if one of the unprotected WP links fails; c_{ij} is the cost assigned to link (i,j) according to FIR algorithm (Eq. 14); and p_{ij} identifies the unprotected links. Once the weight is assigned the PDP is computed. A variation of the Dijkstra algorithm called *Partial Disjoint Path* (Algorithm 5) is used for this process.

In the case of partial network information, c_{ij} is the cost assigned to link (i,j) according to PIR algorithm (Eq. 13), and the weight w_{ij} is computed as follows:

$$w_{ij} = \begin{cases} 0 & \text{if } (i,j) \in WP \text{ and } p_{ij} = 1 \\ c_{ij} & \text{if } (i,j) \notin WP \text{ and } M + S_{ij} - E \geq b \\ \infty & \text{otherwise} \end{cases} \tag{19}$$

In this algorithm, $Cost(v)$ is a vector containing the path cost from s to v; $Pred(v)$ contains the v's predecessor node; and $WPlast(v)$ contains the last WP node visited before treating node v. Q represents the list of adjacent vertices which were not visited yet. Function $min\_cost(Q)$ returns the element $u \in Q$ with the lowest $Cost(u)$; and $adjacency(u, G)$ represents the adjacency list of vertex u. Once the PDP is computed, the BP links are identified. Only links of the PDP, which do not belong to the WP, are used as backup links. Other links are considered to be non-protected at IP/MPLS layer since they are already protected at the optical layer. The reserved capacity depends on the amount of capacity that may be shared in each backup link and the unprotected links.

Algorithm 5 Partial Disjoint Path

```
INPUT
$s$: source node;
$d$: destination node;
$G = (V, E)$: network graph;
$WP$: working path;
ALGORITHM
for all $v \in V$ do
  $Cost(v) = \infty$
  $Pred(v) = s$
  $WPlast(v) = s$
end for
$Cost(s) = 0$
$Q \leftarrow s$
while $(Q)$ do
  $u \leftarrow min\_cost(Q)$
  $Q = Q - \{u\}$
  for all $v \in adjacency(u, G)$ do
    if $(Cost(u) + w_{uv} < Cost(v)$ then
      if $v \in WP$ then
        $WPlast(v) = v$
      else
        $WPlast(v) = WPlast(u)$
      end if
      $Pred(v) = u$
      $Cost(v) = Cost(u) + w_{uv}$
      $Q \leftarrow v$
    end if
  end for
end while
```

4.2 Reliable Services with Fast Protection Routing

Another mechanism to enhance QoS routing performance is the use of the traffic-profile concept to characterise the sensitivity of a class of traffic to failures. Thus, the routing algorithm can act in different ways depending on the traffic class. Protection specific to different classes of traffic has the potential to reduce resource consumption and offers acceptable and appropriate QoS protection to the failed paths.

Our proposed scheme [76] extends previous schemes [74] that consider the application of shared segment protection for fast protection and efficient use of network resources. In this proposal, the needs of different traffic classes according to the classification presented in section 2.4, are taken into account. In addition, knowledge of the already protected links at the lower optical layer, to avoid protection duplications, is also considered. Additionally, this scheme also addresses the selection of the working path to minimize resource consumption; not considered in previous schemes [34, 67–70].

4.2.1 Working Path: The k-Minimum Interference Algorithm

In the proposed schemes, the routing algorithm used to compute the working path, attempts to minimize the resource consumption using minimum interference while also considering the links currently protected at the optical layer. The proposed k-Minimum Interference (KMI) routing algorithm selects the k-paths with minimum interference, using a variation of LMIR. The path is selected according to the traffic class t of the request:

- *HR*: the one with a low number of links to be protected is selected. Note that avoiding a high number of links to protect in the HR case, the number of backup paths with dedicated capacity is also reduced. This minimizes the resource consumption.
- *MR*: the one with minimum interference is selected.
- *LR*: the one with high number of links to protect is selected. Note that LR requests are partially protected, hence for this method, only those links that can reach a shared segment BP with sufficient capacity are protected.

4.2.2 Backup Path: Resilient Partial Disjoint Path Algorithm

A variation of the Partial Disjoint Path (PDP) algorithm to identify the segment backup paths necessary to protect the working path (see section 4.1) is proposed. First, a weight w_{ij} is assigned on each link:

$$w_{ij} = \begin{cases} 0 & \text{if } (i,j) \in WP \text{ and } p_{ij} = 1 \\ M & \text{if } (i,j) \in WP \text{ and } p_{ij} = 0 \text{ and } t = LR \\ c_{ij} & \text{if } (i,j) \notin WP \text{ and } (t = MR \text{ or } t = LR) \text{ and } R_{ij} + S_{ij} - E \geq b \\ c_{ij} & \text{if } (i,j) \notin WP \text{ and } t = HR \text{ and } R_{ij} \geq b \\ \infty & \text{otherwise} \end{cases} \tag{20}$$

Where M is a high value constant ($\neq \infty$) that facilitates the use, in the PDP algorithm, of the unprotected WP links when partial protection (LR) is considered; E is the maximum capacity necessary if one of the unprotected WP links fails; c_{ij} is the cost assigned to link (i, j) according to LMIR algorithm(section 3.1.3); and p_{ij} identifies the unprotected links. Note that, $c_{ij} < M$. Once the appropriate weight is assigned, the Resilient Partial Disjoint Path (RPDP) is computed (see Algorithm 6).

Algorithm 6 Resilient Partial Disjoint Path

```
INPUT
$s$: source node;
$d$: destination node;
$G = (V, E)$: network graph;
$WP$: working path;
ALGORITHM
for all $v \in V$ do
  $Cost(v) = \infty$
  $Pred(v) = s$
  $WPlast(v) = s$
end for
$Cost(s) = 0$
$Q \leftarrow s$
while $(Q)$ do
  $u \leftarrow min\_cost(Q)$
  $Q = Q - \{u\}$
  for all $v \in adjacency(u, G)$ do
    if $(Cost(u) + w_{uv} < Cost(v)$ then
      if $v \in WP$ then
        $WPlast(v) = v$
      else
        $WPlast(v) = WPlast(u)$
      end if
      if $distance(WPlast(u), WPlast(v)) < DIST(t)$ then
        $Pred(v) = u$
        $Cost(v) = Cost(u) + w_{uv}$
        $Q \leftarrow v$
      end if
    end if
  end for
end while
```

In the RPDP algorithm, Algorithm 6, $Cost(v)$ is a vector containing the path cost from s to v; $Pred(v)$ contains the v's predecessor node; and $WPlast(v)$ contains the last WP node visited before treating node v. Q represents the list of adjacent nodes which were not yet visited. Function $min\_cost(Q)$ returns the element $u \in Q$ with the lowest $Cost(u)$; $adjacency(u)$ represents the adjacency list of vertex u; $DIST(t)$ returns the maximum failure notification distance accepted by traffic class t; and $distance(x, y)$ is the maximum failure notification distance between nodes x and y of the WP. Once the PDP is computed, the BP links are identified. The links of the PDP, which do not belong to the WP, are the backup links. Other links can be considered as unprotected at IP/MPLS layer since they are either protected at the optical layer or are unprotected because they are partially protected at the IP/MPLS layer.

The reserved capacity depends on the amount of capacity that may be shared in each backup link and the links that are protected at IP/MPLS layer for the shared backup case. In the dedicated backup case, each backup link reserves b units of capacity.

4.3 Performance Evaluation

This section evaluates the routing schemes presented in this chapter using the simulation model described in Section 3.3. These schemes are evaluated under the static multi-layer simulation scenario for both incremental traffic and limited holding time cases. The metrics of interest to evaluate the algorithm performance are:

- *Request rejection ratio.*
- *Fault notification time.*
- *Restoration overbuild.*

Thirty percent of the network links are considered protected by the optical layer recovery mechanisms, i.e. 70% unprotected links, unless specifically stated otherwise.

4.3.1 Partial Disjoint Path Evaluation

4.3.1.1 QoS Restorable Routing Algorithms

This set of experiments assume that the working path is selected using the WSP algorithm. Thereby, these experiments evaluate the mechanisms to select the backup path to minimize the total reserved spare capacity over all network links in the IP/MPLS layer, taking into account the already protected links at the optical layer. Two new routing schemes based on the proposed PDP routing algorithm are evaluated:

- *Partial Disjoint Path with Full Information Routing* (PDPFIR). This algorithm uses the PDP routing algorithm to compute the backup path and considers full network information.
- *Partial Disjoint Path with Partial Information Routing* (PDPPIR). This algorithm computes the backup path by applying the PDP and considering partial network

information.

Table 10: Routing Schemes for PDP evaluation.

Routing Algorithm	Working Path	Backup Path	Backup path method
PDPFIR	WSP	PDP (Eq. 18)	Global, segment, local
PDPPIR	WSP	PDP (Eq. 19)	Global, segment, local
FIR	WSP	FIR	Global
PIR	WSP	PIR	Global

As listed in Table 10, routing algorithms presented in Section 3.4.2, FIR and PIR, are also considered in order to compare the merits of the new routing schemes. The algorithms are evaluated under incremental traffic assumption, first, considering 30% of protected links and, then, under different protection levels in terms of percentage of unprotected links.

4.3.1.2 Incremental Traffic and 30% Protected Network Scenario

In this set of experiments, 30% of the network links are protected by the optical layer recovery mechanisms, i.e. 70% unprotected links. The protected links have been randomly selected.

As shown in Fig. 42, the PDP with full network information routing algorithm, PDPFIR, outperforms all the other routing algorithms. Although PDPPIR is slightly better than FIR and PIR, its request rejection ratio increases and results in either similar or worst behavior depending on the network topology for medium/high number of requested LSPs. For low requested LSPs, both PDPFIR and PDPPIR routing schemes are significantly better than FIR and PIR. As shown in more detail in Fig. 43, PDPFIR and PDPPIR produce no rejected requests (0% request rejection ratio) while FIR and PIR have already rejected requests.

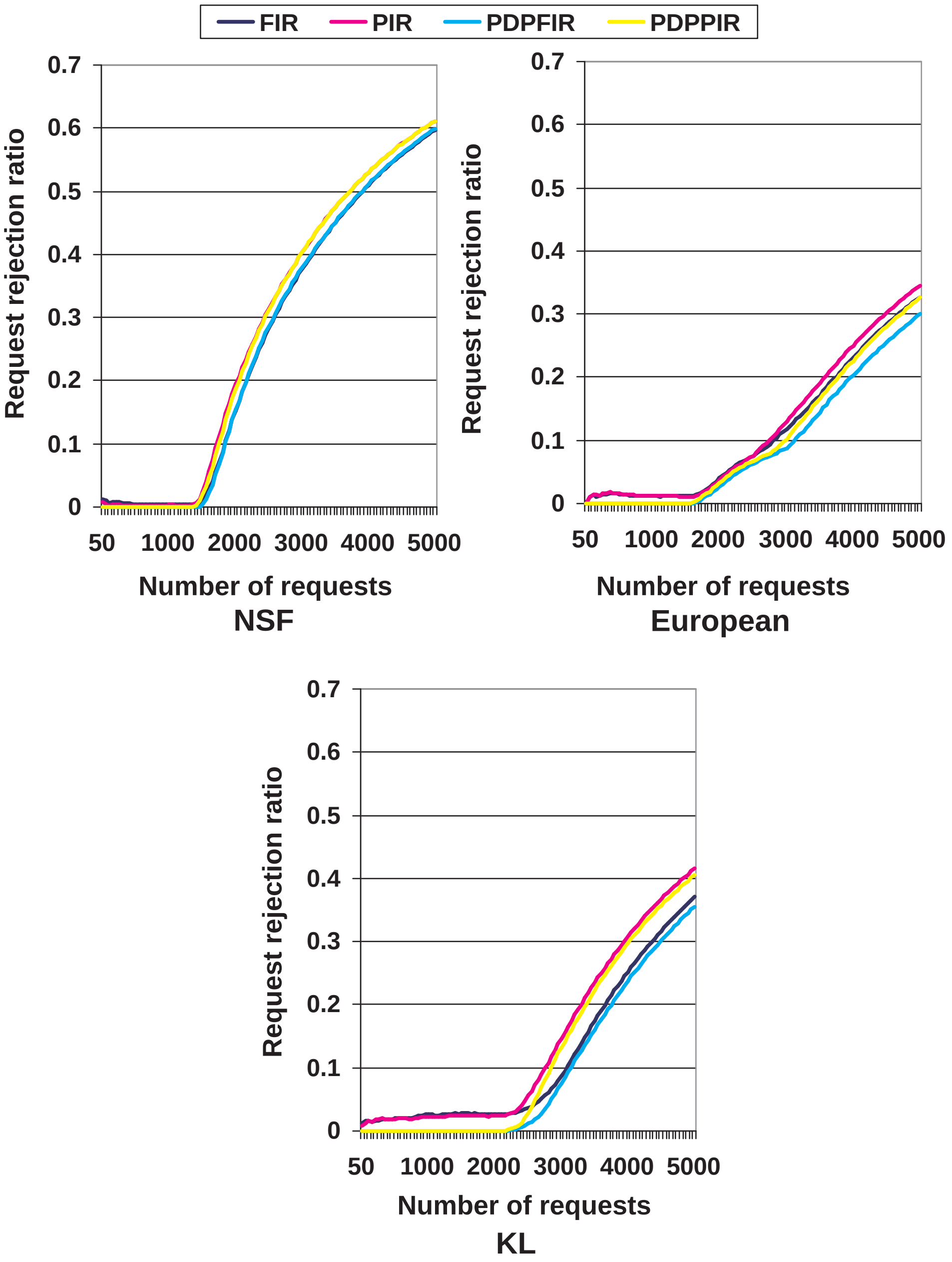


Figure 42: Request rejection ratio for each network topology.

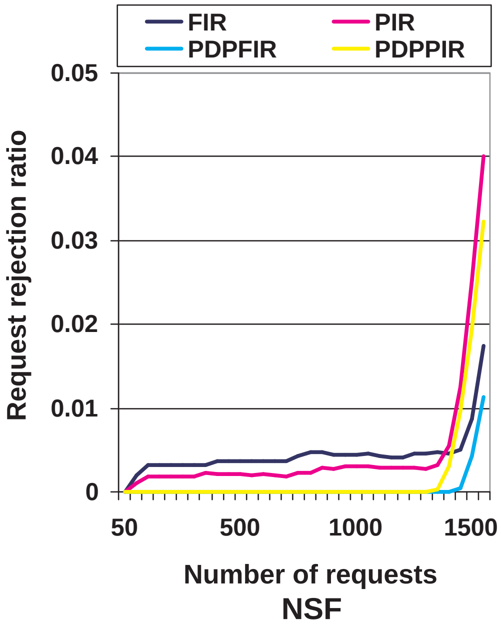


Figure 43: Magnified request rejection ratio for NSF network topology.

As shown in Fig. 43, PDPFIR routing algorithm accepts more requests than the other routing schemes. On the other hand, when PDPPIR routing algorithm starts rejecting requests, it gets dramatically worse. According to the network load, the best choice is always the PDPFIR algorithm. However, for low reuested LSPs, the PDPPIR algorithm could be a better choice because full routing information is not always available. Partial information routing algorithms also simplifies the management of the network.

In terms of restoration overbuild, PDPFIR and PDPPIR outperform FIR and PIR, respectively (see Fig. 44). This improvement is equally proportional for each routing algorithm in each network topology, around 4%/10%/5% in NSF, European and KL networks respectively. As expected from the request rejection ratio results, the routing algorithms achieving the poorest level of sharing are the ones that consider partial network information, i.e. PIR and PIRFP.

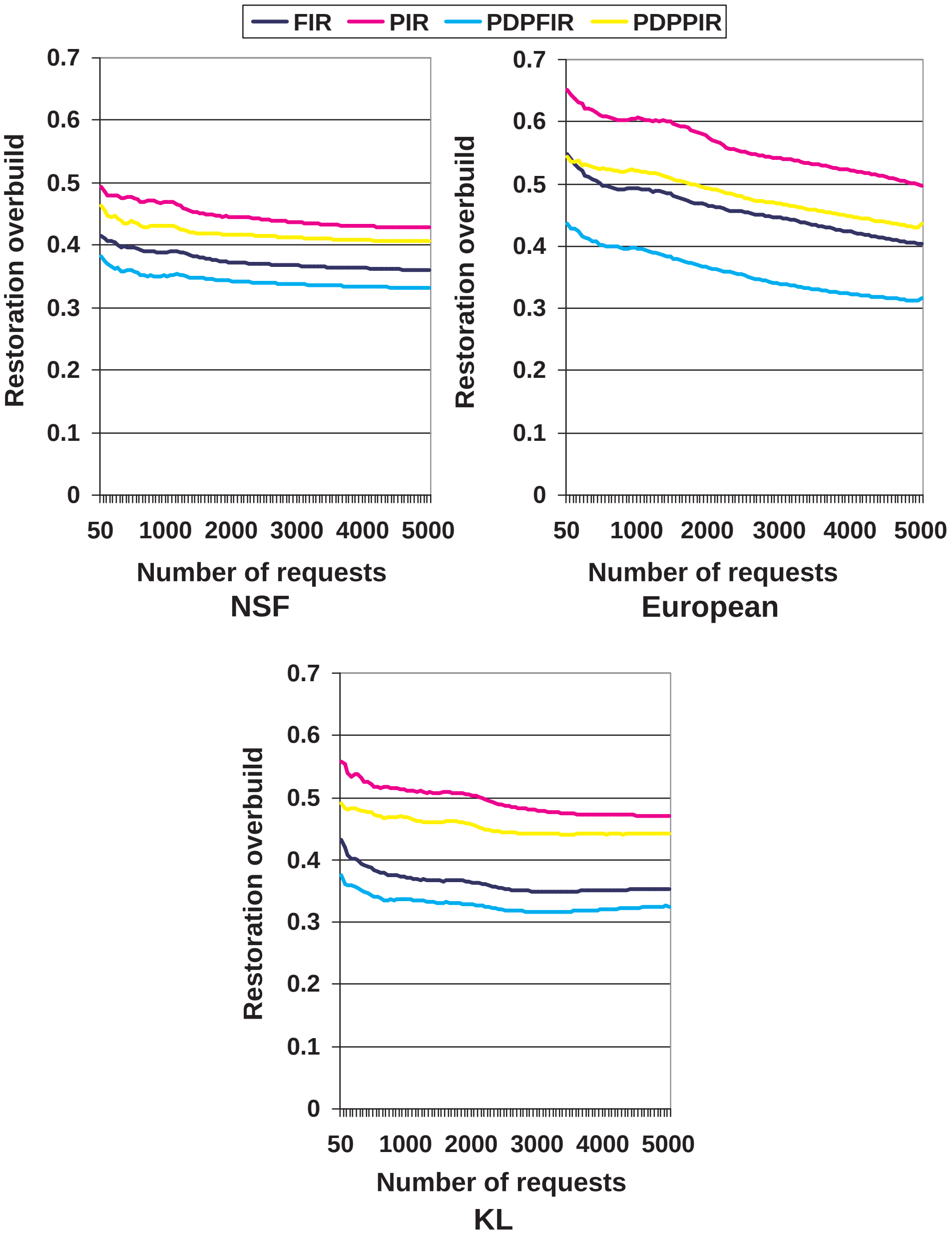


Figure 44: Restoration overbuild for each network topology.

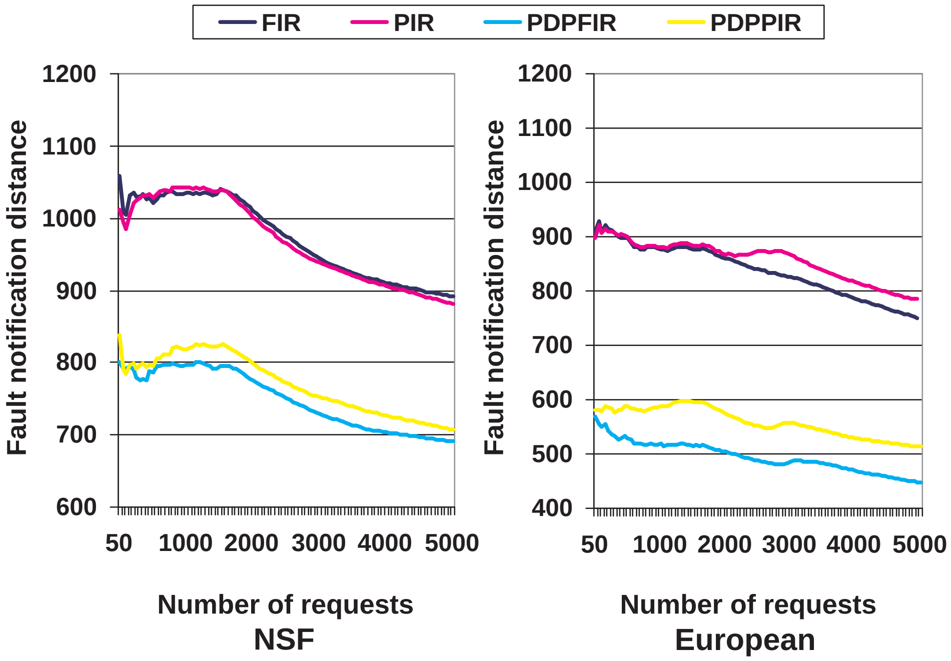


Figure 45: Fault notification distance for each network topology.

Figure 45 shows the improvement of the fault recovery time obtained when PDPFIR and PDPPIR algorithms are applied. Thereby, these algorithms not only result in better resource consumption, moreover, they offer faster recovery time than FIR and PIR algorithms. Note that PDPFIR is faster than PDPPIR.

It can be concluded from these results that the PDPFIR algorithm should be chosen due to faster recovery time. On the other hand, PDPPIR could also be chosen since its recovery time is not larger than the one in PDPFIR. Thus, slightly slower recovery time could be achieved in exchange for simplifying the management of the network, due to less network information flooding.

These results have been obtained conducting the simulations into a 30% optically protected network. Next section evaluates the performance of these algorithms in different protected network scenarios.

4.3.1.3 Incremental Traffic and Different Levels of Protection

Since all the algorithms offer similar performance independently on the network used, in this section the simulations are carried on the European network topology. These experiments proves that the PDP algorithms offer lower request rejection ratio when the number of protected links increases and the request rejection ratio increases when less number of links are protected. Thereby, the following levels of protection are considered: 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% and 10% unprotected links by the optical layer.

Results are presented in Table 11 for up to 1% of rejected requests, the maximum number of requests increases when the number of unprotected links decreases.

Table 11: Maximum number of accepted requests for up to 1% of rejected requests.

Routing Schemes	**Level of protection (% unprotected links)**								
	90%	80%	70%	60%	50%	40%	30%	20%	10%
PDPFIR	1804	1805	1833	1862	1896	1897	2927	2933	3339
PDPPIR	1779	1779	1779	1814	1857	1857	2916	2917	3360

From these results each protection level can be assorted on three groups since their performance is similar:

- 10% of unprotected links with 3339 requests.
- From 20% to 30% of unprotected links with around 2930 requests.
- From 50% to 90% of unprotected links with up to 1897 requests.

This information suggests that there is no need to assign resources at the lower layer to protect some links according to the requested load in the network. For instance, in networks with 2930, the network should be 40% unprotected instead of 30% because the penalty of unprotecting 10% less links is very low. The same explanation can be applied for 50%-90% of unprotected links.

The maximum fault recovery time, evaluated in terms of fault notification distance, is presented in Table 12. As expected the fault notification distance decreases when the number of unprotected links decrease. This is because less links require to be protected at the IP/MPLS layer. Thereby, the working paths have few links to protect and segment and local backups are only set up.

Table 12: Maximum fault notification distance (km).

Routing Schemes	**Level of protection (% unprotected links)**								
	90%	80%	70%	60%	50%	40%	30%	20%	10%
PDPFIR	731	635	585	536	506	399	370	370	370
PDPPIR	758	715	603	559	515	419	370	370	370

4.3.2 Reliable Services with Fast Protection Evaluation

4.3.2.1 QoS Restorable Routing Algorithms

Three new routing schemes based on the two routing algorithms (KMI and RPDP) and the class of service are proposed (see Table 13):

- *Reliable Services with Fast Protection* (RSFP). This algorithm uses KMI to compute the WP and RPDP to compute the BP. This algorithm has been explained previously (see Section 4.2).

- *Semi-Reliable Services* (SRS). This algorithm uses WSP to compute the WP and a variation of RPDP to compute the BP. In SRS, the cost c_{ij} given in Equation 14 is assigned. For this case, dedicated capacity allocation cannot be applied using this method. Therefore, the requirements of HR requests are not achieved.

- *Semi-Reliable Services with Minimum Interference* (SRSMI). This algorithm uses KMI to compute the WP and the variation of the RPDP used on SRS.

Table 13: Routing Schemes for RPDP evaluation.

Routing schemes	**Path**		**Traffic Services**		
	WP	**BP**	**LR**	**MR**	**HR**
RSFP	KMI	RPDP Eq. 20	Partial protection, shared backups, medium, slow recovery time	Full protection, shared backups, fast recovery time	Full protection, dedicated backups, very fast recovery time ($\simeq 0$)
SRS	WSP	RPDP Eq. 14			Full protection, shared backups, very fast recovery time
SRSMI	KMI	RPDP Eq. 14			
NRS	WSP	FIR	Full protection, shared backups, medium, slow recovery time		
NRSMI	LMIR	FIR			

In order to compare the merits of these schemes, two algorithms not oriented to multi-layer and multi-service differentiation are also considered:

- *No-Reliable Services* (NRS). This algorithm has the objective of minimizing resource consumption used by backup paths. Therefore, FIR is used to compute the backup path, whereas the working path is computed using WSP.

- *No-Reliable Services with Minimum Interference* (NRSMI). This scheme takes into account the minimization of interference. LMIR is used to compute the WP and

FIR to compute the backup path.

4.3.2.2 Simulation Results under Incremental Traffic

As shown in Fig. 46, the algorithms that have in their objectives the static multi-layer information, RSFP, SRS and SRSMI, outperform all other routing algorithms. These algorithms are avoiding to protect the links protected by the optical layer, consequently, less amount of resources are reserved at the IP/MPLS layer for protection. Although RSFP is much better than SRS and SRSMI, its request rejection ratio increases and results in worst behavior for high requested LSPs because RSFP reserves dedicated capacity to protect HR traffic. For medium and low requested LSPs RSFP accepts more LSPs. For up to 5% of rejected requests, RSFP sets up around 2000 LSPs, SRS and SRSMI set up around 1800 requests and NRS and NRSMI around 1500 LSPs.

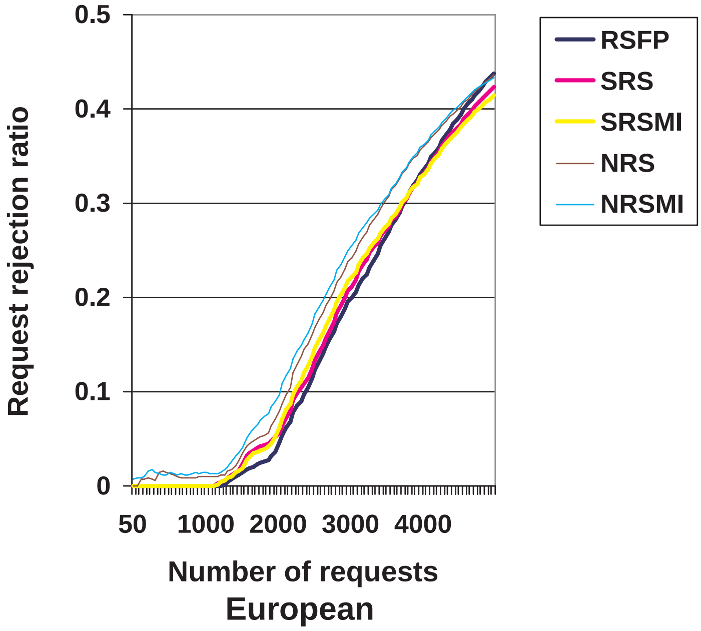


Figure 46: Request rejection ratio.

Table 14 shows these results for up to 1% of rejected requests. RSFP should be the algorithm used to compute the working and backup path because it accepts major number of requests, 1450.

Table 14: Maximum number of accepted requests for up to 1% of rejected requests.

Routing schemes				
RSFP	SRS	SRSMI	NRS	NRSMI
1450	1366	1380	300	169

In terms of restoration overbuild, the RSFP algorithm shares less spare capacity compare to the others algorithms (see Fig. 47). Note that RSFP algorithm is the only algorithm that dedicates spare capacity to protect the HR requests. For that reason, the ratio of spare capacity used to protect working paths is higher than SRS and SRSMI.

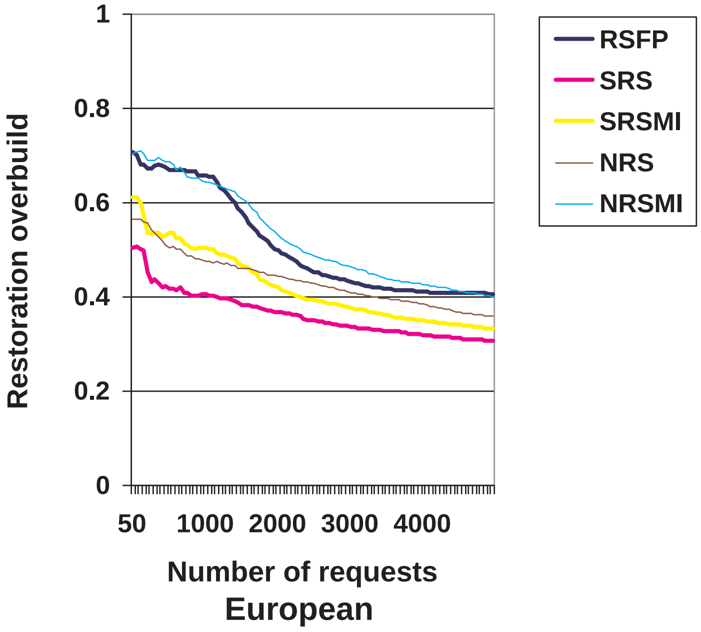


Figure 47: Restoration overbuild.

Figure 48 shows the improvement of the fault recovery time obtained when the proposed RSFP, SRS and SRSMI algorithms are applied. Thereby, these algorithms not only result into better resource consumption, moreover, they offer faster recovery time than NRS and NRSMI algorithms. Note that our main algorithm, RSPF, offers the fastest recovery time.

It can be concluded from these simulations that, for offering up to low rejected requests, the RSFP algorithm should be chosen for faster recovery time and better use of the network resources. These results have been obtained conducting the simulations under incremental traffic. Next section evaluates the performance of these algorithms under limited holding time.

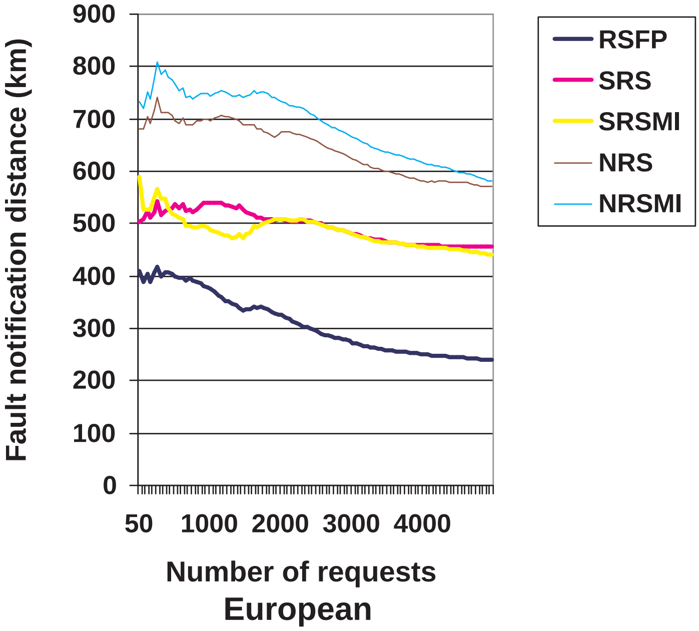


Figure 48: Fault notification distance.

4.3.2.3 Simulation Results under Limited Holding Time

In this simulation, two network scenarios are evaluated. in the first network scenario, the capacity of the links is set to 4800 units, representing OC-48 rates and, in the second, the capacity is set to 1200 units representing OC-12 rates. Each link is bi-directional i.e. they act like two unidirectional links of the same capacity. Requests arrive according to a Poisson process with an average rate λ, and exponentially distributed holding times with a mean value of $\frac{1}{\mu}$. In this set of experiments, $\frac{\lambda}{\mu}$ is 150.

Table 15 shows that the QoSP routing algorithms offer 0 request rejection ratio throughout the experiment for OC-48 link rates. In this case, the best algorithm is RSFP; it takes into account all the QoSP requirements of traffic class that neither SRS nor SRSMI consider. Thereby, although SRS and SRSMI have a suitable performance in terms of network resources, accurate reliability required for HR traffic class is not achieved.

Table 15: Total number of rejected requests through the simulations.

Link	**Routing schemes**				
rates	RSFP	SRS	SRSMI	NRS	NRSMI
OC-48	0	0	0	171	217
OC-12	2837.0	613.0	634.0	926.0	1016.0

On the other hand, when link rate is low OC-12, the network is high loaded, and RSFP offers higher number of rejected requests. On the other hand, SRS and SRSMI algorithms keep performing better than NRS and NRSMI. It can be concluded that the proposed SRS and SRSMI algorithms are independent of the network characteristics and traffic loads. This is not the case of the RSFP due to the dedicated spare capacity of the HR traffic class. For that reason, its performance gets worse when the network load is high.

4.4 Concluding Remarks

In this chapter, as a first stage of this thesis that aims to improve the QoSP routing algorithms for IP/MPLS-based networks, new routing schemes have been proposed. The proposed routing schemes apply segment protection for fast recovery and shared backups to reduce the amount of resources assigned for protection. Although the logical topology where the LSPs are routed is given, the lower layer information, where some logical links are already protected, is considered. Taking into account this information, the routing algorithms avoid to protect at IP/MPLS level those links that are optically protected. In order to carry out this process PDP computation is implemented. Moreover, a new definition of link-disjoint path based on Shared Risk Link Group (SRLG) has been introduced in order to share more spare capacity and, consequently, minimizing the resource consumption. The second proposed scheme defines different levels of reliability and failure impact in terms of recovery time depending on the QoS traffic class requirements. Results have shown that the proposed RSFP algorithm offers the requirements of all the defined traffic classes without adding more resource consumption or increasing the rejected requests of previous proposals that did not consider traffic differentiation. Moreover, our proposed algorithms SRS and SRSMI, improve upon the previous ones in terms of resource consumption or rejected requests.

This chapter has left some questions that the following chapters cover. Our proposed schemes result in efficient resource consumption and recovery time under the assumption of a partial protected optical layer. However, no study related to the impact of protecting at the IP/MPLS or optical layers is done. Next chapters provide 1) an explanation of why the recovery mechanisms cannot be only applied at optical layer, 2) an analysis of the advantages and disadvantages of the recovery mechanisms at each network layer and, finally, 3) some proposals for dynamic multi-layer routing that consider a dynamic cooperation between packet (IP/MPLS) and wavelength (optical) switching domain.

CHAPTER V

MULTI-LAYER SURVIVABILITY: WHERE TO RECOVER?

As presented in Chapter 3, different QoSP routing algorithms are found in the literature in order to offer the reliability required by traffic services. However, they are oriented to a single switching layer: either wavelength (lambda) or packet oriented. In Chapter 4, new QoSP routing approaches that consider multi-layer network recovery have been proposed, although the lower optical layer has been considered pre-established and static. Basically, these proposals are taking advantage of knowing which lightpaths are already protected at the optical layer and which lightpaths are unprotected. Thus, in the IP/MPLS layer, spare capacity is only reserved to protect the *unprotected* lightpaths. However, these approaches only focus on packet switching.

Cooperation between each switching layer should be considered in order to provide protected paths cost effectively. The interoperability between each switching layer, IP/MPLS and optical domains, is enabled by Generalized Multi-Protocol Label Switching (GMPLS). Although there have been efforts in developing routing algorithms that consider both switching layers, multi-layer protection is not considered. This chapter aims to provide the characteristics of each switching layer and their advantages and disadvantages for fault recovery as well as a brief literature survey of the related research in multi-layer routing.

5.1 Multi-layer Survivability Overview

Different recovery strategies are considered in the multi-layer network scenario [1, 29]: *bottom layer*, *bottom-up* and *top layer*.

In the *bottom layer* recovery strategy, the optical layer is responsible for recovering from all failures. In the *top layer* recovery strategy, the IP/MPLS layer is the responsible of the recovery. The most common recovery strategy used is the *bottom-up*. In this case, the bottom layer, i.e. the optical layer, first tries to recover the failed connections. After a hold-off time [29], if the bottom layer has not been successful then the upper layer, i.e. the IP/MPLS layer, tries to recover the failed connections.

The pros and cons of these recovery strategies are depicted in Tab. 16 [1]. Table 16 also shows the preferred values of the performance parameters considered.

Table 16: Comparison of some recovery strategies [1].

Performance Criteria	**Survivability Strategy**			**Preferred Value**
	Bottom Layer	*Bottom-Up*	*Top Layer*	
Switching granularity	Coarse	Coarse	Fine	Coarse
Failure coverage	Low	High	High	High
Required Capacity Resources	Low	High	Low	Low
Service Differentiation	Difficult	Difficult	Easy	Easy
Coordination, Management	Low	High	Low	Low
Failure Scenario	Simple	Simple	Complex	Simple
Recovery close to root	Yes	Yes	No	Yes
Strategy complexity	Low	Medium	Low	Low

5.1.1 Photonic MPLS Router: Packet Switching Capabilities

The new photonic MPLS routers offer packet and wavelength switching [77]. Thus, packet Label Switched Paths (p-LSPs) are routed in the optical network through wavelength paths, called indistinctly in this work lightpath and lambda LSPs (λ-LSPs). The architecture of a photonic MPLS router is shown in Fig. 49 (refer to [78, 79] for more details).

For better utilisation of the network resources, p-LSPs should be efficiently multiplexed into lightpaths and then, these lightpaths should be demultiplexed into p-LSPs at some router. This procedure of multiplexing/demultiplexing and switching p-LSPs onto/from λ-LSPs is called traffic grooming. Traffic grooming is an important issue for next generation optical networks. An in-depth study about the characteristics of different optical grooming switches has been presented in [80]. The photonic MPLS routers have the technology to implement traffic grooming. It consists of a number of Packet-Switching Capable (PSC) ports, p, and number of wavelengths, w [79] (see Fig. 49). The number of PSC indicates how many lambda LSPs can be demultiplexed into this router, whereas the number of wavelengths corresponds to the number of wavelengths connected to the same adjacent router.

Based on these parameters, a new resource constraint is added to the network: p, the number of PSC ports. Three scenarios are associated with according to the following

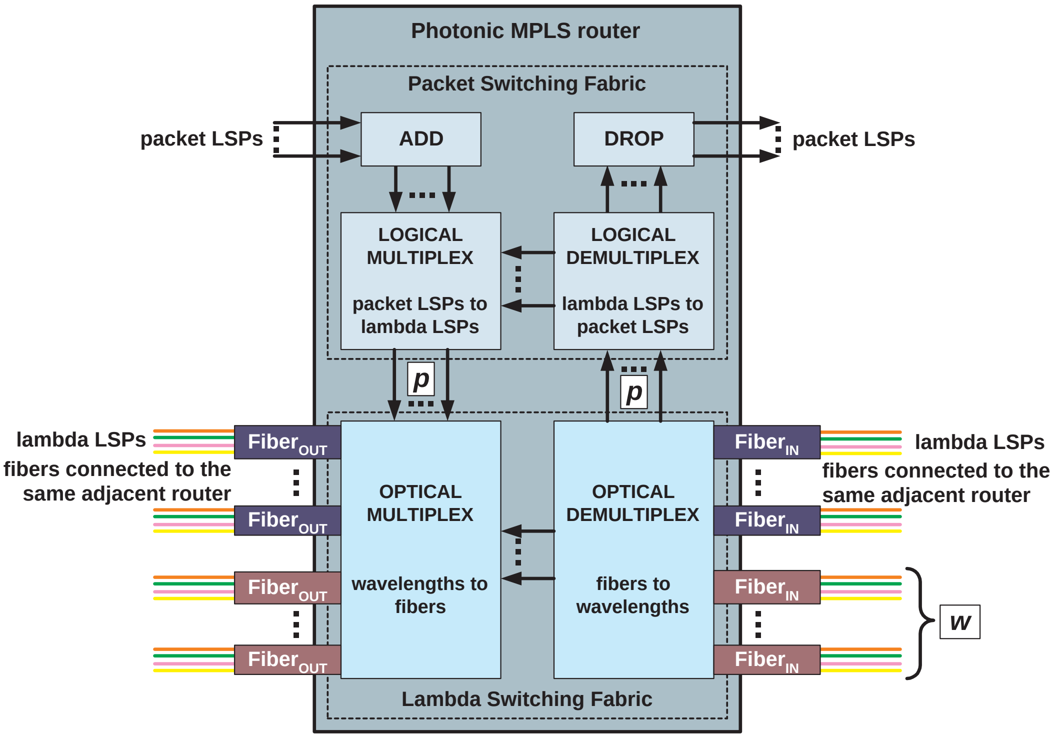


Figure 49: Photonic MPLS router architecture.

switch architectures [80]:

- *Single-hop grooming*: $p = 0$. Using this type of switching architecture, the network does not offer packet switching capability at intermediate nodes. Thus, low speed traffic from source node is multiplexed onto a wavelength switched to the same destination node. Moreover, the protection should be performed either at the optical domain and is λ-LSP oriented or at the IP/MPLS domain and is p-LSP oriented using only path protection (global) strategies.

- *Multihop partial grooming*: $0 < p < w$. In this case, some of the wavelengths may be demultiplexed at the intermediate nodes for switching at finer granularity. Therefore, not all the p-LSP are able to perform segment/local protection.

- *Multihop full grooming*: $p = w$. Every wavelength on each fibre link forms a lightpath between adjacent node pairs. Thus, the logical topology is predetermined and exactly the same as the physical topology. All the protection strategies, i.e. global, segment and local, are suitable for all p-LSP.

Note that although the PSC ports at intermediate nodes allow performing packet segment/local protection, the number of optical-electrical-optical (o-e-o) conversions increases. Thus, the cost of o-e-o conversions must be also considered because they represent a bottleneck to network throughput and also influence the overall delay and buffer usage [25].

5.1.2 Switching granularity

The granularity of the recovery strategy is an important parameter in terms of recovery time and fault management. Diverse switching granularity levels exist into the optical IP/MPLS network scenario. Going from coarser to finer, there is fibre, wavelength (lightpath) and packet (LSP) switching as shown in Fig. 50.

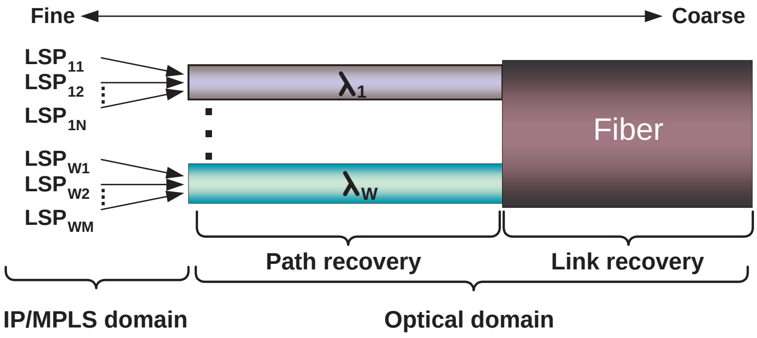


Figure 50: Switching granularity

In the optical layer case, when a failure occurs either all lightpaths that travel along the failed link are simultaneously rerouted (*link recovery*) or each affected lightpath is individually switched on its alternative path (*path recovery*). Therefore, the level of recovery at the optical layer is bundle of lambdas or individual lambdas. This is not the case in the IP/MPLS layer, where the LSPs are multiplexed onto lightpaths. Recovery at the IP/MPLS layer is at the packet level. Since recovering at the optical layer recovers affected connections in-group, the recovery action is also fast and easier to manage than recovering each affected connection individually (p-LSP) in the IP/MPLS layer [29]. On the other hand, the recovery using *bottom-up* strategy first tries to recover the failure with the coarsest granularity (optical layer). If the failure is not solved, the upper layer with finer granularity, i.e. IP/MPLS layer, is used. As shown in Table 16, *bottom-up* recovery strategy results in high resource consumption since both layers are dedicating resources for recovering the failure. The granularity of the nodes has an important role in order to efficiently use the network resources. Thus, if the intermediate nodes have only wavelength granularity, without packet switching capability, all traffic multiplexed to the same lightpath at the source node will be switched to the same destination node.

5.1.3 Required Capacity Resources

Internet Service Providers (ISPs) aim to achieve the required level of protection with minimum resource consumption. Different techniques can be used at different network layers in order to reduce the capacity usage. However, the switching granularity of each layer also affects to the level of resources used to protect the connections. The finer is the granularity; lower is the resource consumption. This section analyzes the amount of network resources used for the working path and the backup path according to the switching granularity.

5.1.3.1 Terminology

For the network resources analysis, the next terminology is used to represent λ-paths:

$$(n_1 \xrightarrow{\lambda_1} n_2 \xrightarrow{\lambda_2} n_3 \xrightarrow{\lambda_3} n_4)$$

Where (n_1, n_2, n_3, n_4) is the ordered sequence of nodes across the λ-path and $\lambda_1, \lambda_2, \lambda_3$ are the wavelengths used in the corresponding links. For instance, link (n_1, n_2) uses wavelength λ_1, link (n_2, n_3) uses λ_2 and, finally, link (n_3, n_4) uses λ_3. For better understanding, the fibre is not represented in this terminology; one fibre is only used in each link. The representation can be easily extended adding the fibre id as follows: $(n_1 \xrightarrow{f_1,\lambda_1} n_2 \xrightarrow{f_2,\lambda_2} n_3 \xrightarrow{f_3,\lambda_3} n_4)$.

5.1.3.2 Amount of Resources for the Working Path

An example comparing the optical and IP/MPLS switching domain for establishing packet LSPs, working paths, is shown in Figs. 51 and 52. In these figures, each physical link has one fibre with two wavelengths α, β with 10 units of capacity. Lets suppose three requests: WP_1, WP_2 and WP_3. Each request has a capacity requirement of 3 units. First consider the set-up when intermediate nodes have only lightpath granularity as shown in Fig. 51 and Tab. 17. Hence, WP_1 is established and the lightpath L_1 $(1 \xrightarrow{\alpha} 4 \xrightarrow{\alpha} 5 \xrightarrow{\alpha} 2)$ is assigned to the working path WP_1. Afterwards, WP_2 is established and lightpath L_2 is assigned to it:$(3 \xrightarrow{\beta} 4 \xrightarrow{\beta} 5 \xrightarrow{\beta} 6)$. Finally, when request WP_3 arrives, it is rejected since there is not sufficient network resources, i.e. a new lightpath cannot be established, since fibre 4-5 does not have more free wavelengths and traffic cannot be multiplexed in node 4 in any used wavelength, though the network has sufficient capacity.

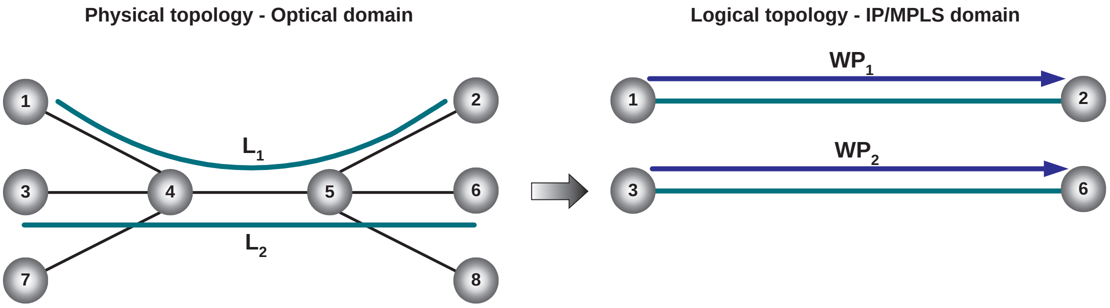

Physical topology - Optical domain
Logical topology - IP/MPLS domain
1
2
3
4
5
6
7
8
L1
L2
WP1
WP2

Figure 51: Optical resource consumption for working path set-up.

Table 17: Optical resource consumption for working path set-up.

Resource	WP_1	WP_2	WP_3	**Total Resources**
Lightpath	$(1 \xrightarrow{\alpha} 4 \xrightarrow{\alpha} 5 \xrightarrow{\alpha} 2)$	$(3 \xrightarrow{\beta} 4 \xrightarrow{\beta} 5 \xrightarrow{\beta} 6)$	Rejected	6 wavelengths
Capacity	3+3+3	3+3+3	-	18 units

On the other hand, if requests are managed at IP/MPLS layer WP_3 may be accepted since traffic grooming is possible at intermediate nodes (see Fig. 52 and Tab. 18). For instance, the WP_2 may be established by using three lightpaths, L_{21} $(3 \xrightarrow{\beta} 4)$, L_{22} $(4 \xrightarrow{\beta} 5)$ and L_{23} $(5 \xrightarrow{\beta} 6)$; and the WP_3 may be established by using two new lightpaths, L_{31} $(7 \xrightarrow{\beta} 4)$ and L_{33} $(5 \xrightarrow{\beta} 8)$ and the existing lightpath L_{22}. Thereby, at node 4 the lightpaths L_{21} and L_{31} are demultiplexed. Then, the WP_1 and WP_3 from the respective lightpaths are multiplexed to L_{22}. At node 5, this lightpath is demultiplexed, and the WP_2 is multiplexed to L_{23} and the WP_3 is multiplexed to L_{33}. Thus, in this case no request is rejected.

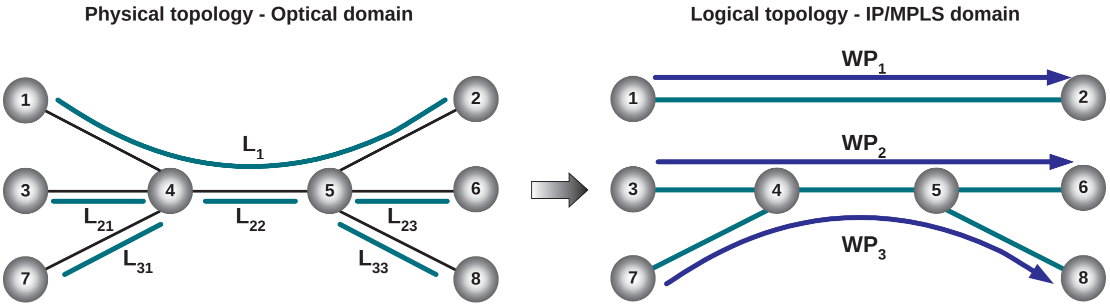

Physical topology - Optical domain
Logical topology - IP/MPLS domain
1
2
3
4
5
6
7
8
L1
L21
L22
L23
L31
L33
WP1
WP2
WP3

Figure 52: IP/MPLS resource consumption for working path set-up.

Table 18: IP/MPLS resource consumption for working path set-up.

Resource	WP_1	WP_2	WP_3	Total Resources
Lightpath	$(1 \xrightarrow{\alpha} 4 \xrightarrow{\alpha} 5 \xrightarrow{\alpha} 2)$	$(3 \xrightarrow{\beta} 4)$ $(4 \xrightarrow{\beta} 5)$ $(5 \xrightarrow{\beta} 6)$	$(7 \xrightarrow{\beta} 4)$ $(4 \xrightarrow{\beta} 5)$ $(5 \xrightarrow{\beta} 8)$	8 wavelengths
Capacity	3+3+3	3+3+3	3+3+3	27 units

Note that intermediate nodes should be provided of Packet Switching Capable ports. Therefore, although at IP/MPLS domain no request is rejected, there is the equipment cost added at intermediate nodes as explained in Section 5.1.1.

5.1.3.3 Amount of Resources for the Backup Path

An example comparing the optical and IP/MPLS switching domain for establishing protected packet LSPs, working and backup paths, is shown in Figs. 53 and 54. In these examples, each physical link represents one fibre. Lets suppose that two protected packet LSPs, WP_1 and WP_2, must be established with a requested capacity of 5 and 6 units respectively and all network links have 2 wavelengths (α, β). First the WP_1 is computed and set-up between node pair (1, 2) with the respective backup path, BP_1. Afterwards, the WP_2 is also set-up between node pair (5, 6) with the respective backup path, BP_2. Lets, first, consider the set-up when intermediate nodes have only wavelength granularity shown in Fig. 53 and Tab. 19. In this case, two λ-paths (lightpaths) are computed with the respective backup λ-path. Since both λ-paths have different source and destination nodes the traffic of both working paths cannot be multiplexed in the same wavelength (λ-path) resulting into a total spare capacity of 33 units of capacity.

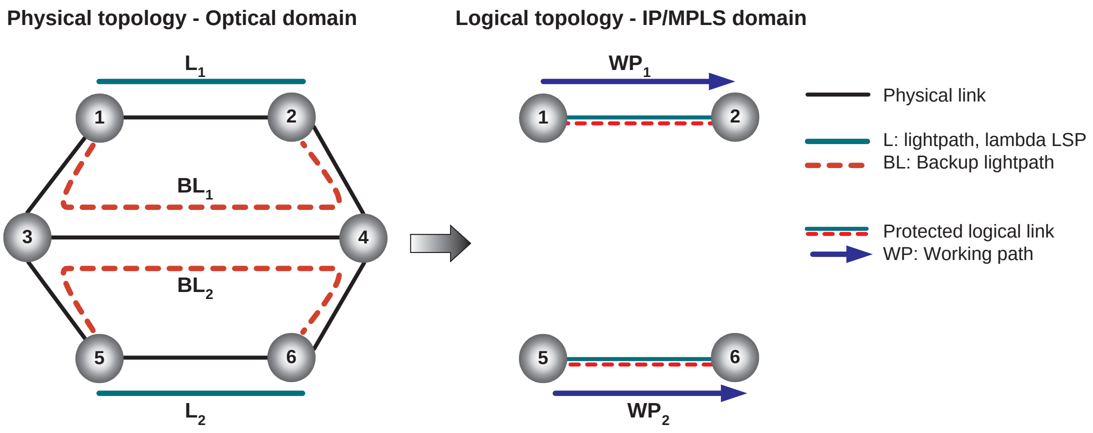


Figure 53: Optical resource requirements for backup paths.

Table 19: Optical resource consumption for backup path set-up.

Resource	Backup of WP_1	Backup of WP_2	Total Resources
Lightpath	$(1 \xrightarrow{\alpha} 3 \xrightarrow{\alpha} 4 \xrightarrow{\alpha} 2)$	$(5 \xrightarrow{\beta} 3 \xrightarrow{\beta} 4 \xrightarrow{\beta} 6)$	6 wavelengths
Capacity	5+5+5	6+6+6	33 units

On the other hand, if the requests are treated at IP/MPLS layer (intermediate nodes with packet switching capabilities), the backup link 3-4, assigned to BP_1, may be shared with the backup path BP_2. Thus, when BP_2 is established only one unit more of extra-spare capacity is reserved in fibre 3-4 wavelength α, resulting on a total of 28 units of spare capacity on the network, 5 units less than the optical case.

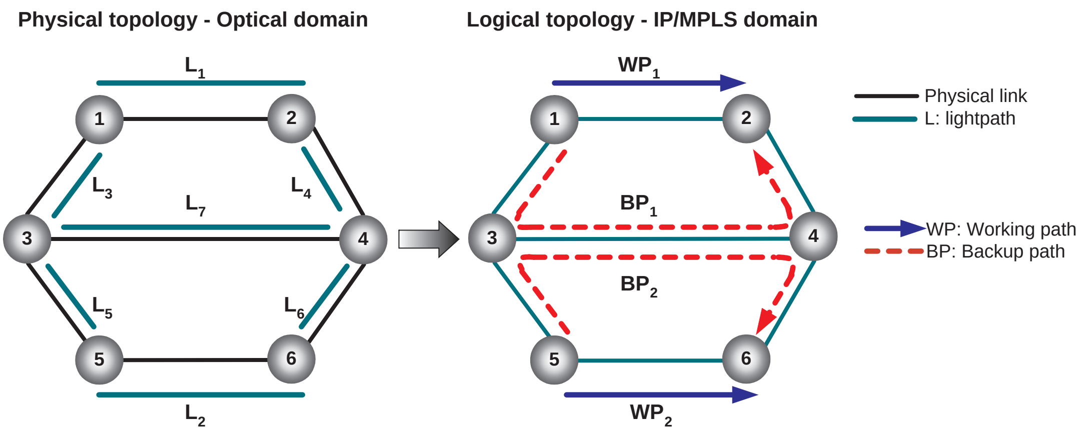


Figure 54: IP/MPLS resource requirements for protected packet LSPs.

Table 20: IP/MPLS resource consumption for backup path set-up.

Resource	Backup of WP_1	Backup of WP_2	Total Resources
Lightpath	$(1 \xrightarrow{\alpha} 3)$	$(5 \xrightarrow{\alpha} 3)$	5 wavelengths
	$(3 \xrightarrow{\alpha} 4)$	$(3 \xrightarrow{\alpha} 4)$	
	$(4 \xrightarrow{\alpha} 2)$	$(4 \xrightarrow{\alpha} 6)$	
Capacity	5+5+5	6+1+6	28 units

Therefore, the IP/MPLS finer granularity results into a better resource consumption. However, finer granularity means at the same time more failure management, i.e. the number of required recovery actions is higher as the following section points out.

5.1.4 Signaling overhead

After a failure occurs, the node that detects the failure needs to notify the failure to the nodes responsible of the switchover. The level of recovery at the optical domain is bundle of λ-LSPs or individual λ-LSPs. Hence, if a single fibre failure occurs, at the most w notification messages should be sent. This is not the case of the IP/MPLS, where p-LSPs are multiplexed onto λ-LSPs. The level of recovery at the IP/MPLS domain is p-LSP. Thus, according to the granularity coefficient around $g \cdot w$ fault notification messages should be sent, where g is the maximum number of p-LSPs that a lightpath can transport.

Note that the number of failed p-LSPs is much larger than the number of failed λ-LSPs. Therefore, since fault recovery at the optical domain recovers the affected connections in-group, the fault recovery action is also faster and easier to manage; the number of required recovery actions is minimal. That is not the case of recovering each affected connection individually (p-LSP) in the IP/MPLS domain [25].

5.1.5 Wavelength Conversion Capability

The ability to convert the wavelength of an incoming signal to another wavelength before leaving the outgoing port is possible if the optical node has Wavelength Conversion Capability (WCC) [29]. However, although the routing is more flexible when the WCC is available, wavelength converters are expensive. On the other hand, if the nodes do not have WCC, then the λ-LSP uses the same wavelength on all the links along the path. The constraint of using the same wavelength along the path is called wavelength continuity constraint and introduces the well-known problem of Routing and Wavelength Assignment (RWA) [26]. This work assumes that all the routers are provided of wavelength converters, thus, the network has full wavelength conversion capability.

5.1.6 Activation of the Recovery Schemes

When an optical failure occurs, such a fibre cut, both IP/MPLS and optical layer detect the failure. In order to avoid the activation of the fault recovery mechanisms of both layers at the same time, two strategies exist [1]: hold-off timer and recovery token signal. In the case of using a *hold-off time*, first the optical layer tries to recover the failure. After the expiration of the holt-off timer, if the optical recovery has succeeded, no recovery actions are needed in IP/MPLS; otherwise, recovery actions in IP/MPLS should be initiated. The hold-off time must be long enough to guarantee that the optical recovery actions have finished.

In the case of using a *recovery token signal*, the optical layer explicitly sends the recovery token to the IP/MPLS layer when it cannot recover all or part of the traffic. When IP/MPLS layer receives the token, the recovery actions are initiated. The advantage of the recovery token signal is that it reduces the delay added to recover the failure at IP/MPLS layer when the optical layer is unable to recover it.

5.1.7 Service Differentiation

Not all the current and future traffic services have the same protection requirements. Some current works differentiate three traffic classes with their protection requirements [14, 76]. Obviously, the optical layer cannot perform different levels of protection according to the traffic classes, since it recovers at level of lambda (see section 5.1.2). This is not the case of the IP/MPLS layer, where protection is at p-LSP level. This allows to protect each p-LSP according to its protection requirements.

An example is shown in Fig. 55 and Fig. 56, where two requests, WP_1 and WP_2 with a requested capacity of 5 and 6 units respectively, are established using the same lightpath. Lets suppose that WP_2 does not have protection requirements. Since the optical layer protects at lightpath level, the backup lightpath recovers all the traffic that the lightpath is carrying on: WP_1 and WP_2, though WP_2 does not have protection requirements.

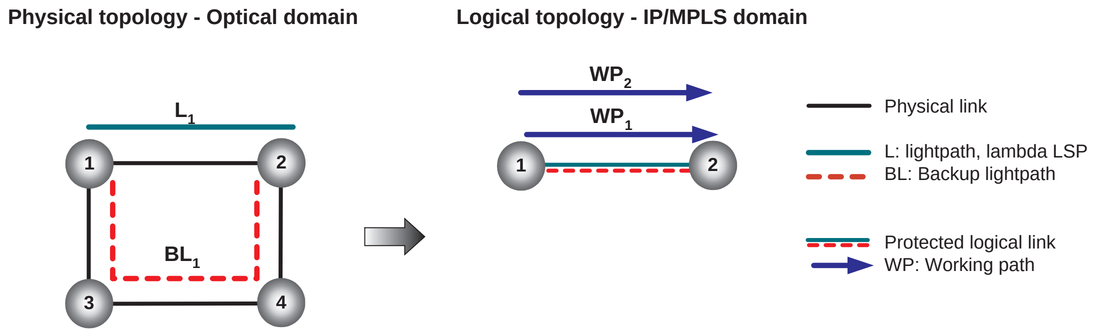


Figure 55: Optical resource consumption when traffic has different protection requirements.

However if the failure is recovered at the IP/MPLS layer, only capacity to recover WP_1 is reserved on the backup path BP_1 (5 units) reducing the resource consumption.

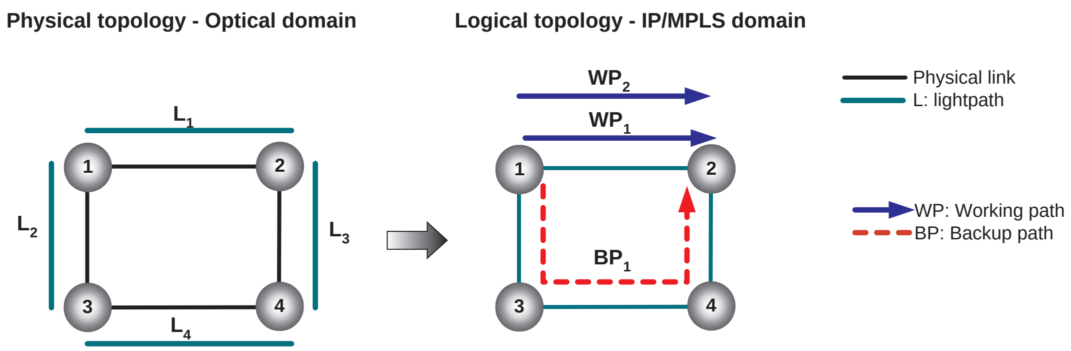


Figure 56: IP/MPLS resource consumption when traffic has different protection requirements.

5.2 *Multi-Layer Routing Schemes*

As shown in section 5.1, recovering at the IP/MPLS layer results into better resource consumption. However, the use of photonic MPLS routers as intermediate nodes adds complexity and management cost. Thus, a trade-off exists between a better filling of the capacity of the lightpaths (granularity) and the potential larger amount of higher layer equipment required [29]. This section provides a brief survey of the literature related to multi-layer routing schemes where some of the parameters presented above are analyzed.

5.2.1 Network Definition

Let $G_P = (V, E_P)$ and $G_L = (V, E_L)$ represent the physical topology and the logical topology respectively, where V is the set of photonic MPLS routers; E_P and E_L are the set of network physical links and λ-LSPs respectively. Each router has p input and output Packet Switching Capable (PSC) ports, where $PSCi(u)$ input ports and $PCSo(u)$ output ports of node u are already not assigned to any λ-LSP. Each physical link has w wavelengths.

The p-LSP request is defined by (s, d, b) where (s, d) is the source and destination node pair; and b, specifies the amount of capacity required for this request.

5.2.2 Grooming Using Auxiliary Graph

Traffic grooming is an important issue for next generation optical networks. The Grooming Using Auxiliary Graph (GUAG) algorithm was proposed by Zhu in [80] and it considers the different switching layers (see Alg. 7).

This algorithm first finds the auxiliary network graph, G', removing all the fibres and

Algorithm 7 Grooming Using Auxiliary Graph

```
INPUT
$s$: source node;
$d$: destination node;
$b$: capacity requested;
$G = (V, E_P \cup E_L)$: network graph;
ALGORITHM
$G' = G - \{$fibres without free wavelengths, lightpaths without enough residual capacity$\}$
$G' = Assign\_weights(G')$
if (exists a direct lightpath from $s$ to $d$) then
  $WP$ = direct lightpath
else
  $paths$= set of the least-cost paths
  if (no full wavelength capacity) then
    $WP$ = select the path from $paths$ with minimal number of free wavelength links.
  else
    $WP$ = select the path from $paths$ which traverses less number of PSC ports.
  end if
end if
if ($WP ==$ NULL) then
  return reject
end if
$postprocessing(WP, G)$
```

lightpaths that are not suitable. Afterwards, a weight is assigned to the link candidates based on the function $Assign\_weights(G')$. Finally, the WP is computed according to the capacity granularity requested and the residual network resources. Hence, if a lightpath, which directly connects the source and destination node pair, exists then this lightpath is selected. Otherwise, from the set of least cost paths, *paths*, 1) the one with minimal number of free wavelengths is selected if the request does not require full wavelength capacity 2) the one that traverses the minimal number of PSC ports is selected, if the request requires full wavelength capacity. Once the working path is computed the *postprocessing* function is executed. This function includes operations such as updating the number of available wavelengths on fibres, updating the number of $PSCi$ and $PSCo$.

In the simulations presented in [80], the authors showed that multihop full grooming, $p = w$, offers the best performance in terms of network blocking probability, wavelength utilisation and resource efficiency. However, it may have scalability problems since a large amount of o-e-o conversions are required.

5.2.3 Dynamic Multi-Layer Routing Schemes

Another framework for dynamic multi-layer routing was proposed by Oki [78, 79]. Oki proposed two different policies to allocate the p-LSPs to an existing λ-LSP. If the $\lambda-LSP$ is not available then *policy 1* selects a sequence of existing λ-LSPs with two or more hops that connects the source and destination nodes (see Alg. 8). *Policy 2* selects and establishes a new one-hop λ-LSP as the new p-LSP (see Alg. 9).

Algorithm 8 Policy 1

```
candidate = any available existing λ-LSP that directly connects source and destina-
tion node pair with enough residual capacity.
if (candidate == NULL) then
   candidate = available existing λ-LSPs that connect source and destination node
   pair with two or H hops, where H is the maximum hop number.
   if (candidate == NULL) then
      candidate = new λ-LSP set up
      if (candidate == NULL) then
         return reject
      end if
   end if
end if
```

Algorithm 9 Policy 2

```
candidate = any available existing λ-LSP that directly connects source and destina-
tion node pair with enough residual capacity.
if (candidate == NULL) then
   candidate = new λ-LSP set up
   if (candidate == NULL) then
      candidate = available existing λ-LSPs that connect source and destination node
      pair with two or H hops, where H is the maximum hop number.
      if (candidate == NULL) then
         return reject
      end if
   end if
end if
```

Simulations presented by the authors showed that the number of PSC ports is a key factor in choosing the appropriate policy. Hence, policy 1 outperforms policy 2 only when p, the number of PSC ports, is small. This is because policy 2 first tries to set up new λ-LSPs first.

5.2.4 Network Survivability

The main drawback of the routing schemes reviewed above is that network connectivity is not guaranteed. An example is shown in Fig. 57. Lets suppose that a new p-LSP between nodes (1,3) is requested and a new λ-LSP (1,2,3), i.e. the λ-LSP$_1$, is set up according to the routing policies presented by Oki [79]. In this example, both policies presented by Oki give the same result. The same procedure is applied to set-up two new LSPs between the nodes (1,5) and (3,5) obtaining λ-LSP$_2$ and λ-LSP$_3$ respectively. Lets consider that the optical fibre (1,2) fails. Automatically the λ-LSPs λ-LSP$_1$ and λ-LSP$_2$ also fail. Considering only the IP/MPLS layer, node 1 is isolated, and the connectivity is lost, whilst the network has still enough resources to recover the failure. For instance, instead of selecting the optical fibres 1-2-5 and 5-2-3 for setting up λ-LSP$_2$ and λ-LSP$_3$ respectively, the optical fibres 1-4-5 and 5-6-2 should be selected. Thus, the connectivity will remain against any single fibre failure.

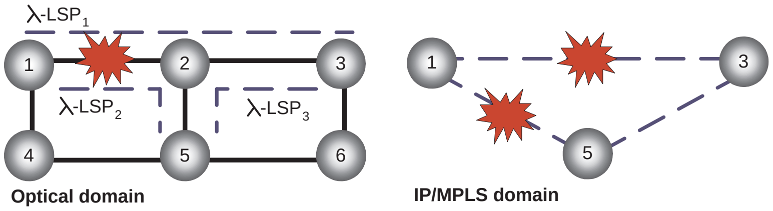


Figure 57: Loss of connectivity at IP/MPLS domain due to a single link failure.

5.3 Concluding Remarks

This chapter has analyzed the advantages and disadvantages of recovering failures at each network layer. Main points to highlight are:

- A new resource constraint is added to the network: the number of PSC ports. The PSC ports at intermediate nodes allow performing packet segment/local protection. However, the number of optical-electrical-optical (o-e-o) conversions increases.
- Better use of network resources is achieved by recovering at IP/MPLS layer due to its finer switching granularity.
- The recovery actions at optical domain are much faster and easier to manage than recovering at IP/MPLS domain, since the affected connections are recovered in group.

Thus, a trade-off exists between the resource consumption and the cost added to the network in terms of recovery time, failure management and node technology. Each layer has its pros and cons. Therefore, a cooperation between both layers seems to be the solution in order to take the advantages of each switching domain. Although this switching cooperation has been considered in the literature, the efforts have been only focused on the development of routing schemes without offering multi-layer protection against single link failures. Next chapter presents and evaluates a new QoSP multi-layer routing scheme that consider a cooperation between IP/MPLS and optical switching domain.

CHAPTER VI

QOSP ROUTING IN THE DYNAMIC MULTI-LAYER SCENARIO

TAKING ADVANTAGE OF IP/MPLS AND OPTICAL SWITCHING DOMAIN

In Chapter 5, some of the existing multi-layer routing schemes that take into account interoperability between each switching domain have been reviewed. However, these schemes are only focused on computing and establishing packet LSPs without offering protection against single fibre failures. This chapter provides new multi-layer routing schemes with protection where the cooperation between switching domains is considered for both p-LSP and backup p-LSP computation. Our proposal is compared to other algorithms that consider either full optical protection or full IP/MPLS protection. The performance of the algorithms are analyzed according to the network variables:

- The number of PSC ports, p.
- The number of wavelengths per fibre, w.

Moreover, the maximum number of hops (number of λ-LSPs than a p-LSP traverses), H, is also analyzed. Note that the maximum number of hops is an upper bound of the number of packet switching operations at intermediate nodes, $H-1$. Reducing H, the number of o-e-o operations is reduced.

6.1 Reliable and Dynamic Multi-layer Routing Scheme

In this section, our proposed QoSP routing scheme presented in [81] is depicted. This proposal is a first order approach that takes into account the multi-layer network scenario. It includes parameters related to the optical switching domain that are not currently considered in IP/MPLS-based QoSP routing algorithms. These parameters are:

- The Packet Switching Capable (PSC) ports.
- The λ-LSP protection status: protected or unprotected.
- The number of free wavelengths per fibre.

This proposal is based on the establishment of link-disjoint lambda/packet LSP pairs: the working lambda/packet LSP and the backup lambda/packet LSP. When a failure occurs at the working λ-LSPs, the traffic is switched to the respective backup λ-LSPs. If no backup λ-LSP exists, the traffic is switched to the respective backup p-LSPs. The main objective is to take advantage of both switching domains and provide protected paths with an efficient use of network resources.

6.1.1 Network Definition

Let $G_P = (V, E_P)$ and $G_L = (V, E_L)$ represent the physical topology and the logical topology respectively, where V is the set of photonic MPLS routers; E_P and E_L are the set of network physical links and λ-LSPs respectively. Each router has p input and output Packet Switching Capable (PSC) ports, where $PSCi(u)$ input ports and $PCSo(u)$ output ports of node u are already not assigned to any λ-LSP. Each physical link has w wavelengths. When a p-LSP is requested, the proposed routing scheme considers both physical links and λ-LSPs, i.e. $E_P \cup E_L$. In order to univocally identify the physical links and existing λ-LSPs that connect node pair (i, j) the 3-tuple (i, j, k) is used. Thus, the link (i, j, k) is a physical link if $k = 0$, otherwise $(k > 0)$ it is a λ-LSP.

Each (i, j, k) λ-LSP has an associated R_{ijk} residual capacity; S_{ijk}^{uv} total capacity reserved to protect the physical link $(u, v, 0)$; and T_{ijk} the total shared capacity allocated in link (i, j, k). Note that $T_{ijk} = \max_{(u,v,0)\in E_P} S_{ijk}^{uv}$. Each (i, j, k) λ-LSP is a sequence of physical links denoted as a set LP_{ijk} and a sequence of wavelengths assigned at each physical link denoted as LW_{ijk}.

The p-LSP request is defined by (s, d, b) where (s, d) is the source and destination node pair; and b, specifies the amount of capacity required for this request. For each request, a working p-LSP (WP) has to be set-up. A backup p-LSP (BP) must be also set-up, whenever the WP has, at least, one unprotected λ-LSP. If there are not sufficient resources in the network, for either the WP or the BP, the request is rejected.

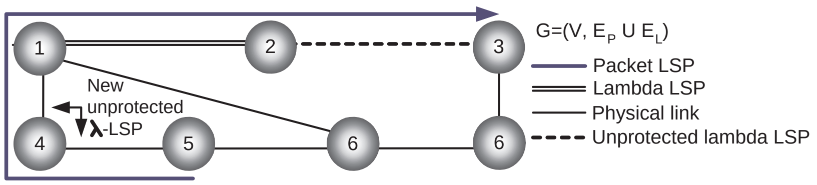


Figure 58: Working p-LSP computation. Creation of a new unprotected λ-LSP using the physical links (5,4) and (4,1).

6.1.2 Working Lambda and Packet LSP Computation

In the proposed scheme, a new procedure to compute the working p-LSP (WP) is presented. In this procedure the following cost parameters are taken into account:

1. The residual capacity of the link candidates, R_{ijk}.
2. The maximum number of hops, H, i.e. maximum number of λ-LSPs that the WP may traverse.
3. The free packet switching ports of each router, $PCSi$ and $PSCo$.

Note that the residual capacity of the physical links with free wavelengths is the capacity of the wavelength. The proposed procedure, called Dynamic Multi-Layer Working Path (DMWP) algorithm (Algorithm 10), computes the min-hop WP based on a variation of the Dijkstra algorithm. In this case, the number of hops coincides with the number of λ-LSPs. Thus, the consecutive sequence of physical links, that constitutes a λ-LSP, is only considered as one hop. The DMWP procedure uses the network graph composed by λ-LSPs and physical links, i.e. $G = (V, E_P \cup E_L)$. This procedure ends when it reaches the destination node or there is no feasible path between source and destination nodes. If a feasible path exists then the procedure may return:

1. A sequence of existing protected λ-LSPs.
2. A sequence of physical links. In this case, a new unprotected λ-LSP is set up between source and destination node.
3. A sequence of physical links, protected and unprotected λ-LSPs. In this case, new unprotected λ-LSPs are setup for each consecutive sequence of physical links as shown in Fig. 58. In this example, a new unprotected λ-LSP is set up with the physical links (5,4) and (4,1).

Algorithm 10 Dynamic Multi-Layer Working Path

```
INPUT
$(s, d, r)$: p-LSP request;
$G = (V, E)$: current network graph;
$H$: maximum hop number;
ALGORITHM
for all $v \in V$ do
  $Cost(v) = \infty$
  $Pred(v) = s$
  $WPlast(v) = s$
end for
$Cost(s) = 0$
$Q \leftarrow s$
while $(d \notin Q$ and $Q \neq \oslash)$ do
  $u \leftarrow min\_cost(Q)$
  $Q = Q - \{u\}$
  for all $v \in adjacency(u, G)$ do
    for all $(u, v, k) \in E$ do
      if $(R_{ijk} \geq b)$ and $((k = WPlast(u) = 0)$ or $(Cost(u) + 1 < Cost(v) < H))$ then
        if $(PSCi(v) > 0$ and $k = 0$ and $WPLast(u) > 0)$ or $(PSCo(v) > 0$ and $k > 0$ and $WPlast(u) = 0)$ or $(k = WPlast(u) = 0)$ or $(k > 0$ and $WPlast(u) > 0)$ then
          $Pred(v) = u$
          $WPlast(v) = k$
          $Q \leftarrow v$
          if not $(k = WPlast(u) = 0)$ then
            $Cost(v) = Cost(u) + 1$
          end if
        end if
      end if
    end for
  end for
end while
```

In the Dynamic Multi-Layer Working Path algorithm (Alg. 10), $Cost(v)$ is a vector containing the path cost from s to v; $Pred(v)$ contains the v's predecessor node; and $WPlast(v)$ contains the identifier k of link (u, v). Q represents the list of adjacent vertices which are not visited yet. Function $min\_cost(Q)$ returns the element $u \in Q$ with the lowest $Cost(u)$; and $adjacency(u)$ is the adjacency list of vertex u in graph G.

6.1.3 Backup Lambda and Packet LSP Computation

Once the WP is known, the backup p-LSP (BP) is computed. Three different procedures could be applied depending on the WP characteristics:

Step 1. If the WP is a sequence of existing protected λ-LSPs, the computation of the

BP is not required.

Step 2. If the WP is a new unprotected λ-LSP and an available and shareable backup λ-LSP exists, this is used to protect the λ-LSP. Otherwise, a new backup λ-LSP is set-up applying DMWP algorithm (Algorithm 10) with $G = (V, E_P)$. A backup λ-LSP is shareable if the new λ-LSP does not belong to the same Shared Rink Link Group (SRLG) [82] of both backup λ-LSP and λ-LSPs protected by this backup λ-LSP. If the procedure fails to find a backup λ-LSP, go to Step 3.

Step 3. If the WP is a combination of protected and unprotected λ-LSPs, then a variation of the Partial Disjoint Path (PDP) algorithm (Alg. 5) is used to compute the BP. The variations are the ones included to the Dijkstra algorithm in order to consider the packet switching ports in the DMWP algorithm (Alg. 10). The PDP may overlap with protected λ-LSPs of the WP, since they are already protected, and the nodes of the WP. Therefore, no extra resources are necessary in the IP/MPLS layer against failure of protected λ-LSPs in the optical layer. When the BP overlaps the WP, more than one segment backup paths are established.

6.2 Multi-layer Protection with Traffic Differentiation Routing Scheme

In this section, our novel reliable and dynamic routing scheme with traffic differentiation is presented. This is a first order approach for dynamic multi-layer routing that takes into account differentiated protection according to the traffic classification presented in section 2.4. This scheme is a variation of the above proposed routing algorithm. Thus, it takes also into account the parameters related to the optical switching domain: the PSC ports, the number of free wavelengths and the λ-LSP protection status.

6.2.1 Reliable Working Packet LSP Computation

A new procedure to compute the working p-LSP (WP) that takes into account the QoS protection requirements of the traffic class is presented. For high and medium reliable traffic, HR and MR, the DMWP algorithm (Alg. 10) is applied. In the case of low reliable traffic, LR, the WP is routed through unprotected λ-LSPs. Thus, if a feasible path exists for LR traffic, the procedure may return:

1. A sequence of existing unprotected λ-LSPs.

2. A sequence of physical links and unprotected λ-LSPs. In this case, new unprotected λ-LSPs are setup for each consecutive sequence of physical links.

Thus, optical protection will be only considered for protecting requests with HR and MR requirements.

6.2.2 Reliable Backup Lambda and Packet LSP Computation

Once the WP is known, the backup p-LSP (BP) is computed. Three different BP computations exist according to the requested level of reliability as follows:

High Reliability (HR). HR traffic requires fast protection. Hence, local protection is used to protect the unprotected λ-LSPs that the WP traverses. The protection will be optical if a shareable or new feasible backup λ-LSP exists. Otherwise, local backup paths at IP/MPLS layer will be computed. If the algorithm fails to find feasible local backup paths, then the request is rejected.

Medium Reliability (MR). The backup path for this traffic class is computed by means of applying the algorithm presented in 6.1.3. Thus, the unprotected λ-LSPs of the WP are either protected 1) at optical layer with path recovery or 2) at IP/MPLS and with global, segment or local backup path methods.

Low Reliability (LR). The WP is protected at IP/MPLS layer. A variation of the backup path routing algorithm presented in 6.1.3 is applied. It only considers the existing unprotected λ-LSPs when computing the BP. If a feasible BP does not exist, then the request is accepted but is unprotected.

6.3 Performance Evaluation

This section evaluates the routing schemes presented in this chapter using the NSF simulation model described in Section 3.3. These schemes are evaluated under the dynamic multi-layer simulation scenario for both incremental traffic and limited holding time cases. The metrics of interest to evaluate the algorithm performance are:

- *Request rejection ratio.*

- *Number of λ-LSPs and backup λ-LSPs.* This metric evaluates the total number of λ-LSPs that have been created during the simulation as well as the number of backup λ-LSPs. This metric is useful for analyzing the amount of protected and unprotected λ-LSPs.

- *Average of λ-LSPs per p-LSP.* This value represents the average number of λ-LSPs per packet LSP.

- *Average of physical links per λ-LSP.* This value is the average number of physical links per λ-LSP.

- *Average number of λ-LSPs protected per backup.* This parameter evaluates the number of λ-LSPs protected by a backup path. This average is the mean of the number of hops of the working path divided by the number of backup paths used to protect the working path. Thus for local protection (IP/MPLS) and path recovery (optical), this average is unity; for each λ-LSP of the working path a backup path is created to protect it.

Since both wavelength and packet switching is considered, the number of PSC operations (o-e-o conversions) may increase according to the protection method applied. Fault notification distance is not considered because the main difference between the proposed algorithms are the number of PSC operations.

6.3.1 Network Topology and Traffic Request Parameters

Each physical link has one bi-directional fibre with the same number of wavelengths in each direction. Each physical links has 12 wavelengths unless specifically stated otherwise. The transmission speed of each wavelength is set to 10 Gbps. The number of PSC ports p is the same in each node. The required p-LSP capacity is set to 500 Mbps.

6.3.2 Dynamic Multi-layer Scheme Evaluation

6.3.2.1 QoSP Routing Algorithms

Our proposed dynamic multi-layer routing scheme with Protection Against Single Fibre Failures (PASFF) is evaluated. PASFF computes the WP using the DMWP algorithm (Algorithm 10) and the BP according to the criteria presented in Section 6.1.3. In order to compare the merits of the new routing scheme, the following algorithms based on Oki policies [79] are also considered:

- Policy 1 with Protection (P1P). The routing policy 1 first tries to allocate the p-LSPs to an existing λ-LSP. If the λ-LSP is not available then a sequence of existing λ-LSPs with two or more hops that connects the source and destination nodes are selected. Otherwise, a new one-hop λ-LSP is established. When a new λ-LSP is created to accommodate the request, a backup λ-LSP is set up (path recovery).

- Policy 2 with Protection (P2P). The routing policy 2 first tries to allocate the p-LSPs to an existing λ-LSP. If the λ-LSP is not available then a new one-hop λ-LSP is established and selected as the new p-LSP. Otherwise, a sequence of existing λ-LSPs with two or more hops are selected. As in the case of the P1P algorithm, a backup λ-LSP is set up when a new λ-LSP is created.

If P1P and P2P fail to find a feasible p-LSP or backup λ-LSP, then the request is rejected. Note that protection is only applied at optical domain in both P1P and P2P algorithms by means of path recovery.

Table 21: Routing schemes for multi-layer protection evaluation.

Routing scheme	**Working path**	**Backup path**	**Protection domain**	**Switching architecture**
PASFF	DMWP (Alg. 10)	DMWP (Section 6.1.3)	IP/MPLS and optical protection	Multihop partial grooming
P1P	Policy 1	Backup λ-LSPs	Optical protection	Multihop partial grooming
P2P	Policy 2	Backup λ-LSPs	Optical protection	Multihop partial grooming
FIR	WSP	FIR	IP/MPLS protection	Multihop full grooming

As shown in Table 21, the Full Routing Information (FIR) algorithm presented in Section 3.4.2 is also considered in order to evaluate the performance of the new routing scheme when only IP/MPLS protection is applied. For the FIR algorithm evaluation, the photonic MPLS routers have a multihop full grooming architecture. Thus, the logical topology is exactly the same as the physical topology, i.e. all λ-LSPs are predetermined and unprotected (see Section 5.1.1).

The following parameters are also analyzed:

- H: The maximum number of λ-LSPs that a p-LSP may traverse. The number of hops is an important parameter since it cuts down the number of intermediate nodes that p-LSPs traverse. The larger is this number, larger is the number of packet switching operations.

- p: The number of PSC ports per node. This parameter is also analyzed in order to evaluate how the PSC ports may affect to each routing algorithm.

- w: The number of wavelengths per fibre. The number of wavelengths is also used to evaluate the behavior of the routing algorithms.

Note that the FIR algorithm is simulated into a *multihop full grooming*. Thereby, its performance is independent of p and w.

6.3.2.2 Simulation Results for Multi-layer Protection

Figure 59 shows the performance of the proposed algorithm PASFF compared to 1) optical oriented routing algorithms with protection, P1P and P2P, and 2) IP/MPLS oriented routing algorithm with protection, FIR algorithm. Results show that the proposed PASFF algorithm outperforms P1P and P2P routing algorithms because of the finer granularity. The P2P algorithm is practically independent of the number of hops because of the first-create procedure used to compute the p-LSP: if there is not a direct existing λ-LSP a new λ-LSP is established. Hence, most of the p-LSPs have low number of hops. However, each λ-LSP may traverse several physical links, consuming high amount of wavelengths. On the other hand, FIR presents a sharp increase in the

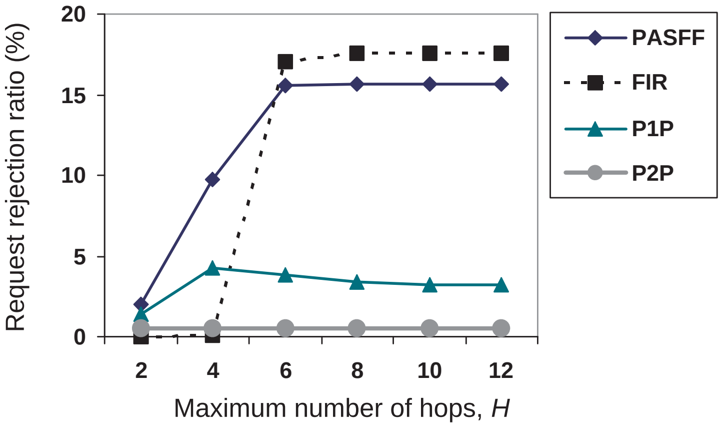


Figure 59: Number of hops analysis ($p = 10$).

request rejection ratio from $H = 6$. Although the diameter of the NSF network, δ, is 3, there are no many disjoint paths with number of hops $\leq H$ and, thus, many requests are rejected for $H < 6$.

Next two results show the influence of the number of PSC ports per node for all routing algorithms when $H = 4$ and $H = 6$ (see Fig. 60). Although the FIR algorithm operates under multihop full grooming ($p = w$), the results are shown in order to present the IP/MPLS bound of the solution in terms of capacity when $H = 6$. Again, the PASFF algorithm results in better use of the network resources compared to P1P and P2P. When p is small, the rejections are due too few available PSC ports and, for all, optical protection is applied.

Figure 61 shows the influence of the number of wavelengths per fibre for all routing algorithms when $p = 10$ and $H = 4$ and $H = 6$. As shown, the number of rejected requests lineally increases for FIR algorithm when $H = 6$. Moreover, since P2P prioritize λ-LSP that directly connects source and destination nodes, it outperforms P1P when $w > 24$. Plus, P2P also offers better performance than PASFF when $H = 4$ for $w > 24$. Note that PASFF and FIR behavior sharply change according to the maximum number of hops (see Fig. 59) while P1P and P2P do not.

From these results, it can be concluded that the PASFF algorithm allows decreasing the rejected requests due to the finer recovery granularity at the IP/MPLS domain. Additionally, PASFF outperforms FIR algorithm when the number of packet switching operations is reduced (number of hops); when H is less than 6. Moreover, when $H \geq 6$,

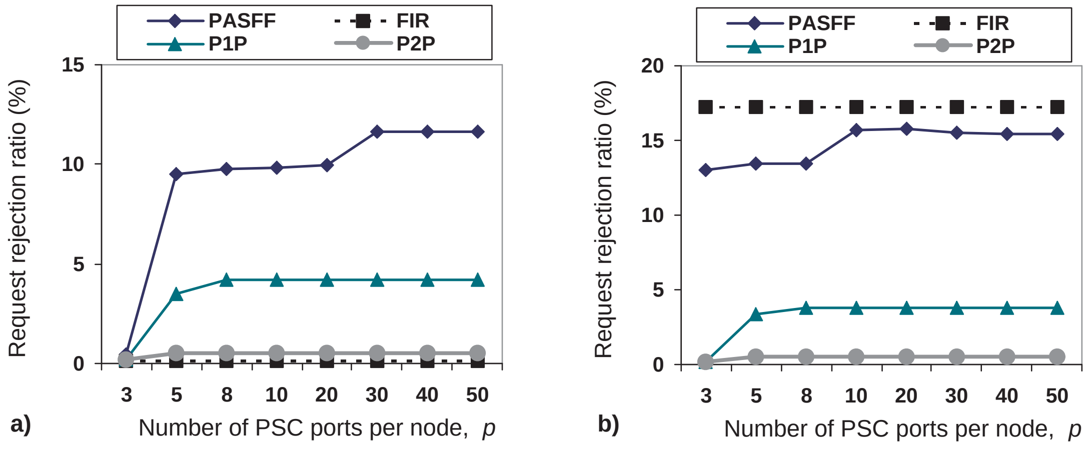


Figure 60: Number of PSC ports per node analysis when a)$H = 4$ b) $H = 6$.

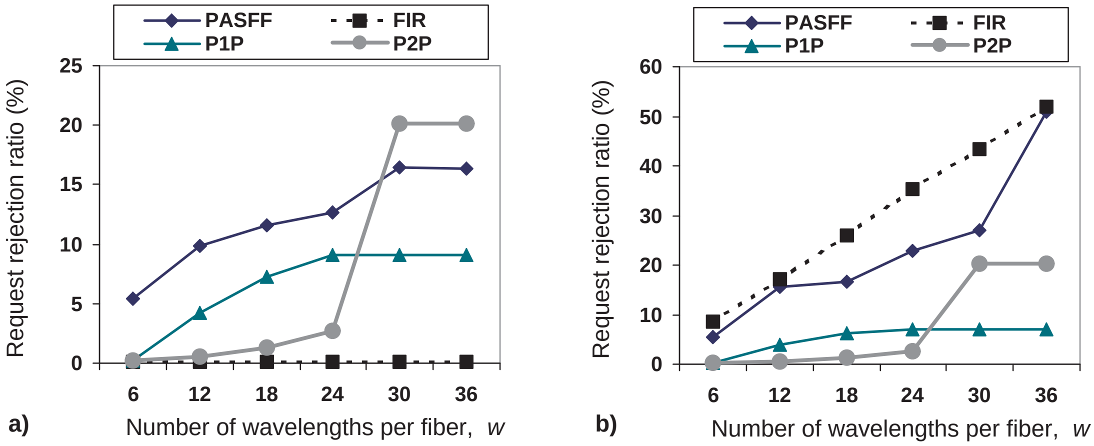


Figure 61: Number of wavelengths per fibre analysis when $H = 10$ and a)$H = 4$ b)$H = 6$.

PASFF only outperforms FIR when the network nodes have high number of PSC ports.

Lets now analyze the resource consumption of each routing algorithm. First, the total number of λ-LSPs and backup λ-LSPs established is evaluated in Fig. 62. Although resources have been analyzed varying H and w, the case of $H = 4$ and $w = 18$ is only plotted for clarity since the behavior of all the algorithms is similar in all cases in terms of network resources. Figure 62a shows the total number of λ-LSPs created. Since FIR operates under *multihop full grooming*, each wavelength is seen as a λ-LSP. Knowing that 1) the number of links of the NSF network is 21, 2) there is a bi-directional fibre per link and 3) each fibre has 18 wavelengths; the total number of λ-LSPs in the network for FIR algorithm is $21 \cdot 2 \cdot 18 = 756$. This number is an upper bound of the maximum number of λ-LSPs that may be established. In the PASFF algorithm case, when the

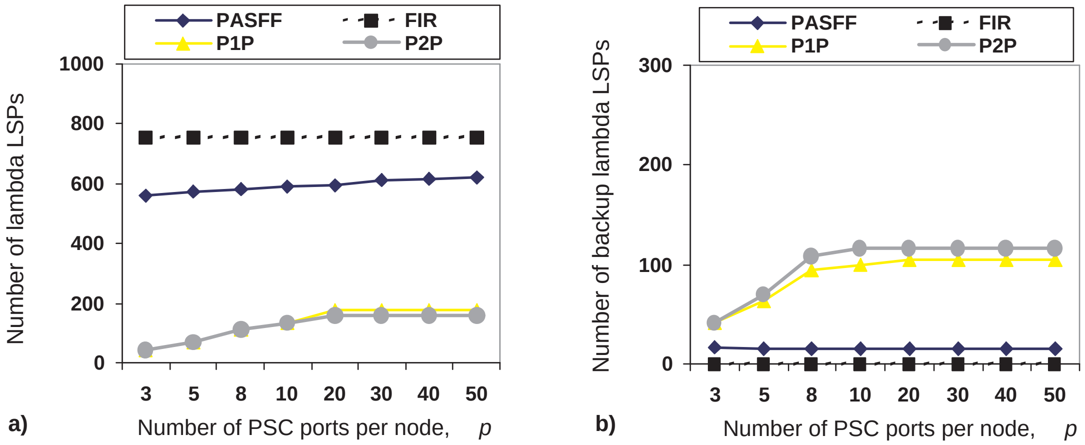


Figure 62: Total number of a) λ-LSPs and b) backup λ-LSPs for $H = 4$ and $w = 18$.

number of PSC ports increases, the number of new λ-LSPs slightly increases. On the other hand, the number of new λ-LSPs sharply increases from $PSC = 3$ to $PSC = 10$ for P1P and P2P algorithms. The number of PSC ports has higher impact to P1P and P2P algorithms because of the full optical protection applied. This is shown in Fig. 62b, where the curve of new backup λ-LSPs has similar behavior than the one of new λ-LSPs for P1P and P2P algorithms. However, although P1P has similar number of new λ-LSP than P2P, it has lower number of new backup λ-LSPs respect to P2P. Note that each new λ-LSP is either protected by a new backup λ-LSP or is sharing an existing backup λ-LSP. Thereby, P1P algorithm shares higher number of backup λ-LSPs. In the case of PASFF algorithm, few λ-LSPs are optically protected because most of the failures are recovered at IP/MPLS domain.

Figure 63 analyzes the average of hops of the logical links (λ-LSPs) and the packet

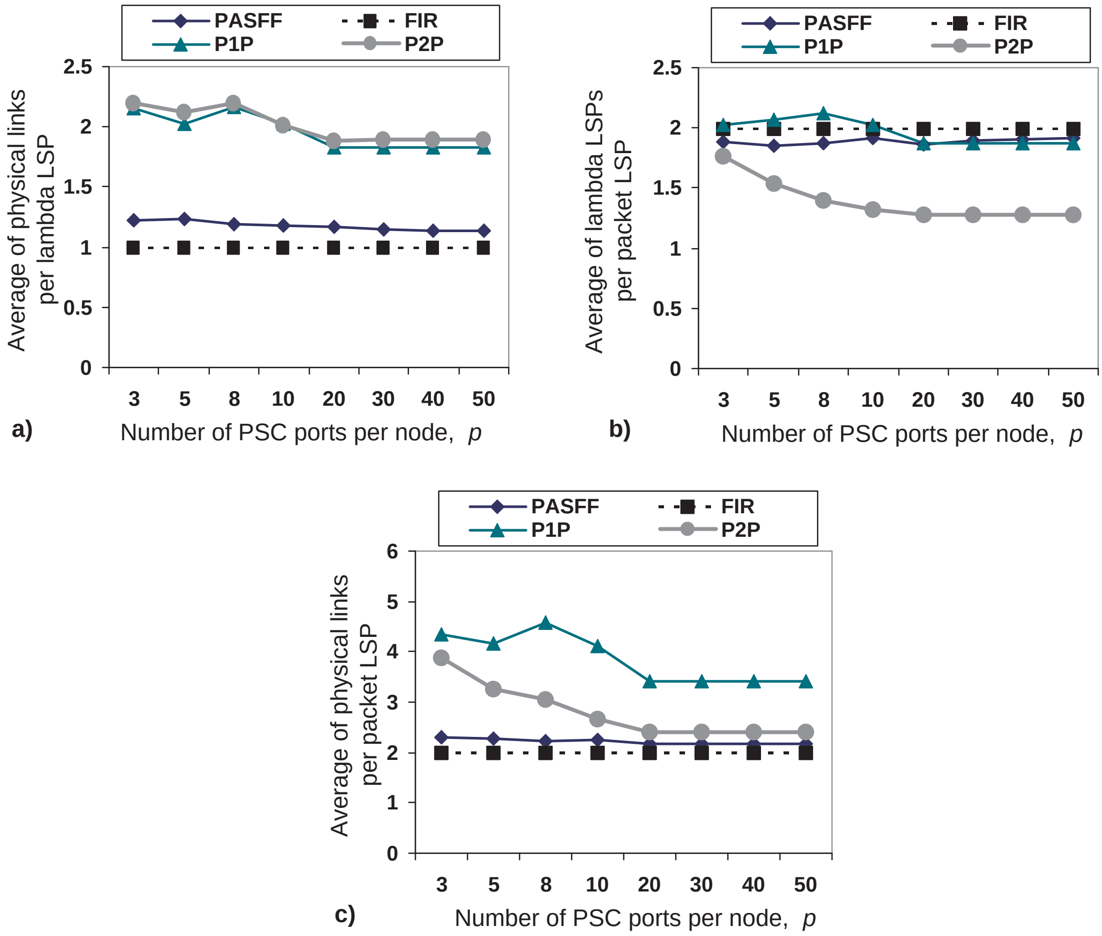


Figure 63: Average of a) physical links per λ-LSP b) λ-LSPs per p-LSP and c) physical links per p-LSP, for $H = 4$ and $w = 18$.

LSPs. According to the number of PSC ports, the average number of hops tends to decrease for all the routing algorithms, excluding FIR. Both P1P and P2P algorithms result into higher average number of physical links per λ-LSP compared to PASFF and FIR, as shown in Fig. 63a. Moreover, P2P results into low average number of λ-LSPs per p-LSP since it gives priority at creating new λ-LSPs for each request, see Fig. 63b. On the other hand, the rest of the algorithms offer an average of two λ-LSPs per p-LSP. Taking into account that $H = 4$ then the p-LSPs may traverse 1, 2, 3 or 4 λ-LSPs. Thus, the theoretical average number is $\frac{1+2+3+4}{4} = 2.5$. Thereby, the new p-LSPs have usually less than 4 λ-LSPs when P1P, FIR and PASFF algorithms are applied. Finally, the total number of physical links per p-LSP is analyzed in 63c. Although for FIR and PASFF, p-LSPs result in low use of physical links, P2P offers similar amount of hops due to its first-create policy when the number of PSC ports increases. Note that the best algorithm in terms of hops is P2P; it requires low amount of packet switching operations. However, it suffers from high request rejection ratio.

It can be concluded that FIR and PASFF are the best algorithms in terms of network resources. The use of IP/MPLS recovery mechanisms with finer granularity than the optical recovery allows better allocation of the p-LSPs. However, when the number of packet switching operations is limited (H), the FIR algorithm may result in the worst request rejection ratio. On the other hand, PASFF still offers better request rejection ratio than P1P and P2P independently of the number of hops. Thereby, full recovery at IP/MPLS is more efficient than applying other routing algorithms. However, for low H, PASFF should be chosen to compute new p-LSPs and their backup; reducing the number of o-e-o operations.

6.3.3 Dynamic Multi-layer Scheme Evaluation with Traffic Differentiation

6.3.3.1 QoSP Routing Algorithms

In this section, the dynamic Multi-layer Protection with Traffic Differentiation (MPTD) routing scheme, described in Section 6.2 is evaluated. In order to compare this new routing algorithm, two algorithms without multi-service differentiation that have been evaluated previously are also considered: PASFF and FIR. These algorithms are, according to H value, offering the best performance in terms of request rejection ratio when no differentiated protection is considered. Table 22 summarizes the characteristics of each routing algorithm according to the level of reliability required by each traffic class.

Table 22: Routing schemes for multi-layer protection with traffic differentiation evaluation.

Routing schemes	Traffic Services			Router functions
	LR	MR	HR	
MPTD	IP/MPLS shared local, segment, global protection, if enough resources. Slow recovery time	IP/MPLS shared local, segment, global protection. Optical path recovery. Medium, fast recovery time.	IP/MPLS shared local protection. Optical path recovery. Fast recovery time.	Multihop partial grooming
PASFF	IP/MPLS shared local, segment, global protection. Optical path recovery. Medium, fast recovery time.			
FIR	IP/MPLS shared local, segment, global protection. Medium/slow recovery time.			Multihop full grooming

In this section the algorithms are evaluated under incremental traffic. The metrics of interest for these simulations are: the request rejection ratio and the average number of λ-LSPs protected per backup. In this set of simulations the parameters H, p, and w are also analyzed.

6.3.3.2 Simulation Results for Multi-Layer Protection with Traffic Differentiation

As shown in Fig. 64, the algorithms that have in their objectives the dynamic multi-layer information, MPTD and PASFF, outperform FIR algorithm when H is small. Moreover, including different levels of protection improves the network resource consumption depending on H. When reducing the number of PSC operations, small H, MPTD offers higher accepted requests than the other algorithms because it does not protect LR requests when there are not sufficient network resources. However, MPTD degrades when H increases, due to the local backup paths established to protect HR requests. The higher is H, higher is the number of local backup paths that must be established to protect the working path.

Next two results show the influence of the number of the PSC ports per node when $H = 4$ and $H = 6$ (see Fig. 65). FIR results are plotted in order to show the IP/MPLS bound of the solution. The MPTD algorithm results in better use of the network resources compared to PASFF through most of the cases, except when $p = 10$ and $H = 6$. When p is small, the rejections are due too few available PSC ports. Thus, for larger p, MPTD also improves FIR.

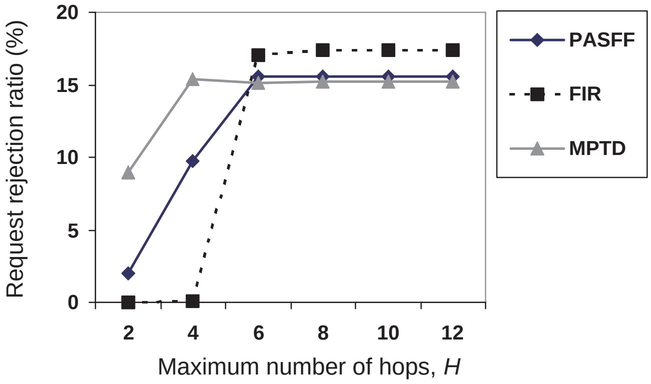


Figure 64: Number of hops analysis ($p = 10$).

Figure 66 shows the influence of the number of wavelengths per fibre when $p = 10$ and $H = 4$ and $H = 6$. MPTP does not necessarily protect the LR requests, thus, it outperforms PASFF as shown in Fig. 66a. On the other hand, when $H = 6$, MPTP does not improve FIR because of the high number of local backup paths that are established.

From these results, it can be concluded that the MPTD algorithm allows decreasing the number of rejected requests due to the differentiated reliability when number of PSC operations, H, is low. However, for larger H the performance of the MPTD decreases due to the high number of local backup paths that should be established in order to protect HR requests.

Finally, the average number of λ-LSPs protected per backup path is evaluated in Fig. 67 according to each traffic class. As expected MPTD creates a backup path per λ-LSP for all HR traffic (IP/MPLS local protection or optical path recovery) offering the faster recovery time. However, it suffers of high resource consumption for large H

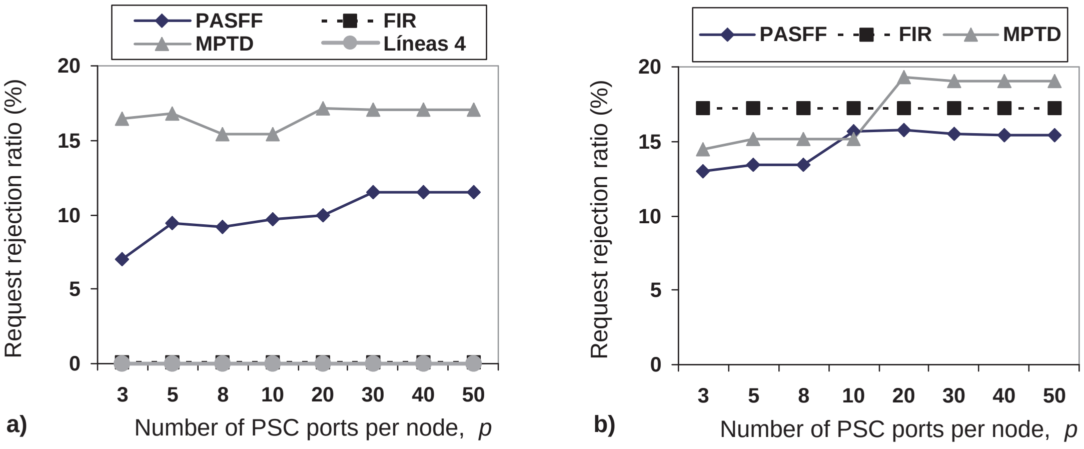


Figure 65: Number of PSC ports per node analysis when a)$H = 4$ b)$H = 6$.

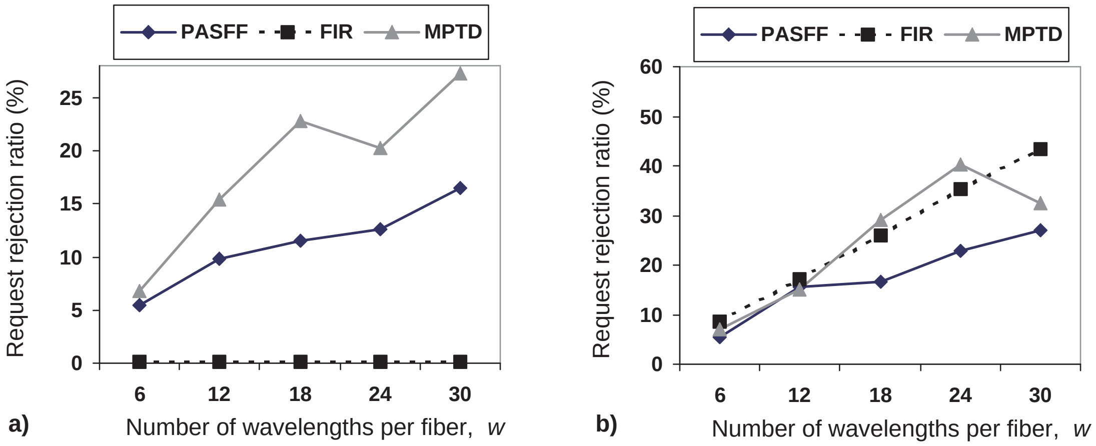


Figure 66: Number of wavelengths per fibre analysis when $H = 10$ and a)$H = 4$ b) $H = 6$.

as shown in Fig. 64. Moreover, MR and protected LR requests offer better performance than FIR. In the FIR algorithm case, the backup paths protect an average of two λ-LSPs, offering the slower recovery time because of the faut notification time. Therefore, the MPTD algorithm facilitates a decrease in rejected requests due to the finer recovery granularity at the IP/MPLS domain. Additionally, MPTD outperforms FIR algorithm when either the number of packet switching operations is reduced or when the network nodes have high number of PSC ports. Moreover, the proposed MPTD meets the QoSP requirements of each traffic class.

6.4 Concluding Remarks

In this chapter, as a last contribution of this thesis new routing schemes have been proposed: the Protection Against Single Fibre Failures (PASFF) and the Multi-layer Protection with Traffic Differentiation (MPTD) schemes. Both proposed routing schemes consider a dynamic cooperation between packet and wavelength switching domain in order to minimize the resource consumption and provide the QoSP requirements of each traffic class.

The finer granularity of both proposed schemes results into better filling of the capacity and less number of rejected requests comparing to routing schemes that apply protection at optical domain. Moreover, the proposed schemes outperform the algorithms that only consider IP/MPLS recovery, when the number of packet switching operations is reduced.

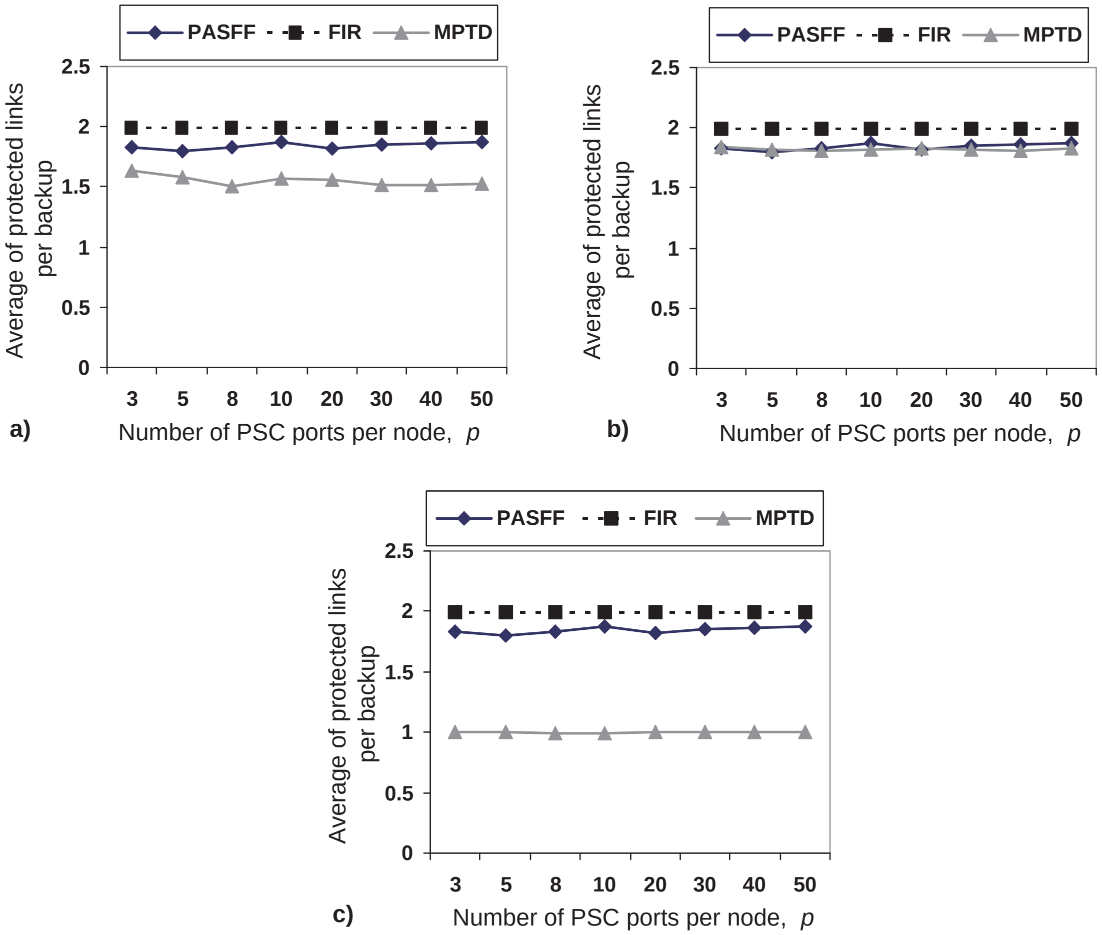


Figure 67: Average of λ-LSPs protected per backup for a) LR traffic b) MR traffic and c) HR traffic.

The second proposed scheme, Multi-layer Protection with Traffic Differentiation (MPTD), takes into consideration different levels of reliability depending on the QoSP traffic class requirements. Results have shown that the proposed MPTD algorithm meets the requirements of all the traffic classes without adding resource consumption and request rejection ratio of previous proposals that did not consider traffic differentiation when the network load is low. Additionally, our proposed algorithm improves upon the previous ones when the number of packet switching operations is reduced, decreasing the number of o-e-o operations.

CHAPTER VII

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this thesis new routing algorithms have been developed to support the Quality of Service with Protection (QoSP) required by the new traffic services in the next generation GMPLS-based optical networks. This chapter discusses the research contributions and the directions in which this research might be focused on the future.

7.1 Research Contributions

Survivability has become a crucial issue for the next generation backbone networks, IP/MPLS over optical network. Single fibre failures occur frequently, causing disruptions in the service of the affected applications. Quality of Service with Protection (QoSP) routing algorithms are defined in order to guarantee that the affected traffic reaches the destination node even though a single fibre failure occurs. The multi-layer network scenario is analyzed in order to make an efficient use of the network resources and avoid protection duplications simplifying the network management.

7.1.1 Survivability Techniques Overview

The two survivability techniques, protection and restoration, have been considered. When recovering the traffic, protection mechanisms are faster than restoration, since it does not need to wait for the establishment and reservation of the alternative path. However, protection mechanisms cost more resources, since it needs to pre-allocate spare capacity for pre-establishing backup paths. The reduction of the recovery time is one of the main aspects to consider in order to reach the level of reliability required by many current traffic services. Reducing the recovery time, traffic delay and packet loss are also reduced. In Chapter 2, the recovery time has been evaluated and the recovery phases that may be reduced have been identified. The reduction of the spare capacity may be achieved by sharing the spare capacity. Shared protection saves a large amount of resources by maintaining the same level of protection for single fibre failures. This is the common case and the only considered in this thesis.

7.1.2 Failure Probability Formalization

The network reliability, in terms of failure probabilities, has been also considered. Reducing the failure probability of a working path, the reliability of the transported traffic increases. This reduction can be offered by taking into account the link failure probabilities. Note that in the multi-layer scenario IP/MPLS optical networks, two kind of links can be identified: the physical links and the logical links (lightpaths). A formalization of the failure probability adapted to the multi-layer network scenario has been presented in Chapter 2.

7.1.3 Traditional QoS and QoSP Routing

In Chapter 3, the following QoS and QoSP routing metrics:

1. shortest path,
2. load balancing,
3. minimum interference,
4. network information available, and
5. recovery time reduction,

have been analyzed and evaluated in terms of network resources, request rejection ratio and recovery time. The shortest path and load balancing metrics implemented by the traditional QoS routing algorithms, such as WSP and SWP, are simple to deploy using the minimum link state information which is provided by current OSPF/IS-IS extensions. However, they do not consider the capacity sharing during the backup path selection and, therefore, the selected path may not reflect the maximum possible capacity sharing. Additionally, although more sophisticated and efficient QoS routing proposals, such as the current minimum interference proposals, aim to reduce the block probability; dedicated spare capacity is only considered resulting into large amount of spare capacity. Significant reductions in spare capacity can be achieved by sharing this capacity. The accuracy and performance of the shared backup routing algorithms are based on the available network information. Partial and full information routing schemes reduce significantly the spare capacity. However, the proposed schemes in the literature are oriented to path protection (global backup path method), reporting high recovery time. Two routing algorithms have been proposed in Chapter 3, where both local backup method and spare capacity have been considered in order to offer fast protection and efficient use of network resources. Results have shown that if the objective is to minimize the blocking probability then partial information routing algorithm should be chosen, simplifying the management of the network. On the other hand, for medium/high network loads, full information routing algorithm and local backup method should be chosen. All these metrics have been evaluated into a single network layer without taking into account the multi-layer network scenario.

7.1.4 QoSP Routing Algorithms for the Static Multi-layer Network Scenario

In the *static multi-layer network scenario*, the logical topology where the LSPs are routed is given and pre-established. The lightpaths are pre-established and some of them are assumed to be already protected at the optical layer. This static information from the lower layers is added in the metrics of the proposed QoSP routing algorithms for IP/MPLS networks presented in Chapter 4. Thus, an enhancement is achieved by avoiding the protection of those lightpaths that are already protected at the optical layer. The lightpath failure probability concept has been introduce to identify the status of the lightpath: protected or unprotected.

In order to deploy this idea, a new definition of link-disjoint path based on Shared Risk Link Group (SRLG) concept is presented. As a novelty, the alternative path is proposed to be a Partial Disjoint Path (PDP). The PDP may overlap the protected lightpaths of the working path. In order to guarantee fast protection, the proposed algorithms also combine segment protection and shared spare capacity, resulting in a suitable fault recovery time and resource consumption. A complete set of simulations has proved the efficiency of the proposed algorithms.

7.1.5 Dynamic Multi-layer Network Scenario

An analysis of the advantages and disadvantages of recovering failures at each network layer has been presented in Chapter 5. When the dynamic multi-layer network scenario is considered, new resource constraints are added to the network: the number of Packet Switching Capable (PSC) ports and the number of wavelengths. The PSC ports at intermediate nodes allow performing packet segment/local protection. However, the number of optical-electrical-optical (o-e-o) conversions increases.

A trade-off exists between the resource consumption and the cost added to the network in terms of recovery time, failure management and node technology. A better use of the network resources is achieved by recovering at IP/MPLS layer due to its finer switching granularity. On the other hand, the recovery actions at optical domain are much faster and easier to manage than recovering at IP/MPLS domain, since the affected connections are recovered in group. A cooperation between both switching layers seems to be the solution in order to take the advantages of each switching domain.

7.1.6 QoSP Routing Algorithms for the Dynamic Multi-layer Network Scenario

In the *dynamic multi-layer network scenario*, cooperation between each switching layer, optical and IP/MPLS, has been considered. In the proposed QoSP routing algorithms, whenever a new Label Switched Path (LSP) request arrives, the decision of setting up new lightpaths, backup lightpaths and backup LSPs is evaluated. Additionally, whenever a LSP is torn-down, the respective lightpaths and backup lightpaths that do not accommodate any other LSP are disconnected. The number of PSC ports constraint is also added to the network design. The proposed QoSP routing algorithms have been compared to other algorithms that consider either full optical protection or full IP/MPLS protection. The performance of the algorithms has been analyzed according to the network constraints: number of PSC ports, number of wavelengths per fibre and number of hops (lightpaths). The maximum number of hops is an upper bound of the number of packet switching operations at intermediate nodes. Reducing the number of hops, the number of optical-electrical-optical operations is also reduced. A complete set of simulations has proved the efficiency of the proposed algorithms. The proposed algorithms result in a reduction in rejected requests; a consequence of the finer granularity associated with IP/MPLS recovery. Moreover, they outperform IP/MPLS protection-oriented algorithms when the number of packet switching operations is reduced, decreasing the number of o-e-o operations.

7.1.7 Traffic Classification

A traffic classification has been presented according to the QoSP requirements of the traffic. Note that when a failure occurs not all the applications affected by the failure require the same QoSP. From the ISPs perspective, low network cost should be provided whilst achieving the survivability requirements of the requests that are expressed in terms of availability and recovery time. Thus, four traffic classes are considered: high, medium, low and null reliability. QoSP routing algorithms that take into account the presented traffic classification have been also proposed and evaluated under the static and dynamic multi-layer network scenarios.

7.2 Future Research Direction

In this section the main issues to cover in further research are presented.

7.2.1 Failure Coverage

This thesis has been focused on protecting traffic against *single fibre failures* since it is the most common impairment in the current networks. Further work should cover next failure scenarios:

- *Node failures.* Protection against node failures should be also considered in both static and dynamic multi-layer network scenario. Although routing algorithms that address the link and node failures [83, 84] exist, they are oriented to one switching layer, either optical or IP/MPLS. This case can be faced as the failure of the adjacent links of the failed node.

- *Dual failures.* Dual failures become more probable in larger networks [85] and should be considered when planning and operating such networks. Moreover, to evaluate the availability of service paths in a survivable network, the way that a given survivability mechanism will react to multiple failures should be analyzed. Amongst multiple failures, dual fibre failures are the ones that contribute the most service outage, so studying this type of failure provides a good comparative estimate of the availability of service [86] Although research proposals for protecting traffic against dual fibre failures exist [87], they do not consider the dynamic multi-layer network scenario presented in this thesis.

7.2.2 Routing Information and Signaling Overhead

Although the QoSP routing algorithms have been analyzed and have shown an improvement respect to other QoS routing algorithms, the signaling information transmitted through the network should also be analyzed as follows:

- *Signaling overhead.* Two classes of messages should be sent: 1) fault notification messages to the the nodes that are responsible of the switchover and 2) activation messages, in order to activate the backup paths. As explained in Chapter 5.1.4, the lower is the number of signaling messages, the easier and faster to manage is the recovery action.

- *Routing information.* The routing information must be updated. Although in Chapter 3.2.3 the routing information has been analyzed for static multi-layer network scenario, a study for the dynamic case should be considered. The IETF CCAMP working group is currently studying the multi-layer network scenario in [88, 89].

Different mechanisms to overcome the time required to send these messages can be proposed for further work. Note that the granularity of the recovery affects to the number of messages to be sent (signaling overhead) as well as the amount of information to update.

7.2.3 Reducing the Traffic Redundancies

One other aspect to be included in this research is the reduction of the *traffic redundancies*. Note that when IP/MPLS connections are set up, a sequence of lightpaths is chosen. However, this set of lightpaths may have some physical links in common. The traffic redundancy of a packet LSP may be defined as the ratio of times that traffic is sent through the same physical link. For a network graph $G = (V, E_P \cup E_L)$, the LSP redundancy, LSP_r can be formulated as follows:

$$LSP_r = \frac{\sum\limits_{p \in E_P} \sum\limits_{l \in LSP} \Omega_{lp}}{\sum\limits_{l \in LSP} h_l} - 1 \tag{21}$$

Where h_l is the number of physical links that the lightpath l traverses and Ω_{lp} returns 1 if the lightpath l of the LSP traverses the physical link p; 0, otherwise.

An example is shown in Fig. 68 into a dynamic multi-layer network scenario. Lets assume that the routing algorithm first tries to allocate the packet LSPs on the existing lightpaths. Three packet LSP requests arrive in the following order: LSP_1 connecting source and destination node pair $(1, 3)$, LSP_2 connecting node pair $(3, 5)$ and, finally, LSP_3 connecting $(1, 5)$. When LSP_3 request arrives, it is routed using the existing lightpaths λ-LSP_1 and λ-LSP_2, which have been created to accommodate the previous requests, LSP_1 and LSP_2 respectively. Thus, the LSP_3 is traversing twice the physical link $(2 - 3)$. Table 23 shows the redundancy ratio of each LSP according to Eq. 21.

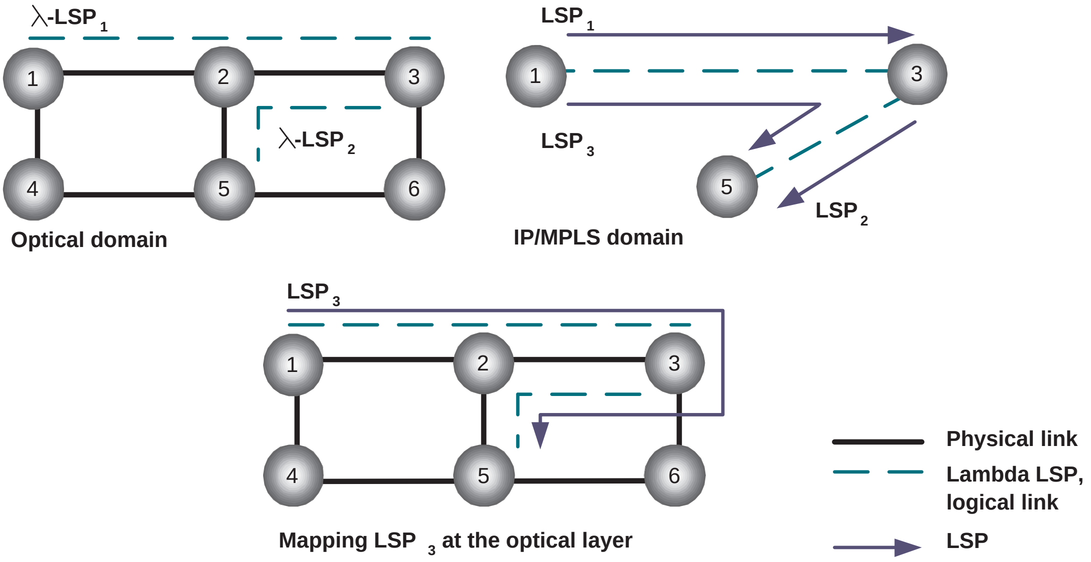


Figure 68: Traffic redundancy. Illustrative example.

Table 23: Traffic redundancy evaluation according to Fig. 68.

LSP	Lightpath	Lighpath links	Redundancy, LSP_r
LSP_1	λ-LSP_1	$(1 \xrightarrow{\alpha} 2 \xrightarrow{\alpha} 3)$	0
LSP_2	λ-LSP_2	$(3 \xrightarrow{\beta} 2 \xrightarrow{\beta} 5)$	0
LSP_3	λ-LSP_1 λ-LSP_2	$(1 \xrightarrow{\alpha} 2 \xrightarrow{\alpha} 3)$ $(3 \xrightarrow{\beta} 2 \xrightarrow{\beta} 5)$	$\frac{1}{3}$

7.2.4 Bounding the Multiple Logical Link Failures

Chapter 1.3 has shown that a physical link failure leads to multiple logical links (lightpaths) failures. When the dynamic multi-layer network scenario is considered, the lightpaths are dynamically established according to the packet LSP request arrivals. Current research in this area does not consider the problem of reducing the impact at the IP/MPLS layer of a single fibre failure in terms of logical link failures.

The maximum impact of any network single link failure, I, may be evaluated as follows:

$$I = \max_{p \in E_P} \Big(\sum_{l \in E_L} \Omega_{lp} \Big) \tag{22}$$

Where Ω_{lp} returns 1 if the logical link l traverses the physical link p; 0, otherwise.

Reducing the impact I can be achieved by balancing the logical links and their backups as shown in Fig.69. When no balancing is considered, Fig.69a, the failure of either the physical link $1-2$ or $2-3$ or $2-5$, leads to two logical failures at IP/MPLS layer: the logical links L_1, L_2; L_1, L_3; and L_2, L_3 respectively. Thus, $I = 2$. On the other hand, considering balancing, Fig.69a, the failure of each link leads to one logical link failure, $I = 1$.

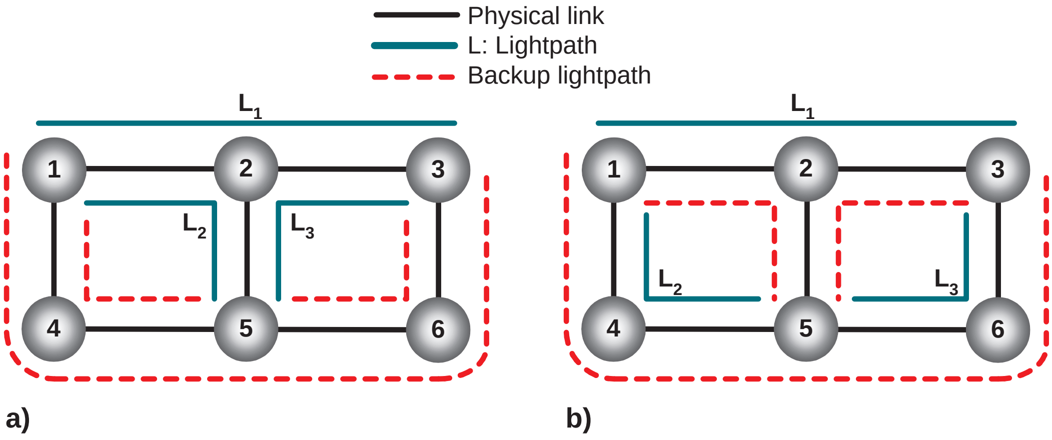


Figure 69: Impact of a single link failure. Illustrative example.

7.2.5 Bi-directional WDM Transmission System

For the dynamic multi-layer network scenario, the routing algorithms should be adapted to the new and future optical technology. One interesting further work should address the routing algorithms for *bi-directional WDM transmission system* [90]. Bi-directional

WDM transmission system uses part of wavelengths in one fibre for transmitting data in one direction and the rest in the opposite direction. It is more important for such a system to properly separate/isolate the wavelengths running in the opposite directions. One technique used to implement this system is by means of introducing *circulators* at the nodes [90]. A circulator is a multi-port device that allows signals to propagate in certain directions based on the port that the signal came from and blocks all transmission in other directions. Thus, if two adjacent nodes, x and y, are provided of circulators, then they can use a wavelength, i, to send information either from x to y or y to x. This property, if considered, may improve the routing performance when dynamic multi-layer network scenario is considered or when photonic MPLS routers are provided of *single-hop grooming*. An example is shown in Fig. 70. In unidirectional sharing, four units of capacity along backup λ-LSP_1 and backup λ-LSP_2 should be reserved. In bidirectional sharing, only two units are needed since the reserved wavelength can be used in either direction. The wavelength used for backup may be shared though the traffic is transmitted in opposite direction. The SRLG concept must be taken into account and redefined. Thus, those paths that are fibre disjoints may share the wavelength used for protection, independently of the direction of the backup traffic on that wavelength. Thereby the number of wavelengths used for protection may be reduced.

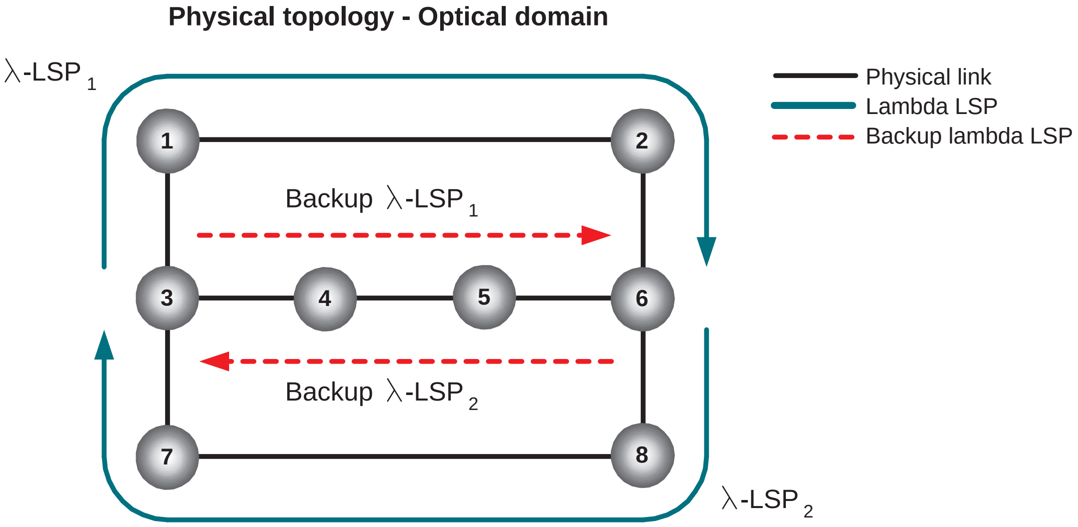


Figure 70: Illustrative example of unidirectional and bi-directional backup sharing.

Table 24: Network resources evaluation according to Fig. 70.

Transmission	Backup ID	Backup lightpath	Number of wavelengths	Total
Unidirectional Sharing	Backup λ-LSP_1	$(3 \xrightarrow{\alpha} 4 \xrightarrow{\alpha} 5 \xrightarrow{\alpha} 6)$	$1+1+1$	6
	Backup λ-LSP_2	$(3 \xleftarrow{\alpha} 4 \xleftarrow{\beta} 5 \xleftarrow{\alpha} 6)$	$1+1+1$	
Bi-directional Sharing	Backup λ-LSP_1	$(3 \overset{\alpha}{\longleftrightarrow} 4 \overset{\alpha}{\longleftrightarrow} 5 \overset{\alpha}{\longleftrightarrow} 6)$	$1+1+1$	3
	Backup λ-LSP_2			

7.2.6 Optimization Models

Enhancing current QoSP routing algorithms to offer better network protection has been one of the main objectives of this thesis. The complexity of the routing algorithms depends on the method and objectives of the routing applied. If the goal of the algorithms is to find an optimal solution, the utilisation of network optimization models should be considered. The solution is achieved by applying off-line routing algorithms. Further work in the analysis of current network optimization models and the application of these models to the proposed schemes should be considered for future work in order to evaluate the performance of the presented algorithms respect to the optimal case.

APPENDIX A

NETWORK TOPOLOGIES

A.1 NSF network

The NSF network topology [78] has 14 nodes and 21 physical links, as shown in Fig. 71. Each adjacent node pair is connected through a bi-directional physical link that consists of two fibers. Table 25 shows the link length in miles. The average node degree is 3 and the network diameter is 3.

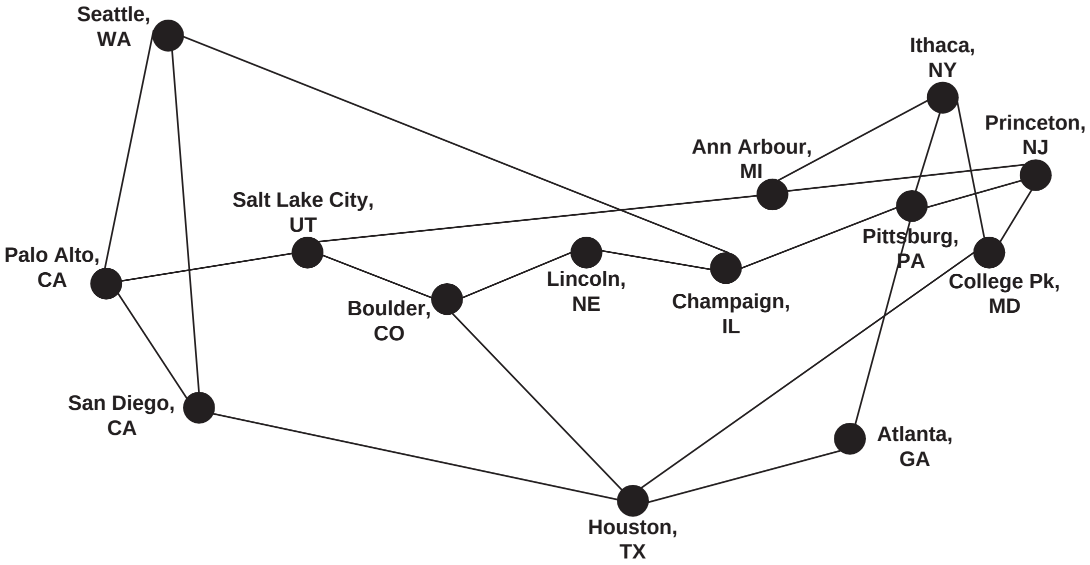


Figure 71: NSF network.

Table 25: NSF network: distance matrix (miles).

Name		13	12	11	10	9	8	7	6	5	4	3	2	1	0
Seattle, WA	**0**	0	0	0	0	0	0	1600	0	0	0	0	1000	700	0
Palo Alto, CA	**1**	0	0	0	0	0	0	0	0	0	0	500	400	0	
San Diego, CA	**2**	0	0	0	0	0	0	0	0	1200	0	0	0		
Salt Lake City, UT	**3**	0	0	0	1300	0	0	0	0	0	400	0			
Boulder, CO	**4**	0	0	0	0	0	0	0	400	800	0				
Houston, TX	**5**	1200	0	0	0	700	0	0	0	0					
Lincoln, NE	**6**	0	0	0	0	0	0	500	0						
Champaign, IL	**7**	0	0	0	0	0	500	0							
Pittsburgh, PA	**8**	0	200	200	0	500	0								
Atlanta, GA	**9**	0	0	0	0	0									
Ann Arbor, MI	**10**	0	500	400	0										
Ithaca, NY	**11**	200	0	0											
Princeton, NJ	**12**	100	0												
College Pk, MD	**13**	0													

A.2 European network

The European network topology [91] has 19 nodes and 39 physical links, as shown in Fig. 72. Each adjacent node pair is connected through a bi-directional physical link that consists of two fibers. Table 26 shows the link length in kilometers. The average node degree is 4.1 and the network diameter is 4.

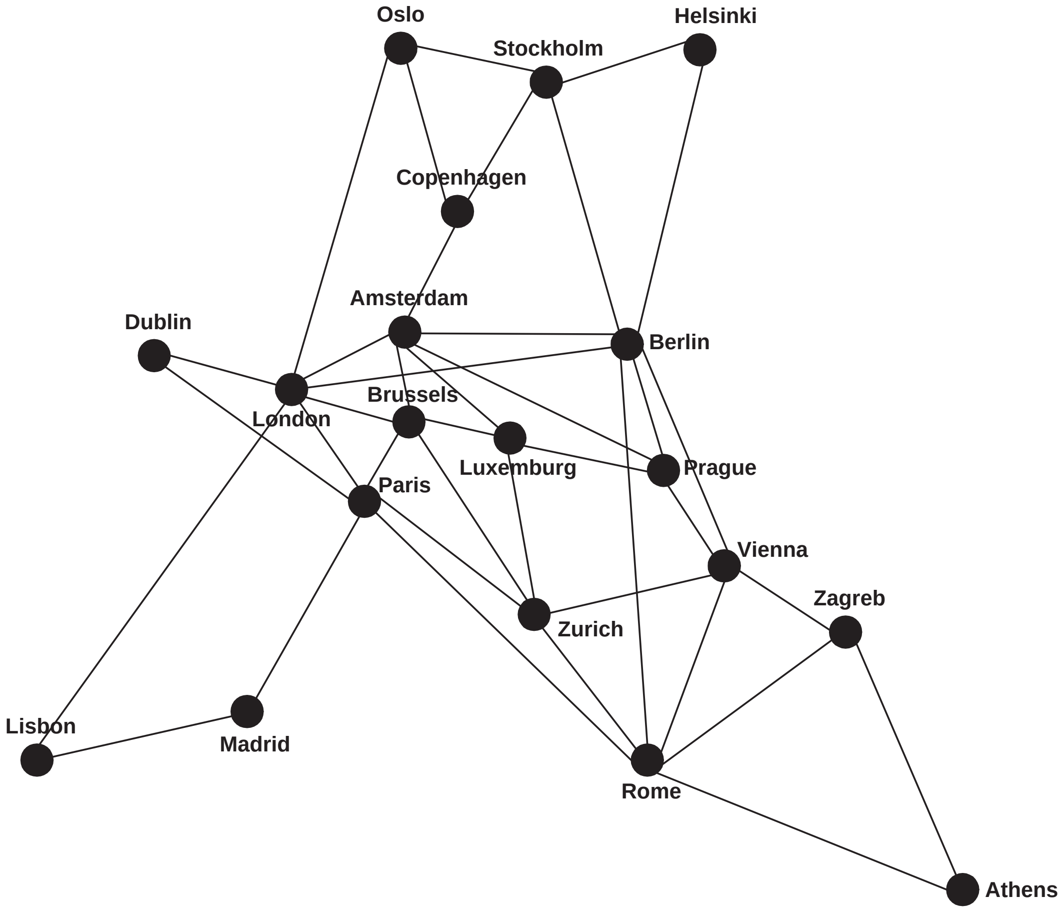


Figure 72: European network.

Table 26: European network: distance matrix (km).

Name		18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Luxembourg	**0**	0	0	0	0	0	560	0	0	0	0	160	0	0	0	0	0	340	220	0
Brussels	**1**	0	0	0	0	0	0	0	0	0	0	470	280	0	0	0	340	250	0	
Amsterdam	**2**	0	600	0	0	600	750	0	0	0	0	0	0	0	0	0	375	0		
London	**3**	1190	0	0	0	1000	0	0	0	0	0	0	500	0	1600	500	0			
Dublin	**4**	0	0	0	0	0	0	0	0	0	0	0	690	0	0	0				
Lisbon	**5**	0	0	0	0	0	0	0	0	0	0	0	0	500	0					
Madrid	**6**	0	0	0	0	0	0	0	0	0	0	0	1070	0						
Paris	**7**	0	0	0	0	0	0	0	0	0	660	500	0							
Zurich	**8**	0	0	0	0	0	530	0	0	0	220	0								
Rome	**9**	0	0	0	0	690	660	0	530	1500	0									
Athens	**10**	0	0	0	0	0	0	0	1100	0										
Zagreb	**11**	0	0	0	0	0	0	250	0											
Vienna	**12**	0	0	0	0	500	280	0												
Prague	**13**	0	0	0	0	280	0													
Berlin	**14**	0	0	875	1650	0														
Helsinki	**15**	0	0	1300	0															
Stockholm	**16**	560	560	0																
Copenhagen	**17**	500	0																	
Oslo	**18**	0																		

A.3 KL network

The KL network topology [37] has 15 nodes and 28 physical links, as shown in Fig. 73. Each adjacent node pair is connected through a bi-directional physical link that consists of two fibers. The average node degree is 3.73 and the network diameter is 4.

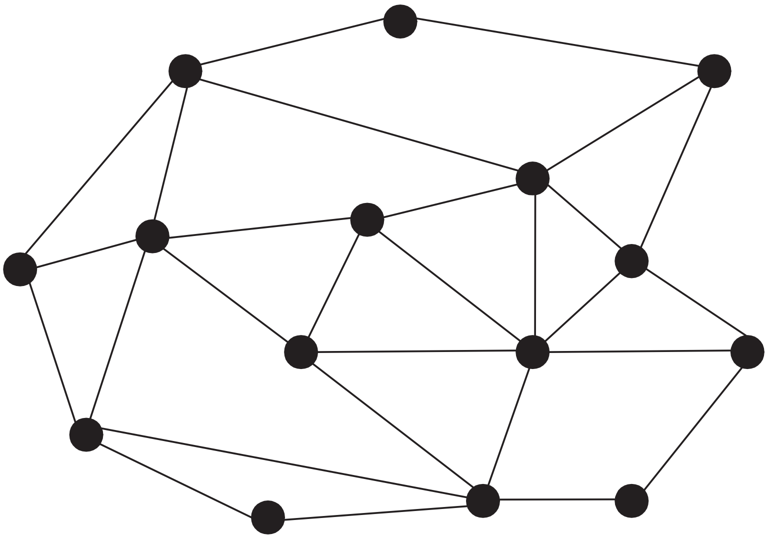


Figure 73: KL network.

APPENDIX B

PUBLICATIONS AND PROJECTS

Throughout this thesis many publications has been published and presented. In this chapter the main publications related with this work are listed.

B.1 Publications

B.1.1 Journals and books

E. Calle, J. L. Marzo, ***A. Urra***. "Protection performance components in MPLS networks. **Computer Communications Journal**, Elsevier, vol. 27, issue 12, pp. 1220-1228, July **2004**.

E.Calle, J.L. Marzo, ***A. Urra***, P.Vila, S.Cots. "Encaminamiento con calidad de servicio y proteccion en redes GMPLS sobre WDM". **Boletin de RedIRIS**, Ed. Red.es / RedIRIS, no. 70-71, pp. 48-51, Dec. 2004/Jan. **2005**.

A. Urra, E. Calle, J. L. Marzo. "Multiagent system for controlling GMPLS network protection". **Artificial Intelligence Research and Development**, Frontiers in Artificial Intelligence and Applications. IOS Press, pp. 256-264, ISBN 1-58603-378-6, **2003**.

B.1.2 International Conferences

A. Urra, E. Calle, J. L. Marzo. "Multi-Layer Network Recovery: Avoiding Traffic Disruptions against Fiber Failures". In Proceedings of the workshop on Evolution toward Next Generation Internet, ICCS 2006, **LNCS**, Reading, UK, May **2006**.

A. Urra, E. Calle, J. L. Marzo. "On-line and dynamic multi-layer routing with protection". In Proceedings of the Workshop on Design of Next Generation Optical Networks: from the Physical up to the Network Level Perspective, **COST 291**, Ghent, Belgium, Feb. **2006**.

A. Urra, E. Calle, J. L. Marzo. "Enhanced multi-layer protection in multi-service GMPLS networks". In Proceedings of **IEEE GLOBECOM**, St. Louis, USA, Nov. **2005**.

A. Urra, E. Calle, J. L. Marzo, "Partial disjoint path for multi-layer protection in GMPLS networks". In Proceedings of Design of Reliable Communication Networks (**DRCN**), Island of Ischia, Italy, Oct. **2005**.

A. Urra, E. Calle, J. L. Marzo. "Enhanced Protection using shared segment backups in a multiservice GMPLS-based networks". In Proceedings of IEEE Symposium on Computers and Communications (**IEEE ISCC**), Cartagena, Spain, Jun. **2005**.

A. Urra, E. Calle, J. L. Marzo, P. Vila. "Minimum interference routing adding shared segment protection in GMPLS-based networks". In Proceedings of International Symposium on Performance Evaluation of Computer and Telecommunication Systems (**SPECTS**), Philadelphia, USA, Jul. **2005**.

A. Urra, E. Calle, J. L. Marzo. "Adding new components to the Knowledge Plane in GMPLS over WDM networks". In Proceedings of IEEE International Workshop on IP Operations & Management (**IEEE IPOM**), Beijing, China, Oct. **2004**.

E. Calle, J.L. Marzo, ***A. Urra***, Ll. Fabrega. "Enhancing fault management performance of two-step QoS routing algorithms in GMPLS". In Proceedings of **IEEE ICC**, Paris, France, Jun. **2004**.

J. L. Marzo, E. Calle, ***A. Urra***, P. Vila. "Failure recovery time minimization in GMPLS-based networks using segment protection". Invited paper. In Proceedings of **JICCSE**, Al-Salt, Jordan, Oct. **2004**.

E. Calle, J. L. Marzo, ***A. Urra***, "Evaluating the probability and the impact of a failure in GMPLS based networks". In Proceedings of Design of Reliable Communication Networks (**DRCN**), Alberta, Canada, **2003**.

E. Calle, J. L. Marzo, ***A. Urra***, P. Vila. "Enhancing MPLS QoS routing algorithms by using the Network Protection Degree paradigm". In Proceedings of **IEEE GLOBECOM**, San Francisco, USA, Nov. **2003**.

E. Calle, J. L. Marzo, ***A. Urra***. "Protection performance components in MPLS networks". In proceedings of Symposium on Performance Evaluation of Computer and Telecommunication Systems (**SPECTS**), Montreal, Canada, July **2003**.

B.1.3 Other publications

A. Urra, J. L. Marzo, M. Sbert, E. Calle. "Estimation of the probability of congestion using Monte Carlo method in OPS networks". In proceedings of IEEE Symposium on Computers and Communications (**IEEE ISCC**), Cartagena, Spain, June **2005**.

D. Merida, R. Fabregat, ***A. Urra***, A. Bueno. "Analysis and regeneration of hypermedia contents through Java and XML tools". In proceedings of **ITCC**, Las Vegas, USA, **2003**.

D. Merida, R. Fabregat, C. Arteaga, ***A. Urra***. "X-SHAAD: an XML implementation for hypermedia systems modeling through SHAAD". In Proceedings of **ICWE**, LCNS 2722, pp. 245-254, Oviedo, Spain, **2003**.

C. I. Pea de Carrillo, R. Fabregat, ***A. Urra***, M. Valles, J. L. Marzo. "Shared whiteboard manager and student notebook for the PLAN-G telematic platform". **Computers and Education, Towards an Interconnected Society**, Kluwer Academic Publishers, **2001**.

REFERENCES

[1] D. Colle, S. Maesschalck, C. Develder, P. V. Heuven, A. Groebbens, J. Cheyns, I. Lievens, M. Pickavet, P. Lagasse, and P. Demeester. Data-centric optical networks and their survivability. *IEEE J. Select. Areas Commun.*, 20(1):6–20, January 2002.

[2] D. Crawford. Fiber optic cable dig-ups: causes and cures. In *Network reliability: a report to the nation - Compendium of technical papers, National Engineering Consortium*, 1993.

[3] J. Y. Wei. Advances in the management and control of optical Internet. *IEEE J. Select. Areas Commun.*, 20(4):768–785, May 2002.

[4] M. Yoo and C. Qiao. Optical burst switching for service differentiation in the next-generation optical internet. *IEEE Commun. Mag.*, 39(2):98–104, February 2001.

[5] N. Ghani, S. Dixit, and T. Wang. Oon ip-over-wdm integration. *IEEE Commun. Mag.*, 38(3):72–84, March 2000.

[6] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y. Rekhter. Generalized Multiprotocol Label Switching: An overview of routing and management enhancements. *IEEE Commun. Mag.*, 39(1):144–150, January 2001.

[7] B. Davie and Y. Rekhter. *MPLS Technology and Applications.* Morgan Kaufmann Publishers, 2000.

[8] S. Yao, B. Mukherjee, and S. Dixit. Advances in photonic packet switching: An overview. *IEEE Commun. Mag.*, 38(2):84–94, February 2000.

[9] E. Mannie. Generalized multi-protocol label switching (GMPLS) architecture. IETF RFC 3945, October 2004.

[10] B. Rajagopalan, J. Luciani, and D. Awduche. IP over optical networks: A framework. IETF RFC 3717, March 2004.

[11] M. Gurusamy and C. S. Murthy. *WDM Technology and Issues in WDM Optical Networks.* Prentice Hall PTR, 2002.

[12] V. Sharma and E. A. Varvarigos. An analysis of limited wavelength translation in regular all-optical WDM networks. *J. Lightwave Technol.*, 18(12):1606–1619, December 2000.

[13] B. Ramamurthy and B. Mukherjee. Wavelength conversion in WDM networking. *IEEE J. Select. Areas Commun.*, 16(7):1061–1073, September 1998.

[14] N. Yamanaka, M. Katayama, K. Shiomoto, E. Oki, and N. Matsuura. Multi-layer traffic engineering in photonic-GMPLS-router networks. In *Proceedings of the IEEE GLOBECOM*, November 2002.

[15] Frank Gonzales, Chia-Hwa Chang, Liang-Wu Chen, and Chih-Kuang Lin. Using multiprotocol label switching (mpls) to improve ip network traffic engineering.

[16] E. Rosen, A. Viswanathan, and R. Callon. Multiprotocol label switching architecture. IETF RFC 3031, January 2001.

[17] E. Rosen, D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li, and A. Conta. MPLS label stack encoding. IETF RFC 3032, January 2001.

[18] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas. Ldp specification. IETF RFC 3036, January 2001.

[19] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, and A. Malis. Constraint-based LSP setup using LDP. IETF RFC 3212, January 2002.

[20] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE: Extensions to RSVP for LSP tunnels. IETF RFC 3209, December 2001.

[21] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin. Resource reservation protocol (RSVP) – version 1 functional specification. IETF RFC 2205, September 1997.

[22] L. Berger. Generalized multi-protocol label switching (GMPLS) signaling functional description. IETF RFC 3471, January 2003.

[23] J. Lang. Link management protocol (lmp). IETF RFC 4204, October 2005.

[24] G. Apostolopoulos, D. Williams, S. Kamat, R. Guerin, A. Orda, and T. Przygienda. QoS routing mechanisms and OSPF extensions. IETF RFC 2676, August 1999.

[25] Q. Zheng and G. Mohan. Protection approaches for dynamic traffic in IP/MPLS over WDM networks. *IEEE Commun. Mag.*, 41(5):S24–S29, May 2003.

[26] W. D. Grover. *Mesh-based survivable networks: Options and strategies for optical, MPLS, SONET, and ATM networking.* Prentice Hall PTR, 2004.

[27] ITU800. Terms and definitions related to quality of service and network performance including dependability. ITU-T E.800.

[28] R. Bhandari. *Survivable networks: Algorithms for Diverse Routing.* Kluwer Academic Publishers, 1999.

[29] J.P. Vasseur, M. Pickavet, and P. Demeester. *Network recovery: Protection and restoration of optical, SONET-SDH, IP and MPLS.* Morgan Kaufmann Publishers, 2004.

[30] B. Rajagopalan, D. Pendarakis, D. Saha, R. S. Ramamoorthy, and K. Bala. IP over optical networks: Architectural aspects. *IEEE Commun. Mag.*, 38(9):94–102, September 2000.

[31] V. Sharma and F. Hellstrand. Framework for Multi-Protocol Label Switching (MPLS)-based recovery. IETF RFC 3469, February 2003.

[32] E. Calle, J.L. Marzo, A. Urra, and L. Fabrega. Enhancing fault management performance of two-step QoS routing algorithms in GMPLS. In *Proceedings of the IEEE ICC*, June 2004.

[33] E. Calle, J. L. Marzo, and A. Urra. Protection performance components in MPLS networks. *Computer Commun. Journal*, 27:1220–1228, July 2004.

[34] D. Xu, Y. Xiong, and C. Qiao. Novel algorithms for shared segment protection. *IEEE J. Select. Areas Commun.*, 21(8):1320–1331, October 2003.

[35] D.A. Schupke, C.G. Gruber, and A. Autenrieth. Optimal configuration of p-cycles in WDM networks. In *Proceedings of the IEEE ICC*, 2002.

[36] W. Lai and D. McDysan. Network hierarchy and multilayer survivability. IETF RFC 3386, November 2002.

[37] M. Kodialam and T. V. Lakshman. Restorable dynamic quality of service routing. *IEEE Commun. Mag.*, 40(6):72–81, June 2002.

[38] M. Kurant and P. Thiran. On survivable routing of mesh topologies in ip-over-wdm networks. In *Proceedings of the IEEE Infocom*, 2005.

[39] L. Hundessa and J. Domingo-Pascual. Reliable and fast rerouting mechanisms for a protected label switched path. In *Proceedings of the IEEE Globecom*, 2002.

[40] C. Huang, V. Sharma, K. Owens, and S. Makam. Building reliable MPLS networks using a path protection mechanism. *IEEE Commun. Mag.*, 40(3):156–162, March 2002.

[41] J. Moy. OSPF version 2. IETF RFC 2328, April 1998.

[42] R. Rabbat and C. F. Su. Fault notification and service recovery in WDM networks. In *White paper*, 2003.

[43] G. Li and J. Yates and. Experiments in fast restoration using gmpls in optical/electronic mesh networks. *OFC*, 4:PD34–1–PD34–3, March 2001.

[44] E. Calle, J.L. Marzo, A. Urra, and P. Vila. Enhancing MPLS QoS routing algorithms by using the network protection degree paradigm. In *Proceedings of the IEEE Globecom*, November 2003.

[45] E. Calle. *Enhanced fault recovery methods for protected traffic services in GMPLS networks*. PhD thesis, University of Girona, February 2004.

[46] M. Tacca, A. Fumagalli, A. Paradisi, F. Unghvary, K. Gadhiraju, S. Lakshmanan, A. de Campos Sachs S. M. Rossi, and D. S. Shah. Differentiated reliability in optical networks: Theoretical and practical results. *J. Lightwave Technol.*, 21(11):2576–2586, November 2003.

[47] T. Kuwabara, Y. Mitsunaga, and H. Koga. Calculation method of failure probabilities of optical fiber. *J. Lightwave Technol.*, 11(7):1132–1138, July 1993.

[48] J. L. Marzo, E. Calle, C. Scoglio, and T. Anjali. QoS on-line routing and MPLS multilevel protection: a survey. *IEEE Commun. Mag.*, 41(10):126–132, October 2003.

[49] AGH University of Science and Technology. *Design and Engineering of the Next Generation Internet, towards convergent multi-service networks*, andrzej jajszczyk edition, December 2004.

[50] A. Nucci, N. Taft, C. Barakat, and P. Thiran. Controlled use of excess backbone bandwidth for providing new services in IP-over-WDM networks. *IEEE J. Select. Areas Commun.*, 22(9):1692–1707, November 2004.

[51] A. Autenrieth and A. Kirstdter. Engineering end-to-end IP resilience using resilience-differentiated QoS. *IEEE Commun. Mag.*, 40(1):50–57, January 2002.

[52] Ch. V. Saradhi, M. Gurusamy, and L. Zhou. Differentiated QoS for survivable WDM optical networks. *IEEE Commun. Mag.*, 42(5):S8–S14, May 2004.

[53] D. Rossier-Ramuz, D. Rodellar, and R. Scheurer. Dynamic protection set-up in optical VPN using mobile agent ecosystem. In *Proceedings of the DRCN*, October 2001.

[54] F. Le Facheur and W. Lay. Requirements for support of diff-serv-aware MPLS traffic engineering. IETF RFC 3564, July 2003.

[55] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. An architecture for differentiated services. IETF RFC 2475, December 1998.

[56] M. Kodialam and T. V. Lakshman. Minimum interference routing with applications to MPLS traffic engineering. In *Proceedings of the IEEE Infocom*, 2000.

[57] E. W. Dijkstra. A note on two problems in connexion with networks. In *Numer. Math.*, volume 1, pages 269–271, 1959.

[58] R. Guerin, A. Orda, and D. Williams. QoS routing mechanisms and OSPF extensions. In *Proceedings of the IEEE Globecom*, 1997.

[59] M. Kodialam and T. V. Lakshma. Integrated dynamic IP and wavelength routing in IP over WDM networks. In *Proceedings of the IEEE Infocom*, 2001.

[60] I. Iliadis and D. Bauer. A new class of on-line minimum-interference routing algorithms. In *Proceedings of Networking Conference, LNCS 2345*, 2002.

[61] Bin Wang, Xu Su, and C. L. P. Chen. A new bandwidth guaranteed routing algorithm for MPLS traffic engineering. In *Proceedings of the IEEE ICC*, 2002.

[62] F. Sun and M. Shayman. Minimum interference algorithm for integrated topology control and routing in wireless optical backbone networks. In *Proceedings of the IEEE ICC*, 2004.

[63] G. B. Figueiredo, N. L. S. Fonseca, and J. A. S. Moneiro. A minimum interference routing algorithm. In *Proceedings of the IEEE ICC*, 2004.

[64] D. Dunn, W. Grover, and M. MacGregor. Comparison of k-shortest paths and maximum flow routing for network facility restoration. *IEEE J. Select. Areas Commun.*, 2(1):88–89, January 1994.

[65] J.W. Suurballe and R.E. Tarjan. A quick method for finding shortest pairs of disjoint paths. *Networks*, 14:325–336, 1984.

[66] K. Kar, M. Kodialam, and T. V. Lakshman. Routing restorable bandwidth guaranteed connections using maximum 2-route flows. In *Proceedings of the IEEE Infocom*, 2002.

[67] P.H. Ho, J. Tapolcai, and T. Cinkler. Segment shared protection in mesh communications networks with bandwidth guaranteed tunnels. *IEEE/ACM Trans. Networking*, 12(6):1105–1118, December 2004.

[68] G. Li, D. Wang, C. Kalmanek, and R. Doverspike. Efficient distributed path selection for shared restoration connections. In *Proceedings of the IEEE Infocom*, 2002.

[69] Murali Kodialam and T. V. Lakshman. Dynamic routing of bandwidth guaranteed tunnels with restoration. In *Proceedings of the IEEE Infocom*, 2000.

[70] L. Li, M. Buddhikot, C. Chekuri, and K. Guo. Routing bandwidth guaranteed paths with local restoration in label switched networks. In *Proceedings of the IEEE ICNP*, 2002.

[71] S. Ramanathan. Multicast tree generation in networks with asymmetric links. *IEEE/ACM Trans. Networking*, 4(4):558–568, August 1996.

[72] P.H. Ho, J. Tapolcai, and H. T. Mouftah. On achieving optimal survivable routing for shared protection in survivable next-generation internet. *IEEE Trans. Rel.*, 53(2):216–225, June 2004.

[73] J. Li and K. L. Yeung. A two-step approach to restorable dynamic QoS routing. In *Proceedings of the IEEE ICC*, June 2004.

[74] A. Urra, E. Calle, and J.L. Marzo. Partial disjoint path for multi-layer protection in GMPLS networks. In *Proceedings of the DRCN*, October 2005.

[75] P. Sebos, J. Yates, G. Hjalmtysson, and A. Greenberg. Routing bandwidth guaranteed paths with local restoration in label switched networks. In *Proceedings of the OFC*, 2001.

[76] A. Urra, E. Calle, and J.L. Marzo. Enhanced multi-layer protection in multi-service GMPLS networks. In *Proceedings of the IEEE Globecom*, November 2005.

[77] K. Sato, N. Yamanaka, Y. Takigawa, M. Koga, S. Okamoto, K. Shiomoto, E. Oki, and W. Imajuku. GMPLS-based photonic multi-layer router (hikari router) architecture: An overview of traffic engineering and signaling technology. *IEEE Commun. Mag.*, 40(3):96–101, March 2002.

[78] E. Oki, K. Shiomoto, M. Katayama, W. Imajuku, N. Yamanaka, and Y. Takigawa. Performance evaluation of dynamic multi-layer routing schemes in optical IP networks. *IEICE Trans. Commun.*, E87-B(6):1577–1583, June 2004.

[79] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku, and Y. Takigawa. Dynamic multilayer routing schemes in GMPLS-based IP+Optical networks. *IEEE Commun. Mag.*, 43(1):108–114, January 2005.

[80] K. Zhu, H. Sang, and B. Mukherjee. A comprehensive study on next-generation optical grooming switches. *IEEE J. Select. Areas Commun.*, 21(7):1173–1186, September 2003.

[81] A. Urra, E. Calle, and J.L. Marzo. Multi-layer network recovery: Avoiding traffic disruptions against fiber failures. In *Proceedings of the ICCS, LNCS*, May 2006.

[82] D. Xu, Y. Xiong, C. Qiao, and G. Li. Failure protection in layered networks with shared risk link groups. *IEEE Network*, 18(3):36–41, May 2004.

[83] S. Kim and S. S. Lumetta. Addressing node failures in all-optical networks. *Jounal of Optical Networking*, 1(4):154–163, April 2002.

[84] K. P. Gummadi, M. J. Pradeep, and C. S. Ramamurthy. An efficient primary-segmented backup scheme for dependable real-time communication in multihop networks. *IEEE/ACM Transactions on Networking*, 11(1):81–94, February 2003.

[85] J. Zhang and B. Mukherjee. A review of fault management in WDM mesh networks: Basic concepts and research challenges. *IEEE Networks Magazine*.

[86] J. Doucette, M. Clouqueur, and W. D. Grover. On the availability and capacity requirements of shared backup path-protected mesh networks. *Optical Networks Magazine*, 20(1):29–44, December 2003.

[87] D. Schupke and R. Prinz. Performance of path protection and rerouting for WDM networks subject to dual failures. In *Optical Fiber Communications Conference*, March 2003.

[88] K. Shiomoto, D. Papadimitriou, J.L. Roux, M. Vigoureux, and D. Brungard. Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN). In *Internet draft, draft-ietf-ccamp-gmpls-mln-reqs-00.txt*, January 2006.

[89] J.L. Roux, D. Brungard, E. Oki, D. Papadimitriou, K. Shiomoto, and M. Vigoureux. Evaluation of existing GMPLS protocols against multi layer and multi region networks (MRN/MLN). In *Internet draft, draft-ietf-ccamp-gmpls-mln-eval-00.txt*, January 2006.

[90] J. Li and K. L. Yeung. Efficient path protection using bi-directional WDM transmission technology. In *Proceedings of the IEEE Globecom*, November 2005.

[91] A. Fumagalli, I. Cerutti, M. Tacca, F. Masetti, R. Jagannathan, and S. Alagar. Survivable networks based on optimal routing and wdm self-healing rings. In *Proceedings of the IEEE Infocom*, 1999.