BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
DEPARTMENT OF TELECOMMUNICATIONS

Assessing and Guaranteeing Availability in Networks with Multiple Failures

Ph.D. Thesis

by

Zsolt Pándi

Advisors:

Dr. Tien Van Do, BUTE
Dr. Andrea Fumagalli, UTD
Dr. Marco Tacca, UTD

Budapest, Hungary

April 24, 2006

The reviews of the dissertation and the report of the thesis discussion are available at the Dean's Office of the Faculty of Electrical Engineering and Informatics of the Budapest University of Technology and Economics.

Az értekezés bírálatai és a védésen készült jegyzőkönyv elérhető a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karának Dékáni Hivatalában.

Contents

List of Tables

List of Figures

List of Abbreviations

ASON	Automatically Switched Optical Network
BSS	Backup Sharing Set problem
DiR	Differentiated Reliability
DPM	Disjoint Path-pair Matix
DPP	Dedicated Path Protection
DWDM	Dense Wavelength Division Multiplexing
eDiR	Extended Differentiated Reliability
FIB	Forwarding Information Base
FIT	Failure In Time
GMPLS	Generalized Multi-Protocol Label Switching
ILP	Integer Linear Programming
IS	Maximum Independent Set problem
IS-IS	Intermediate System to Intermediate System
ITU	International Telecommunication Union
LSA	Link State Advertisement
ME(O)MS	Micro Electro-(Opto-)Mechanical Systems
MIS	Maximal Independent Set problem
MPLS	Multi-Protocol Label Switching
MTBF	Mean Time Between Failures
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
MUT	Mean Uptime
OSPF	Open Shortest Path First
OXC	Optical Cross-Connect
PGF	Probability-Generating Function
QoP	Quality of Protection
QoS	Quality of Service
RSVP	ReSource reserVation Protocol
RWA	Routing and Wavelength Assignment

S(B)PP	Shared (Backup) Path Protection
ShUT	Sharing Unavailability Threshold
SLA	Service Level Agreement
SRLG	Shared Risk Link Group
SRRG	Shared Risk Resource Group
TE	Traffic Engineering
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing

Summary of Notations

q_x is the probability that component x is failed.

$p_x = 1 - q_x$ is the probability that component x is operating.

$q_n^{(A)}$ is the probability that exactly n components are failed in set A.

$q_{n+}^{(A)} = 1 - \sum_{i=0}^{n-1} q_i^{(A)}$ is the probability that at least n components are failed in set A.

$G(V, E)$ is a graph composed of the set of vertices, or nodes, V and the set of edges $E = \{(v_1, v_2) | v_1, v_2 \in V\}$. Note that the edges of graph G may be either directed or undirected depending on the definition used in the context.

$d = (n_{\mathbf{s}}^{(d)}, n_{\mathbf{d}}^{(d)}, t_{\mathbf{a}}^{(d)}, t_{\mathbf{h}}^{(d)}, r^{(d)})$ is a call request, also referred to as a demand, which needs a connection to be set up between source node $n_{\mathrm{s}}^{(d)}$ and destination node $n_{\mathrm{d}}^{(d)}$. The time of arrival of the demand is $t_{\mathrm{a}}^{(d)}$ and the holding time of the connection is $t_{\mathrm{h}}^{(d)}$. The probability that any failure in the network interrupts the connection should be no higher than $r^{(d)}$.

$L_{\mathbf{w}}^{(d)}$ is the set of links used by the working lightpath assigned to demand d.

$L_{\mathbf{b}}^{(d)}$ is the set of links used by the backup lightpath assigned to demand d.

$L_{\mathbf{p}}^{(d)} \subset L_{\mathbf{w}}^{(d)}$ is the set of links used by the working lightpath of demand d that are protected by the backup lightpath of demand d.

$L_{\mathbf{u}}^{(d)} = L_{\mathbf{w}}^{(d)} \setminus L_{\mathbf{p}}^{(d)}$ is the set of links used by the working lightpath of demand d that are not protected by the backup lightpath of demand d.

$N_{\mathbf{w}}^{(d)}$ is the set of nodes used by the working lightpath assigned to demand d.

$N_{\mathbf{b}}^{(d)}$ is the set of intermediary nodes used by the backup lightpath assigned to demand d. That is, $n_{\mathrm{s}}^{(d)} \notin N_{\mathrm{b}}^{(d)}$ and $n_{\mathrm{d}}^{(d)} \notin N_{\mathrm{b}}^{(d)}$.

$N_{\mathbf{p}}^{(d)} \subsetneq N_{\mathbf{w}}^{(d)}$ is the set of nodes used by the working lightpath of demand d that are protected by the backup lightpath of demand d. Note that $n_{\mathrm{s}}^{(d)} \notin N_{\mathrm{p}}^{(d)}$ and $n_{\mathrm{d}}^{(d)} \notin N_{\mathrm{p}}^{(d)}$.

$N_{\mathbf{u}}^{(d)} = N_{\mathbf{w}}^{(d)} \setminus N_{\mathbf{p}}^{(d)}$ is the set of links used by the working lightpath of demand d that are not protected by the backup lightpath of demand d. Note that $n_{\mathrm{s}}^{(d)} \in N_{\mathrm{u}}^{(d)}$ and $n_{\mathrm{d}}^{(d)} \in N_{\mathrm{u}}^{(d)}$.

$\mathcal{W}^{(d)} = L_{\mathbf{w}}^{(d)} \cup N_{\mathbf{w}}^{(d)}$ is the set of links and nodes used by the working lightpath of demand d.

$\mathcal{B}^{(d)} = L_{\mathbf{b}}^{(d)} \cup N_{\mathbf{b}}^{(d)}$ is the set of links and nodes used by the backup lightpath of demand d.

$\mathcal{P}^{(d)} = L_{\mathbf{p}}^{(d)} \cup N_{\mathbf{p}}^{(d)}$ is the set of links and nodes of the working lightpath of demand d that are protected by the backup lightpath of demand d.

$\mathcal{U}^{(d)} = L_{\mathbf{u}}^{(d)} \cup N_{\mathbf{u}}^{(d)}$ is the set of links and nodes of the working lightpath of demand d that are not protected by the backup lightpath of demand d.

$b = (e, w)$ is a backup resource in a WDM network, where capacity is of wavelength granularity, i.e., b is wavelength channel w on link e.

$D^{(b)}(t)$ is the set of active demands whose backup lightpath uses backup resource b at time t. Note that $D^{(n)}(t),\ n \in V$ is interpreted as the set of active demands whose backup lightpath traverses node n.

$S_{\mathbf{w}}^{(b)}(t)$ is the set of links used by the working lightpaths assigned to demands in $D^{(b)}(t)$.

$D^{(b)}(t, d) = D^{(b)}(t) \setminus \{d\}$ is the set of active demands other than d whose backup lightpath uses backup resource b at time t.

$S_{\mathbf{w}}^{(b)}(t, d)$ is the set of links used by the working lightpaths assigned to demands in $D^{(b)}(t, d)$.

$u^{(b)}(t, d)$ is the sharing unavailability of backup resource b as seen by demand d at time t.

λ is the call request arrival intensity.

$\hat{x}$ is an upper bound on quantity x.

$\check{x}$ is a lower bound on quantity x.

$\bar{x}$ is the average value of quantity x.

Preface

Service outages in telecommunications networks represent higher risk than ever before because of the unprecedented extent of traffic concentration on high capacity links and because of the unacceptable potential economic consequences. Moreover, an ever-growing part of revenues of network operators depends on the compliance with service level agreements signed with customers. Hence the improvement of the survivability of telecommunications networks has received increased attention recently.

As a result of significant amount of research in the field today there is a plethora of various resilience schemes that attempt to mitigate the impact of failures. The majority of works on resilience schemes quantifies failure resilience of connections as the highest number of components that may fail at the same time so that the connection is restored by means of diversion to backup paths. However, this quantification is hard to translate to guarantees on downtime, i.e., availability guarantees. These guarantees are important because service level agreements are based on them, as they are easy to interpret by customers. In addition, failure multiplicity based categorization does not provide fine control on the trade-off between resource consumption and the provided level of availability.

The importance of network availability assessment and methods for providing connections with guaranteed availability has thus increased, as well. Telecommunications networks comprise an increasing number of components, and it is easy to see that, especially on the continental scale, the probability that multiple components are failed is significant. Therefore, it is important to address multiple failure scenarios in availability computations.

However, reliability theory has shown long ago that problems related to connection availability computation are hard to solve in the general case. One may thus resort to examining special cases when it is possible to compute exact values, or try to devise conservative estimation methods that may be used for bounding connection availability.

This thesis deals with availability computations that consider multiple failure scenarios, as well. An efficient algorithm is presented first that computes exact failure stratum probabilities in order to support network analysis methods that are based on the principle of state space sampling.

Two on-line connection provisioning methods are proposed next, the extended differentiated reliability method and the sharing unavailability threshold based method,

which enter an area hardly covered by works in the literature so far. Both provisioning methods use *conservative* connection availability estimation techniques that are suitable for application in a *dynamic* traffic scenario.

The extended differentiated reliability method is then used to assess the significance of node equipment failures in all-optical networks, the probability of which is often considered negligible in the literature. Some problems to be solved when applying the proposed methods are proven to be NP-Complete, which is used to justify heuristic solutions. The two provisioning methods are also compared with each other in terms of feasible availability guarantees and blocking probability.

An important limitation to the applicability of the presented methods, as well as to that of other methods in the literature that assume complete knowledge of the network state, is the inaccuracy of network state information. Link state routing protocols are used in order to maintain a link state database, which is a representation of the network as seen by a particular node. These databases reflect changes in the network with a certain delay, and there is a non-zero probability that inaccurate information has to be used at nodes for connection admission decisions. A theoretical model is proposed to estimate this probability, and the performance of a real-world link state routing protocol, namely OSPF extended with TE capabilities is assessed with the model.

Acknowledgements

When one tries to write the acknowledgements section of one's own PhD thesis, it often sounds like an endless list of cheesy credits instead of a true expression of gratefulness. Indeed, I found myself in a difficult situation when I started to think about what to put in here trying to avoid that fate. My notorious keenness on attempting to find original ideas let me come up with the following. Feel free to skip the page if you find it inappropriate.

Tien, Andrea, Marco, Lena, Paolo, Laci, Tivadar, Zoli, Gábor, Csaba, Tihamér, Reni, Tibor, Mom and Dad. All of these people probably know how invaluable their contribution was to my success. And so do I. No words of gratitude are sufficient to match that. I am proud to have them around me.

Chapter 1

Introduction

Telecommunications networks, similarly to other fundamental products of engineering work, remain largely invisible to most people if operated appropriately. Their existence and the extent to which our everyday life relies on them only becomes apparent when an outage disrupts the basic services that we take for granted, such as telephone, television or the Internet. It is, therefore, crucial to examine how the survivability of telecommunications networks and services can be improved and guaranteed.

In a reliability value chain reliable[1] services are operated in a reliable network, which comprises reliable equipment. This value chain emphasizes the hierarchy that determines service reliability perceived by users or customers. While increasing equipment reliability is equally important in increasing perceived service reliability, in what follows emphasis is put on the network level.

Even as short as two decades ago the majority of telecommunications traffic comprised telephone conversations. The primary concern with telephone service was service availability[2], which was limited by two factors: reliability of the telephone network and the relation between offered traffic and network capacity. While both of them may be a cause of service outage, at the network level they are related to two distinct phenomena: network failures and call blocking, respectively. To minimize the probability of service outage caused by either one of these phenomena while maximizing operator revenues has remained a fundamental objective of telecommunications engineering ever since.

Nowadays, the continuously increasing volume of traffic transported by telecommunications networks is concentrated to such a high extent that failures potentially affect several users. A variety of already existing and foreseen applications, such as medical imaging or grid computing [9, 45, 83], or even traffic grooming operated in an overlay model [85] may impose strict availability requirements on the network for short-lived connections. Currently, the emerging virtual private networking (VPN) technology shows that already there is a growing financial interest in services of increased, and most importantly, of guaranteed reliability.

Service contracts between service providers and users specify the parameters related

[1]The word *reliable* is being used here in its most general meaning: something is reliable if it is suitable or fit to be relied on.

[2]The word *availability* is being used here in its most general sense: something is available if it is present or ready for immediate use.

to quality of service (QoS) and their acceptable ranges. Such quality definitions are called service level agreements (SLA). They serve as the basis of revenues for service providers, because the amount of money customers are charged is in general affected by both the level of guarantees in the SLA and the violations of the SLA. The traffic of all of the service classes is usually served by the same network. It is, therefore, of primary importance to address the question of network reliability on a user-by-user basis as well, instead of applying global reliability performance measures exclusively.

The traffic offered by the emerging new services is far less predictable than that of the telephone service. The lack of predictability derives from both randomness of actual data transmission parameters and the uncertainty of demand patterns, i.e., potential drifts in stationary patterns. Two solutions are available to tackle this problem. The first one, which stems from telephony, is capacity overbuild, while the second one is increased network intelligence and flexibility. Given the high capital expenses required in telecommunications the second one is more attractive to network operators, since it changes the cost structure so that some of the initial capital expenses are converted to future operational expenses. Moreover, network intelligence is also the key to future-proof networks.

Recent advances in lightwave technology, e.g. the increasing cost-effectiveness of (dense) wavelength division multiplexing (WDM and DWDM), and in photonic switching, such as the already off-the-shelf Micro Electro-(Opto-)Mechanical Systems (MEOMS) technology, enable networks that provide high bitrate connectivity with sufficient flexibility. Due to the inherent properties of the technology the implementations available today and in the foreseeable near future operate in the connection-oriented discipline. Thus the same phenomena as in case of telephone networks, namely network failures and call blocking have to be dealt with, however, by means of fundamentally different methods.

Standards demonstrate the interest and support of industry in the field. Two prominent examples of this are automatically switched optical networks (ASON) [103] and Generalized Multi-Protocol Label Switching (GMPLS) [58]. The concept of ASON is in fact an architecture based on optical components, that has the property of flexibility. On the other hand, GMPLS is envisioned as the control plane for future transport networks. While these standards give an adequate description of some of the techincal details, they do not address how networks should be operated in a resource-efficient manner.

The literature has proposed several solutions to provide survivability in connection-oriented optical networks. These solutions adapt the same basic idea: a certain amount of resources is dedicated to each connection to establish connectivity, and some backup resources are also relied on either at time of a network failure or earlier. The solutions thus range from fully dynamic re-routing of connections to protection switching according to the level of pre-provisioning of backup resources. Another classification criterion is

the length of the path of the connection intended to be covered individually by backup resources. According to this one, it is possible to distinguish path protection/restoration, sub-path or segment protection/restoration and span protection/restoration [33, 81]. With respect to the latter categorization one has to mention that names of certain schemes may serve as an umbrella term for methods that pertain to multiple categories, one example of which is the remarkable p-cycle approach, proposed first in [32].

The ultimate goal of network operators is to maximize revenues by means of maximizing the number of revenue generating contracts. QoS guarantees require that network resources be allocated to services. As network capacity is finite, it is also the interest of operators to decrease the amount resources dedicated to each service. Maintaining service level guarantees and decreasing resource consumption are thus contradictory requirements.

It is also the interest of network operators to decrease business risk by employing methods that are well-studied, well-understood, bulletproof and have sufficient standardisation and vendor support. In other words, network operators tend to avoid using highly complex, proprietary ways in spite of their willingness to squeeze out more revenues from their infrastructure. As a consequence, the principle of Occam's razor should be followed, and avoidable complexity should be avoided.

1.1 Thesis motivation

Computing availability in networks is indeed difficult in most cases. Since networks are most often modeled by graphs, availability studies in networks may be easily translated to graph theory problems. Reliability theory often uses the same instrument to model systems, and so its abstract results apply to the case of networks, as well. Reliability theory has answers to some relevant questions [93]; however, the exact solution of most of the problems of practical interest remains difficult. One of them is computing connection availability.

One may pursue different approaches to approximate or to bound the results of exact computations. If the network is modeled as a sytem of finite states then state-space sampling methods may be applied, such as the well-known Monte Carlo method [20], stratified sampling [10] or adaptive approximation [53]. These capitalize on randomly selecting a subset of the total failure state space of the network and on evaluating network performance (or connection availability) only in the selected states. Stratified sampling is driven by the cumulated probability of some special subsets in the total network state space called failure strata. A failure stratum is usually defined as the set of failure states of the network where the number of failed components is the same. The total probability of the failure states that belong to the same failure stratum is called failure stratum probability. These probabilities may also be useful for selecting the training set

of adaptive approximation. [10] proposes an algorithm for obtaining these probabilities; however, it is of fairly high complexity.

Another approach to pursue is to derive bounds on the value in question, i.e., connection availability. The principle of inclusion-exclusion is often used for such purposes, that is, when the value to be estimated is a probability [21]. However, one issue with the application of the inclusion-exclusion principle and other general bounding methods may be the need to specify or compute joint probabilities of events. If backup resources are shared in telecommunications networks then the events to deal with are often not independent, and due to the lack of field data on the probabilities of such events and the difficulty to estimate these probabilities the practical applicability of general methods is limited. Nevertheless, by means of taking advantage of the knowledge on the structure of the problem one may devise specific methods for obtaining bounds.

It is an area hardly covered by literature so far. Publications tend to categorize connection availability according to the number of simultaneous failures that a connection survives. The reason for this is twofold. Firstly, failure multiplicity is a picturesque description of the failure state of a network and thus offers an easy grip on how protection/restoration mechanisms should be operated. Secondly, failures eventually mean failure of certain components to network operators, which is easy to interpret in terms of frequency and required reactions. For example, from the aspect of customers an explicit guarantee of single failure restorability is not very meaningful, whereas an upper bound on the connection downtime is.

A notable exception is the proposal of the differentiated reliability (DiR) principle, that appeared first in [28]. It was later accompanied by the notion of quality of protection (QoP) [30]. The idea raised in [28, 30] is the differentiation of the guaranteed level of connection availability according to individual requirements associated with connections. While [30] introduced a more conceptual model, [28] and subsequent works specifically addressed probabilistic guarantees, yet mostly for single failure scenarios, i.e., probabilistic guarantees remained conditional.

The majority of protection methods proposed in the literature mostly target at routing single failure robust connections, that is, they only guarantee the survival of the connection in case of single component failures. This approach is structurally reasonable and economic with respect to resource usage and, surprisingly, usually provides robustness against a significant portion of failures of higher multilicity, as well. With the growth of networks in capacity, scale and number of components, the probability of failures of higher multiplicity increases significantly. Consequently, applications with higher availability requirements do need probabilistic guarantees that address also multiple failures, therefore, it is necessary to propose methods capable of providing them.

Scalability becomes an important issue with network growth in many aspects. Distributed network intelligence is one answer to this challenge, which, unfortunately, intro-

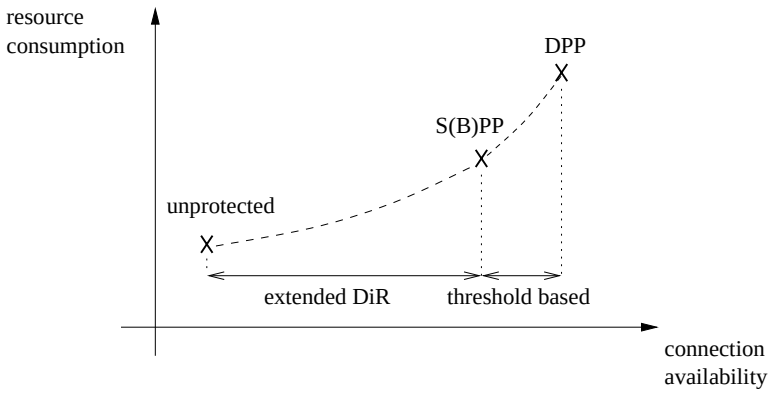


Figure 1.1: Context of provisioning methods proposed in the thesis

duces performance impairments due to the unavoidable delays in network state information dissemination. Up-to-date information is vital for a connection provisioning scheme that guarantees service parameters. Decisions have to be made based on the current knowledge about network configuration and both guarantees and performance are affected by wrong decisions. Standardized link state routing protocols, such as OSPF [66] and IS-IS [68], or more precisely, their respective extensions for traffic engineering (TE) purposes are used for link state information distribution, and it is important to know their performance limitations when QoS guarantees have to be fulfilled.

1.2 Thesis objectives

The general objective of the thesis is to overcome the discussed shortcomings of current work related to (connection) availability computation and connection provisioning with guaranteed availability.

A general network model is considered, in which there are independent, two-state components. Both links and nodes are considered to be failure-prone except where due notice is made in order to simplify the discussion.

Based on this network model the first goal is to propose a computationally efficient algorithm for computing the probabilities of failure strata. Such an algorithm may then be applied to systems of known structure to demonstrate that probabilities of failure strata may be computed efficiently exploting knowledge on system structure.

The main goal of the thesis is to propose methods for connection provisioning with availability guarantees based on shared (backup) path protection (S(B)PP) that are capable of keeping resource usage lower that that of dedicated path protection (DPP). The DiR principle is a suitable candidate for differentiation of connection availability. By means of extending the DiR principle to multiple failures for providing absolute probabilistic guarantees it is possible to scale connection availability from the unprotected

case to the shared backup path protected case with full backup resource sharing.

In order to provide availability guarantees beyond these, additional limitations have to be introduced on the sharing of backup resources. This may be accomplished by introducing a threshold parameter in the connection admission control algorithm. In return for the higher expected complexity this solution has the potential to offer higher availability guarantees (see Figure 1.1).

The importance of node failures is often overlooked or deemed insignificant in the literature. It is therefore an additional goal of the thesis to analyze the impact of node failures on end-to-end connection availability in all-optical networks.

The last goal of the thesis is to address the applicability of the proposed algorithms and methods. Of primary importance is the problem of outdated state information, because most connection provisioning methods rely on complete knowledge of the network state. This assumption, however, is not realistic, and a theoretical model is necessary to estimate the probability that outdated information is used in connection admission decisions. Thus, the performance and limitations of connection provisioning methods that rely on a suitable traffic engineering extension of the OSPF routing protocol for link state database maintenance may be analyzed.

Even though the proposed solutions may be generalized, it is out of the scope of the present work to devise and discuss in detail potential ways to guarantee connection availability with other protection and/or restoration methods. Elaborating solutions that are directly applicable to other networking technologies — including those based on packet-switched operation — also remains out of the scope of the thesis, as well as the assessment of the difference in between *guaranteed* and *actual* connection availability. QoS parameters except for availability and call blocking, such as recovery time, are not covered by this thesis either. Optical networks with wavelength conversion capabilities are neither addressed in this study. Instead, these points are set forth as future research topics.

1.3 Methodology

First, open problems identified by section 1.2 are formulated using mathematical notation. This initial step is inevitable for accurate and unambiguous description of conditions, concepts and relationships.

Models are then constructed using the introduced notations with help of graph theory and probability theory. Elemental techniques and results of algebra and queuing theory are also applied whenever necessary to derive either exact or approximate solutions or to express performance parameters. Some of the encountered problems are proved to be difficult using the results of computation theory, which is used to justify heuristic solutions.

The elaborated and proposed provisioning methods are tested by means of simulations. A simulator is implemented based on the general principles of event driven simulation [27] and parameters of the simulations are chosen to reflect networks of reality. Simulation results are presented with appropriate confidence intervals obtained using statistical methods.

Conclusions are drawn based on either mathematical proofs or simulation results, and limitations to the applicability of the proposed methods are also discussed.

1.4 Thesis structure

This thesis is organized into seven chapters, each one divided into several sections. A brief description of chapters is now presented. After Chapter 2, which serves as a general background to the rest of the thesis, each chapter may be read either individually or sequentially.

Chapter 2 discusses the background of the thesis. It is a review of literature related to reliability modeling, availability computations and network survivability.

Chapter 3 introduces a computationally efficient algorithm for exact computation of failure stratum probabilities in a set of independent two-state components.

Chapter 4 presents a connection provisioning method that extends the DiR principle to multiple failure scenarios and node failures. Using the presented method all-optical networks are analyzed with different node equipment. The NP-Completeness of the arising RWA problem is proved.

Chapter 5 introduces a more complex, sharing unavailability threshold based connection provisioning method to reach availability guarantees beyond those offered by the extended DiR method. The selection of appropriate threshold values is addressed and the advantages and disadvantages of the two connection provisioning methods are compared. The NP-Completeness of determining the maximum number of connections that may share a certain backup resource is proved.

Chapter 6 discusses the applicability of the algorithms and methods presented in the thesis. Most importantly, the concern of decisions based on outdated information in a distributed environment is addressed and an analytical model is proposed for evaluating the probability of such decisions. The performance of OSPF with extensions for traffic engineering is analyzed using the presented model.

Chapter 7 gives a summary of the contributions of the thesis and proposes future research directions related to the presented results.

This thesis ends with two appendices. Appendix A collects topology data of networks referred to throughout the thesis along with their respective references. Appendix B discusses the details of the simulator used for obtaining simulation results.

Chapter 2

Survivability of Telecommunications Networks

Since engineers and researchers working on problems related to survivability sometimes use concepts in an imprecise way — in spite of the existence of a common terminology —, the basic definitions relevant with respect to the discussion in this thesis are reviewed first.

2.1 Definitions related to reliability theory

Definition 1. *The expected length of the time elapsed in the operating state between successive failures of a component or a system is called* mean time between failures (MTBF).

This value is often specified by manufacturers, because it only depends on the inherent structural properties of the component or the system, provided that the operating conditions are met.

Definition 2. *The expected length of the time elapsed in the failed state between the onset of a failure event and the restoration of the operating state (repair) of a component or a system is called* mean time to repair (MTTR).

This value primarily depends on operating practice, but other factors may also have influence on it, such as the time necessary to obtain spare components.

Note that both MTBF and MTTR are mean values; therefore they only give limited information about the distribution of the respective time intervals.

The frequency of failure events is often expressed by engineers using a special unit, which is defined as follows.

Definition 3. 1 FIT (failure in time) = $\frac{1}{\text{MTBF}}$ $[10^{-9}]$h.

The following two concepts are of special importance in reliability engineering.

Definition 4. *The probability that up to time t a component or a system is continuously in the operating state is called* reliability *and is denoted by $r(t)$.*

Reliability has the following properties:

- $r(0) = 1$, that is, at the beginning every component or system is assumed to be operating.
- $\lim_{t\to\infty} r(t) = 0$, i.e., each component or system fails sometime.
- $r(t)$ is monotonously decreasing.

Definition 5. *The probability that at time t a component or a system is currently in the operating state is called* availability *and is denoted by $a(t)$.*

Availability has the following properties:

- $a(0) = 1$, that is, at the beginning every component or system is assumed to be operating.
- $\lim_{t\to\infty} a(t) = A$. If exists, A is the asymptotic availability of a component or a system.

There is a relation in between asymptotic availability and the quantities defined in Definition 1 and in Definition 2, which may be expressed as:

$$A = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}. \tag{2.1}$$

[43] gives standardized definitions for the concepts discussed so far. The definitions used in this thesis agree with those in [43] for repaired items with non-zero time to restoration with the following additional remarks:

- [43] uses the term *instantaneous availability* for $a(t)$.
- MTBF corresponds to mean operating time between failures in [43] and equals mean time to failure (MTTF).
- [43] makes the following distinction with respect to (2.1). If "function-preventing" preventive maintenance actions are not permitted, i.e., the item operates continuously then (2.1) remains unchanged. However, when preventive maintenance is also permitted, then MTBF should be replaced by mean uptime (MUT) in (2.1), because MUT excludes the time when the item is in the operating state, but is not actually functioning due to the maintenance being carried out.

A good example of the difference between reliability and availability is the following. On the one hand, it is extremely hard to construct a system in which $r(t) = 1$ for high values of t, because it essentially requires components of high reliability. On the other hand, it is relatively easy to construct a system with high asymptotic availability even using components of low reliability if one can minimize repair time, according to (2.1).

The rest of the thesis is concerned with availability of components and systems, as defined in Definition 5.

2.2 Reliability modeling and computation of availability

Reliability and availability of different systems has been in the focus of interest for a long time and has given rise to stand-alone fields in theory and engineering: reliability theory and reliability engineering.

The instruments of reliability theory may be used to build models of real systems with a variety of assumptions. A system in general is considered to comprise stateful components. The functionality of the system depends on the state of its components. The general goal is to derive system-level parameters based on knowledge about individual components, system structure and functionality constraints.

Components may either change states independently of each other, or dependence may be assumed, as well. The number of states of individual components may vary from two (*operating* or *failed*) to any arbitrary number.

Several modeling methods have been proposed that are capable of deriving system-wide reliability metrics, and the different methods have different descriptive power. An interesting discussion of this can be found in [57]. Different tools are also available that help in constructing system models and deriving metrics of interest, e.g. [1, 84].

In case of telecommunications networks the solution of these modeling problems often boils down to determining probabilistic quantities derived from graphs. Reliability theory has several related results, two excellent and comprehensive overviews of which may be found in [7, 93].

The practical problem of determining connection availability is matched closest by the K-terminal reliability problem in reliability theory.

Definition 6. *Let a network be represented by a graph $G(V, E)$, where V is the set of nodes and E is the set of independent, two-state communication links between them, each of which is associated with an asymptotic unavailability value. For a subset of nodes $K \subset V$ and a node $s \in K$ the* K-terminal reliability *is the probability that there exist operating paths from s to each node in K* [7].

Of special interest are two derived quantities: 2-terminal reliability (when $|K| = 2$) and all-terminal reliability (when $K = V$). In the general case the computation of K-terminal reliability is NP-hard [7], however, for some restricted classes of graphs, such as series-parallel graphs, polynomial algorithms exist.

The K-terminal reliability problem corresponds to a path existence problem that assumes that availability values are given for all of the links, which fail independently. If backup resources are shared among multiple connections, however, it is difficult to derive exact unavailability values for the shared resources, because activation of the backup resource by other connections also contributes to the unavailability. The evaluation of this contribution alone is complex, because of the interdependence of connections that

share backup resources, which also leads to the loss of independence of link "failures" in this model (c.f. section 2.4.3).

Consequently, it is desirable to use approximation methods for deriving connection availability. Various boundig techniques exist. For example, a general abstract analytical technique is discussed in [21] based on the principle of inclusion-exclusion. Unfortunately, the concept of backup resource sharing is not addressed by general methods in its entirety, because it is specific to certain areas of engineering.

Therefore, instead of general abstract methods often specific methods are used, which exploit the exact knowledge about the structure of the system. Such methods include various state-space sampling techniques and specific bounding methods.

In case of state-space sampling a subset of (relevant) failure states of the system is selected and the value of the reliability metric of interest is bounded using the result of an exhaustive evaluation of the selected failure states and the probabilities of the examined states. Failure state selection may be driven by different methods [10, 20, 53], the choice of which also has an influence on the goodness of the bounds.

Some specific connection availability estimation methods are presented in section 2.4.3, only a few of which are suitable for providing conservative bounds, i.e., lower bounds that always avoid excess. Chapters 4 and 5 propose and analyze such techniques, which are suitable for application in a scenario when the set of connections in the network dynamically changes.

2.3 Data on failures in telecommunications networks

Failures in telecommunications networks may be related to cables and equipment. Statistical data on cable and equipment failures is of extremely high value to both network operators and equipment vendors because of competitive issues and security reasons. As a consequence, it is against the interest of these companies to publish such data, and that is why it is difficult to find reliable references.

In the case of cables, the inherent asymptotic unavailability can be expressed using the empirical formula [47]

$$\frac{hxy}{1+hxy}, \tag{2.2}$$

where x, y and h are the failure rate of a cable span of unit length, the MTTR and the length of the cable span, respectively. Note that the span failure rate is considered to be directly proportional to the length of the cable span, while MTTR is independent of it.

Based on statistical data, $4 * 10^{-6}$ is a reasonable estimate of the inherent asymptotic unavailability of a cable of 1km, which is used throughout this study. Assuming MTTR=12 h this estimation yields a failure rate of 2.92/yr/1000 km. Note that

Network	$q_{1+}^{(E)}$
US	0.123987
EU	0.0955112
Italian	0.0232946
metropolitan	0.00411098

Table 2.1: Probabilities that at least one link is failed in different networks

MTTR of cables is generally assumed to be higher than that of equipment deployed at network nodes. The values used in this study are taken from network planning experience accumulated during the cooperation between Magyar Telekom (formerly Matáv), the Hungarian incumbent telecommunications service provider and the Department of Telecommunications at BUTE.

[100] is the standard reference for cable failure statistics used in the literature. The values published in [100] are based on statistical data collected by Bellcore during the operation of a fiber optic network. Other sources include [2, 41, 56]. The cited works report failure rates of 0.8–2.75/yr/1000 km with MTTR values in the range 8–24 h. Therefore, the value used in this study mathces available data sources.

With regard to equipment failures two primary failure causes may be distinguished: hardware and software. With respect to hardware components of photonic technology one may refer to very few additional sources [42, 61, 99, 108] beside the ones already cited for cable failure data. Out of these sources [108] is used in this study.

It is next to impossible to find data on software related failures. The reason for this is that while in case of hardware components standard procedures exist to determine and/or derive reliability characteristics, e.g. aging tests or prediction models based on statistical data [42, 99], the software industry has been unable to agree on a suitable methodology so far.

[70] concludes that software related failures represent a significant portion of total network failures based on the analysis of trouble ticket data from two research oriented national networks. However, it is only [49] that ventures an educated guess of failure rate of software running in IP routers, which may or may not apply to optical equipment. As a consequence, software related failures are not modeled in the present study, except for section 6.2, where IP routers are assumed to be deployed at network nodes.

2.3.1 Quantitative assessment

In order to assess the importance of multiple failures the following simple experiment is carried out. Only link failures are assumed and the probability that there is at least one failed link is computed in the networks used as reference throughout the present

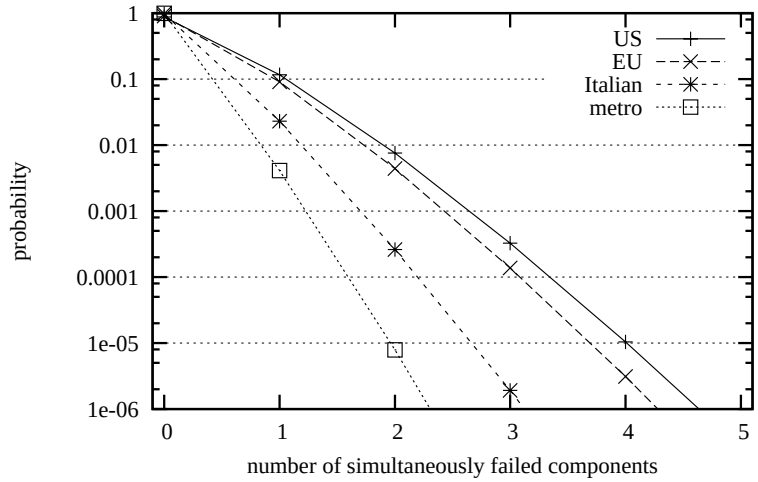


Figure 2.1: Probabilities of failures of different multiplicity in different networks

study. The asymptotic unavailabilities of individual links are determined using (2.2). The results are displayed in Table 2.1.

The probability that the network is in a failure state is surprisingly high (around 0.1) in case of networks of continental scale. It is, therefore, interesting to examine the contribution of failures of different multiplicity to this total value. Figure 2.1 shows the results of a more detailed analysis that uses the same assumptions as before. Note the probabilities are only interpreted at integer values along the horizontal axis on Figure 2.1, and individual points are only connected to make trends visible.

The asymptotic probability that exactly two components are failed is higher than 0.001 in the US and the European networks, and it is above 0.0001 in the Italian national network. The asymptotic probability of failure states with three failed components in the continental topologies is also not negligible. It means that even if only lower values than the often quoted "five nines" of availability arise as a real requirement for connections, the presented results justify the concern about the impact of multiple failures.

2.4 Work related to survivability analysis and guarantees in connection-oriented telecommunications networks

Since telecommunications networks comprise failure-prone components, it is inevitable to address the problem of establishing survivable connections. It is an objective which already existed long ago and numerous attempts have been made to incorporate this objective in network dimensioning. The notion of routing dependable connections in a dynamic traffic scenario appears first in [63].

2.4.1 A note on network dimensioning vs. dynamic traffic scenarios

When the performance of telecommunications networks has to be optimized, two fundamental scenarios are assumed in general.

In a network dimensioning scenario there is an *a priori* knowledge on the set of demands to be served by the network. This information may be the output of some traffic estimation or statistical prediction method, or it may be specified exactly as part of the problem statement. The problem is then to answer the following questions:

- Where should one install cables to establish direct connectivity between network nodes?
- What is the required capacity of these cables?
- How are demands routed in the topology?

while the following constraints apply:

- All demands have to be served and their QoS requirements must be fulfilled.
- Overall network cost should be minimal.

Obviously, the costs of installing cables and equipment, and the availability measures of cables and equipment are given.

In a dynamic traffic scenario, however, there is no *a priori* knowledge on the set of demands to be served by the network. The network topology, on the other hand is given in this scenario. Connection requests arrive in a sequence, and release assigned network resources after the connection holding time expires. The problem is then to answer the following questions at the arrival of each demand:

- Is it possible to serve the demand given the current state of the network?
- How can the current demand be routed in the topology?

while the following constraints apply:

- The demand may only be served if QoS requirements are guaranteed.
- The ratio of rejected demands to arrived demands should be minimal over a long sequence of demands.

Of course, the availability measures of cables and equipment are given.

The problem of survivable connection establishment arises in both scenarios as discussed above. Moreover, it is defined in the two scenarios with similar conditions and constraints. Consequently, resilience schemes proposed for any of the two scenarios are of interest, and are reviewed in what follows.

Guaranteed survivability requires that backup resources are assigned to connections at the time of routing the connection. In case of purely dynamic restoration methods[1] it is extremely difficult to guarantee connection availability, therefore, these methods remain out of the scope of interest of the current study.

The majority of the literature classifies the grade of resilience of connections according to the maximum number of simultaneous component failures that do not disrupt the connection. Hence most proposed methods guarantee single or dual failure restorability.

2.4.2 Guaranteed single failure resilience

First, the schemes proposed in the literature that guarantee single failure restorability are reviewed.

Perhaps the most simple scheme is the one called link protection [81]. When connections are routed with link protection a working path is chosen first that connects the source and destination nodes of the connection request. Then a backup path is sought for each link on the working path between the end nodes of the particular link. This backup path is activated if the link fails. Resources used by backup paths may be either dedicated to the connection or may be shared by multiple connections.

Two distinct backup paths may share resources, i.e., rely on the same wavelength on the same link in the same direction, on the condition that the sets of components to be protected by them are disjoint. Otherwise, the failure of a single component in the intersection of the two sets would lead to a backup resource sharing conflict. In case of link protection these sets comprise single links only, however, this general sharing rule applies to all schemes that guarantee single failure resilience.

Note that it is not always possible to find a "detour" backup path for each link. Link protection, if the necessary backup paths exist, is robust against any single link failure, but the scheme does not provide resilience against node failures, and its resource consumption is significant even if backup resources are shared.

A well-known scheme, which is also widely applied in practice, is path protection [33, 81, 109]. In case of path protection, after having found the working path one end-to-end backup path is searched for that is disjoint from the working path. This backup path is activated if any of the links or nodes on the working path fails, except for the failure of the source or destination nodes. Similarly to link protection, the backup resources may be either dedicated or shared. In case of the former the scheme is called dedicated path protection (DPP), while in case of the latter the scheme is referred to as shared (backup) path protection, or S(B)PP for short. Path protected connections are resilient to any

[1] Purely dynamic restoration methods are fully re-active methods. They take measures to find a new route for a connection only after a failure disrupts the connection, without any pre-assignment of network resources to connections. These methods are very flexible and resource-efficient, however, they are mainly suitable for best-effort services only.

single node or link failure on their working path. Shared (backup) path protection is considered to be among the most resource-efficient protection schemes [22].

There are several other schemes that divide the working path into smaller segments and provide protection to the different segments individually. Conceptually the simplest one is called sub-path protection, which is proposed first in [3]. When a connection is routed using this scheme the working path is divided into non-overlapping segments, and a disjoint backup path is sought for each segment between the end nodes of the segment. The division of the working path to segments may be arbitrary. However, various constraints may also be applied, e.g. the number of segments or the length of the segments may be limited. Backup paths assigned to the same connection may share resources, if not prohibited by other constraints, as well as backup paths that belong to different connections. The former is called intra-sharing, while the latter is termed inter-sharing by [111]. The connections routed with sub-path protection are resilient to each single node and link failure on their working path except for that of the nodes at segment endpoints. The resource-efficiency of sub-path protection is somewhere between that of shared link protection and shared path protection [71–73].

Another segment based method that aims at single failure resilience is called protection using multiple segments (PROMISE) [111]. The authors of [111] propose to divide the working path to a sequence of overlapping segments so that each segment may only overlap its adjacent segments, and each overlap consists of two adjacent nodes and one link between them. Then a backup path is sought for each of these segments so that inter- and intra-sharing are both allowed. The selection of the working path is a critical point in the operation of the algorithm, because backup paths may not be available for an arbitrary working path. A comprehensive study on this is also presented by the authors in [110]. The connections routed with PROMISE may survive any single node or link failure along the working path.

A remarkable scheme introduced in [32] is called p-cycles. Originally proposed for link protection only, the concept is extended to node and path segment protection in [92] and even a failure independent variant has been proposed lately [51]. The scheme is designed and introduced for network dimensioning, but it has been extended to the dynamic traffic scenario, as well [31]. According to the original idea p-cycles are pre-configured protection cycles that provide protection for one unit of working capacity for links on the cycle and two units of working capacity for so-called straddling links. A link is called a straddling link, if both of its end nodes are on the cycle but the link itself is not an on-cycle link. The reason for this is that in case of the failure of an on-cycle link, the rest of the cycle is activated as a backup path, whereas in case of the failure of a straddling link the whole cycle is available for protection. The idea that distinguishes p-cycles is that instead of assigning backup resources to connections they assign backup resources to working capacities used by connections, similarly to link protection. The

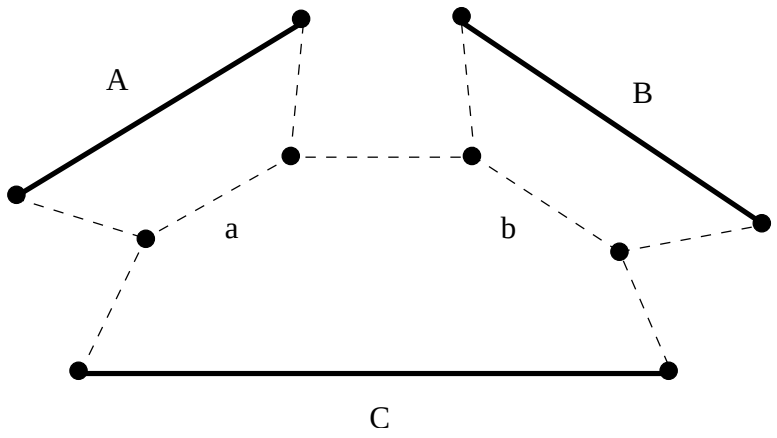


Figure 2.2: An example of backup resource sharing

resource-efficiency of p-cycles approaches that of shared (backup) path protection.

2.4.3 Analysis and enhancements of multiple failure resilience

Even though the methods reviewed so far only guarantee single failure resilience of connections, they provide solutions that are resilient against a portion of failures of higher multiplicity, as well. There are two reasons for this. The first one is the redundant structure of the assigned backup resources. E.g. even a simple shared (backup) path protection survives multiple simultaneous failures on the working path. The second reason is the fact that the assigned resources do not include all of the network components, and, therefore multiple failures that involve up to one component assigned to the connection are obviously survived, as well.

This observation is very important and is the basis of proposed methods in this thesis. However, when backup resources are shared the problem of quantifying multiple failure resilience of a connection becomes overwhelming. In what follows the problem is demonstrated with a simple example using shared (backup) path protection in a dynamic traffic scenario.

Consdier the scenario on Figure 2.2. Solid and dashed lines represent working and backup paths, respectively. Assume that connection A has just arrived and the on-line call admission control algorithm has to evaluate whether the depicted routing fulfils the survivability requirement. There are already two connections, B and C set up in the network. Connection B shares backup resource b with connection C. Similarly, connection A would share backup resource a with connection C. Shared backup resources of a connection are not always available, because another connection may activate them due to a component failure that affects that particular connection. Therefore, to compute the availability of connection A the probability has to be determined that connection C

will activate backup resource a. It will not only be the function of the failure probabilities of the working resources used by connection C, connection B might also interfere. Thus, a large number of admitted connections will have an indirect influence on each other through sharing.

Consider now the same situation from the viewpoint of connection C. Its survivability will change if connection A is admitted using the routing on Figure 2.2, as the perceived availability of resource a will decrease. Consequently, an on-line call admission control algorithm has to ensure also that survivability requirements of already admitted demands remain fulfilled.

Another related problem that can be demonstrated is the following. Consider the case when the working resources of connection B and C are failed simultaneously. A sharing conflict arises and the connection to survive will be the one that attempts to activate b, the shared backup resource, first. As a consequence, the order of failures is important in this case, which further increases the complexity of an exact analysis.

In conclusion, when backup resources are shared among multiple connections, an on-line call admission control algorithm, upon arrival of a new connection demand, not only has to check that there are resources available to route the incoming connection demand satisfying its availability requirement, but it must also check that the new connection demand does not decrease the availability of already active connections below their availability requirement.

Even though analysis of multiple failure resilience is not a problem specific to shared (backup) path protection, the majority of the literature considers S(B)PP only.

[22] proposes a straightforward experimental approach for the analysis, which considers double failures only in a network dimensioning context. This exhaustive failure state-space sampling considers each failure state in which exactly two components are failed and aims at computing dual failure restorability for each dual failure scenario. Dual failure restorability is defined as the ratio of non-restorable paths to the ratio of affected paths, and is thus not a connection-level availability measure. Basically the same approach is used in [14] and [88], although it is applied to preplanned and dynamic span restoration and p-cycles, respectively.

[71–73] analyze an extended set of failure states to derive connection availability for protection schemes with backup resource sharing in a dynamic traffic scenario. The proposed technique is to partition the total failure state space according to the failure states of the working path of the analyzed connection. Then, using the total probability theorem, bounds on *instantaneous connection availability* may be derived by examining the necessary conditions for the connection to remain uninterrupted considering the effect of backup resource sharing. If the sequence of failures in multiple failure scenarios is assumed to be favorable to the connection then an upper bound, in the opposite case a lower bound is obtained on connection availability. An important observation of [71]

is that if backup resources are shared, then asymptotic connection availability depends on the current network state, which changes as new connections are set up and old ones are torn down. The analysis method proposed in [72] is suitable for showing the amount of uncertainty in availability estimations that do not consider failure sequences. Nevertheless, the time complexity of the presented technique is exponential in the general case, which may be easily changed to polynomial in return for relaxing the tightness of bounds in a reasonable way.

[94, 112, 113] consider an availability estimation model for shared (backup) path protection. Assuming independent components path availability is computed in [94, 112, 113] as the product of availabilities of links, and connection availability is computed as that of a parallel system which comprises the working and the backup paths. In order to take into account the effect of shared backup resources, a multiplier is introduced for backup path availability which corresponds to the probability that the backup path is not available because another connection is using any of the shared resources. All of [94, 112, 113] give simple estimations of the multiplier based on the number of connections that share a given backup resource. This can only be accepted in special cases, one of which is used in [113] to verify the correctness of the formula by simulation. However, in the general case neither of these methods provides a conservative estimation of connection availability.

[112, 113] define the studied problem in a network dimensioning context, whereas [94] assumes dynamic traffic and simplifies the case to full wavelength conversion capable networks. Even though the authors claim that the latter assumption may be easily relaxed, it does not seem to be straightforward, and they do not address the problem elsewhere.

[40] introduces the concept of link and resource availability (LRA) to express the probability that at least one wavelength is available on a given link. However, this LRA value is then used only to influence connection routing. Connection availability is estimated according to a similar model as in [94, 112, 113], but the multiplier is assumed to be 1. Hence this estimation is neither conservative.

A similar approach of availability estimation is followed in [46], but a more sophisticated formula is derived for the multiplier to adjust backup path availability. The formula depends on the size of the sharing group, which contains connections whose backup paths share resources with the backup path of the connection to be analyzed. The formula contains a term that estimates the probability that the analyzed connection can use its backup resources in spite of a failure that affects the working paths of connections in the sharing group. This term attempts to capture the total effect of all possible failure scenarios and failure sequences as follows.

The working paths of the connections in the sharing group are assumed to have equivalent failure probability (the average of the group), and the number of simulta-

neously failed connections is assumed to be binomially distributed. Connections in the sharing group are assumed to have access to the shared backup resource with equivalent probability in case of a sharing conflict. The multiplier is thus obtained as one over the expected number of connections competing for the shared backup resources in case of a sharing conflict multiplied by the probability that a sharing conflict occurs. Because of using average values instead of worst-case assumptions or bounds the method obtains a non-conservative estimation of connection availability.

The backup path availability multiplier approach is further improved in [5]. Assuming link failures only [5] introduces a true lower bound on connection unavailability. Instead of using the number of connections in the sharing group the probability that there is at least one link failure that affects any of the working paths of connections in the sharing group is used. This probability is a very good estimate of the probability of the occurence of a sharing conflict, and yields a conservative estimation of connection availability, albeit it is not proved by the authors.

An improvement over the approach of [5] is outlined in [102], which attempts to analyze failure sequence dependence. In order to make the problem tractable [102] proposes to divide the connections in the sharing group to two subgroups. The first one contains connections that always overtake the analyzed connection when a sharing conflict occurs, while the second one contains connections always overtaken by the analyzed connection in case of a sharing conflict. However, it seems to be difficult even to make such a division, because the order of activation attempts most probably depends on the location of the failure, and so a connection-based division may not be the best approach.

Failure sequences up to double failures are dealt with in [62] in a more appropriate way in order to estimate connection availability when shared (backup) path protection is used. [62] proposes a Markov chain to model network failure states up to double failures using individual component MTTR and MTBF values and assuming exponential distributions. Using the derived failure sequence probabilities connection availability may be better estimated. However, the estimation in [62] is neither conservative since it considers the probability of failures of higher multiplicity to be zero.

Finally, [61] attempts to visualize connection availability by introducing the concept of connection availability maps. The maps contain "isoprobability" curves of connection availability as a function of hop count and total connection distance for unprotected and dedicated path protected connections based on a simple model. [61] is also one of the few works that consider node failures besides link failures.

The reviewed analysis methods are useful on the one hand to provide insight to connection availability and, on the other hand, to improve the survivability of protection schemes. Both [22] and [90] discuss such improvements of shared (backup) path protection and p-cycles, respectively, in a network dimensioning context. Note that using protection schemes that guarantee single failure resilience is highly recommendable, since

they entail far less complexity than schemes that guarantee resilience against failures of higher multiplicity. At the same time the availability of connections may be kept at surprisingly high levels. Chapters 4 and 5 discuss this idea further and introduce efficient methods for dynamic connection provisioning with guaranteed availability using shared (backup) path protection.

2.4.4 Guaranteed multiple failure resilience

The first approach to extend single failure resilient methods in order to provide multiple failure resilience is, in fact, a correction, which ensures true single physical failure resilience. Cables that are represented by logically independent links in a network model may be installed in reality so that they partly share common sheats or ducts. There may be several reasons for this phenomenon related mainly to cost-efficiency, but the consequences are more important. If a single failure hits the common part of the paths, which is likely to cut through the whole sheath or duct, it leads to the simultaneous failure of two links that are modeled as independent ones. The solution is a refined network model in which shared risk link groups (SRLG's) are defined as suggested by [11]. A shared risk link group is a set of links that may be disrupted by a single physical failure. The maximal number of shared risk link groups in a network equals the maximal number of physical components that may be shared among different links.

The concept is generalized to include other types of resources in [18], hence the name change to shared risk resource groups (SRRG's). If the working and backup paths are SRRG-disjoint then no single physical failures will disrupt the connection. This is a straightforward extension to any of the protection schemes introduced above; however, there are a few schemes introduced specifically with SRLG constraints.

[37] proposes a protection scheme called short leap shared protection, which is among the first ones proposed for single SRLG failure resilience. The working path is divided similarly to how PROMISE works [111], but SRLG-disjoint backup paths are sought for working path segments instead of link disjoint ones. Segment shared protection, proposed by [38] is based essentially on the sime idea; however, the method discussed in [38] is developed for bandwidth guaranteed tunnels and is thus more general. Moreover, [38] addresses the problem of optimal routing — also considered in [111] —, as well, and proposes a heuristic routing algorithm.

Another straightforward approach is presented by [19], called sub-graph routing. Due to the fact that physical failures may be localized it is possible to identify the failed SRLG. Therefore, a set of subgraphs of the original network model may be maintained, each one of which represents the state of the network when a particular SRLG is failed. These network configurations may be computed and optimized in advance, and if a failure occurs, the network state has to be changed to match the one recorded in the

subgraph corresponding to the failure. When optimizing network states certain constraints have to be observed to minimize the amount of necessary reconfigurations [19].

True multiple (physical) failure resilient methods are, in general, of higher complexity, and therefore of less interest. Nevertheless, there are a few works dedicated to this area.

As an example, [15] discusses ILP models for the dimensioning of completely and partially dual-failure restorable networks that rely on a link restoration scheme. [15] finds that the capacity required for complete dual-failure restorability is extremely high when compared to the capacity requirements of single-failure restorable designs. However, even in networks designed for single-failure restorability it is possible to selectively ensure dual-failure restorability with minor capacity additions. This is motivated by the interest of network operators expressed in the form of a question as "How to earn additional revenues from investing in survivability?" [15].

Network dimensioning with selective restorability enhancements is in the focus of [52], as well. The authors show a technique for p-cycles that is capable of differentiating the QoP. Moreover, by means of introducing pre-emption and backup capacity re-use, the proposed technique can selectively guarantee double failure resilience of connections. A special property of p-cycles is exploited for the purpose, namely, the fact that two units of capacity is available on each cycle for the protection of straddling flows. Thus, if a straddling connection of high priority is forced to use the cycle due to a failure on its working path and a second failure hits the cycle, the connection is allowed to pre-empt other connections potentially using the other part of the cycle as a backup path to restore connectivity. The idea basically trades some of the non-guaranteed structural dual-failure resilience of lower priority connections for selective guarantees on dual-failure resilience. The most important point is that no significant additional capacity is required compared to traditional single-failure resilient p-cycle designs. However, high priority connections must be in straddling relation to all cycles used for protection.

Specifically double failure resilient shared (backup) path protection schemes are proposed in [36, 97], but the solution of [97] may be theoretically extended to withstand failures of higher multiplicity, as well. Both works illustrate the fact that sharing rules for backup resources become highly complex when more than single failure resilience has to be guaranteed. Moreover, the capacity requirements of these schemes is very large compared to that of only single failure resilient ones, and they imply high connectivity in the network, which also limits their applicability.

2.4.5 Guaranteed availability

As already mentioned, the idea of connection level availability differentiation appeared first in [28]. After the introduction of the DiR concept in [28] the authors of [29] apply

the concept to shared (backup) path protection. The basic idea of the DiR concept is to change the assumption that each connection requires protection against every single failure along its working path. If availability requirements of connections are lower, but high enough that a single working path does not suffice, the backup path may be necessary to withstand the failure of only a subset of links used by the working path. This leads to somewhat relaxed sharing rules, which in turn yields increased capacity efficiency, or lower blocking in a dynamic traffic scenario [98].

The DiR concept is originally introduced with conditional failure probabilities, i.e., only single failure scenarios are assumed, and the conditional availability requirement of connections is guaranteed. This is extended to absolute failure probabilities in [75] with the simplifying assumption that any multiple failure leads to the disruption of any connection. In other words, guaranteed availability in terms of the maximum allowed number of simultaneously failed components is directly translated to an availability guarantee. While this ensures absolute availability guarantees, the estimation method is overly pessimistic, and there is plenty of room for improvements.

The papers [40, 60, 94, 112, 113] address connection establishment with availability guarantees in both network dimensioning and dynamic traffic contexts with shared backup resources. However, as they are based on non-conservative availability estimation methods discussed above ([60] uses the method proposed in [62]), their guarantee on connection availability is not a hard guarantee.

[102] proposes a network dimensioning method based on the conservative connection availability estimation method in [5]. The problem of dynamic connection provisioning with availability guarantees is only defined as a future research direction in [102].

To sum up, several publications address the establishment of connections with guarantees on failure resilience. These guarantees often appear as the highest component failure multiplicity that connections can withstand, and only a few works define availability guarantees among the objectives. The key to guaranteed availability is a conservative method for connection availability estimation. To the best of the author's knowledge so far only two such methods have been published [97, 102], none of which is directly applicable to a dynamic traffic scenario. It is, therefore, interesting to investigate the problem.

Chapter 3

Efficient Computation of Multi-Component Failure Stratum Probabilities

This chapter presents a simple method to compute the probability of multi-component failures of any arbitrary multiplicity under the assumption that individual component failures in the system are statistically independent [76]. In addition, the proposed method permits to determine the stratum probabilities for stratified sampling with linear complexity both in the number of components and in the number of strata.

3.1 Related work

Reliability analysis of systems in general — and of data networks, in particular — has been a topic of interest for decades. An example is the *K-terminal reliability*, which is the probability that a path exists in a graph from a source to each member of a set of nodes K. Notable special cases are *two-terminal reliability* and *all-terminal reliability*. Computing the exact values for these quantities is proved to be NP-hard in general [80]. However, a number of time efficient algorithms is available to produce bounds [93].

An interesting approach to analyze systems that are subject to multiple failures is based on stratified sampling [10]. In this approach, the amount of necessary computations is limited by means of evaluating only a subset of all the possible failure states of the system. First, failure states are grouped into subsets according to some criterion, e.g., based on the number of simultaneously failed components in the failure state, in other words, the failure multiplicity. Each subset of failure states is called a stratum. Then, samples of failure states are chosen from the different strata according to the stratum probabilities. Determining the stratum probabilities is thus a critical step in this method.

Stratum probabilities quantify the impact of failures of higher multiplicity on the system under study. They may even help determine the necessary methods and/or structural changes to ensure or improve system robustness. They may be computed using an algorithm that is linear in the number of components and quadratic in the

number of strata [10]. The algorithm in [10] requires that trees are built corresponding to the enumeration of failure scenarios, and derives a recursive formula for each of the stratum probabilities.

In what follows a solution of decreased computational complexity is presented, which facilitates an attractively simple failure multiplicity analysis for series-parallel systems. The solution makes use of probability-generating functions (PGF's) described next. Note that probability-generating functions are used here because they offer a compact, straightforward and widely understood notation of the quantities and operations that appear in the following discussion.

3.2 Proposed method

Assume that every component has two states, i.e., it is either *operating* or *failed*. Component failure events are assumed to be statistically independent. For every component x, it is assumed that q_x — the probability that component x is failed at any time — is known. q_x is thus the asymptotic unavailability of component x.

Consider a set of N statistically independent, two-state components, denoted by A. Let $q_n^{(A)}$ $(0 \leq n \leq N)$ be the probability that exactly n components in set A are failed. In other words, $q_n^{(A)}$ is the probability of the n^{th} multi-component failure stratum over set A. The quantities $q_n^{(A)}$ define a probability mass function over the set $\{0, 1, \ldots N\}$.

Assume that A and B are two disjoint sets of N and M statistically independent components, respectively. $q_n^{(A \cup B)}$ is obtained as the convolution sum of $q_n^{(A)}$ and $q_n^{(B)}$, i.e.,

$$q_n^{(A \cup B)} = \sum_{i=\max(0,n-M)}^{\min(n,N)} q_i^{(A)}\, q_{n-i}^{(B)}. \tag{3.1}$$

Let the probability-generating function of the probability mass function $q_n^{(A)}$ be defined as

$$Q^{(A)}(z) = \sum_{i=0}^{N} q_i^{(A)} z^i. \tag{3.2}$$

$Q^{(A)}(z)$ is always a polynomial, and the values $q_i^{(A)}$ are the coefficients of the respective power terms of z. It is known that the convolution in (3.1) can be expressed using the probability-generating functions as

$$Q^{(A \cup B)}(z) = Q^{(A)}(z)\; Q^{(B)}(z). \tag{3.3}$$

Lemma 1 (Convolution of truncated PGF's)**.** *If the probabilities $q_i^{(A \cup B)}$ are to be computed only for $0 \leq i \leq K < N + M$, the truncated probability-generating functions of the component distributions $Q^{(A)}(z)(mod\, z^{K+1})$ and $Q^{(B)}(z)(mod\, z^{K+1})$ suffice.*

Proof. Truncation of the probability-generating functions yields that on the right hand side of (3.1) the probabilities $q_i^{(A)}$ and $q_j^{(B)}$ are present only for $0 \leq i \leq \min(N, K)$ and $0 \leq j \leq \min(M, K)$, respectively. This set of probabilities suffices to determine $q_i^{(A \cup B)}$ for $0 \leq i \leq K$. □

3.2.1 Computing stratum probabilities

Given a set A of N statistically independent components, the probabilities $q_i^{(A)}$ $(0 \leq i \leq K)$ for the first $K+1$ strata may be computed by applying lemma 1, iteratively. Based on this observation the following algorithm is designed.

Let $Q^{(A)}(z, i)$ be a temporary variable of polynomial type computed during the i^{th} iteration. Let x_i be the i^{th} component in set A. Let $Q^{(\{x_i\})}(z) = (1 - q_{x_i} + q_{x_i} z)$.

Algorithm 1 Failure stratum probability computation

1: $i = 0$
2: $Q^{(A)}(z, 0) = 1$
3: **while** $i < N$ **do**
4: $\quad i = i + 1$
5: $\quad Q^{(A)}(z, i) = Q^{(A)}(z, i-1)\ Q^{(\{x_i\})}(z) \pmod{z^{K+1}}$
6: **end while**

In fact, $Q^{(A)}(z, i)$ gives the truncated probability-generating function of $q_n^{(\{x_j \in A | j \leq i\})}$ as a consequence of lemma 1. Therefore, the coefficients of $Q^{(A)}(z, N)$ are exactly $q_i^{(A)}$ for $0 \leq i \leq K$.

The computation of $Q^{(A)}(z, i-1)$ in step 5 requires $O(K)$ number of operations, as $Q^{(A)}(z, i-1)$ is a polynomial of order K at most, and $Q^{(\{x_i\})}(z)$ is a first order polynomial. Consequently, the algorithm requires $O(KN)$ operations for a set of N components. Note that the probability of a failure event of multiplicity higher than K in set A can be expressed as $1 - \sum_{i=0}^{K} q_i^{(A)}$, which can be used to provide bounds on availability.

For comparison, the algorithm in [10] requires $O(KN)$ operations to compute each $q_i^{(A)}$, and $O(K^2 N)$ operations for the entire set of probabilities.

3.2.2 Decomposition of component sets

Eq. (3.3) shows how the stratum probabilities of a composed set can be computed. It is worth mentioning that if the stratum probabilities are known for both sets $A \cup B$ and A, it is possible to compute those of set B without iterating through B component by component. If truncation is not used, the probability mass function $q_i^{(B)}$ may be obtained trivially from (3.3) by a simple polynomial division. However, if only truncated

probability-generating functions are available, a system of linear equations must be solved. The following notation is introduced to deal with the latter case.

Let C_K denote a $(K+1) \times (K+1)$ matrix, in which the subdiagonal elements are 1, and the others are 0. That is, $C_K(i,j) = 1$ if $j = i-1$, and $C_K(i,j) = 0$ otherwise. Let C_K^i be the i^{th} power of C_K. It follows from the definition that C_K is nilpotent, as $C_K^i = 0$ for $i > K$. Let $\mathbf{q}_K^{(A)} = \{q_i^{(A)}\}$ be a $(K+1)$-element column vector so that if set A contains N components and $N < K$ then $q_i^{(A)} = 0$ for $i > N$. Let a $(K+1) \times (K+1)$ lower triangular Toeplitz matrix $M(\mathbf{q}_K^{(A)})$ be composed of $\mathbf{q}_K^{(A)}$ as $\sum_{i=0}^{i=K} q_i^{(A)} C_K^i$. The coefficients of the truncated probability-generating function $Q^{(A \cup B)}(z) (\text{mod } z^{K+1})$ may then be expressed as:

$$\mathbf{q}_K^{(A \cup B)} = M(\mathbf{q}_K^{(A)})\ \mathbf{q}_K^{(B)}. \tag{3.4}$$

Lemma 2 (Inverse of lower triangular Toeplitz matrices)**.** *If $q_0^{(A)} \neq 0$, then the inverse of the lower triangular Toeplitz matrix $M(\mathbf{q}_K^{(A)})$ exists and can be expressed in the form*

$$M^{-1}(\mathbf{q}_K^{(A)}) = M(\mathbf{p}_K^{(A)}) = \sum_{i=0}^{K} p_i^{(A)} C_K^i,$$

where

$$\begin{aligned} p_0^{(A)} &= \frac{1}{q_0^{(A)}}, \quad \textit{and} \\ p_i^{(A)} &= -\sum_{j=1}^{i} \frac{q_j^{(A)}}{q_0^{(A)}} p_{i-j}^{(A)}, \quad \textit{for } 1 \leq i \leq K. \end{aligned}$$

Note that lemma 2 is a variant of a well-established result concerning lower triangular Toeplitz matrices [16].

If the conditions of lemma 2 are met, (3.4) can be solved for the unknowns $q_i^{(B)}$ by multiplying both sides from left by $M^{-1}(\mathbf{q}_K^{(A)})$. The inversion using lemma 2 and the multiplication requires $O(K^2)$ operations, compared to the $O(KM)$ operations that are required when running the algorithm for set B of M components.

3.3 Structured component sets

The approach presented in section 3.2 may be used to compute end-to-end availability of structured component sets in a compact and elegant way as follows. An example of a structured component set is the k-out-of-n:G system, which is operating if at least k of all the n components are operating. In what follows only the series (n-out-of-n:G) and parallel (1-out-of-n:G) systems are derived. Other configurations, e.g. k-out-of-n:G systems, may be derived in a similar way. In what follows the vector based notation introduced in section 3.2.2 is used for the ease of discussion.

Let $G = (V, E)$ be an undirected and connected graph used to model a structured system. V is the set of nodes. E is the set of edges. Each edge in E represents a two-state component. The states of the components (edges) are statistically independent. Each node in V represents a connection in between components represented by edges incident to the node. Let $s(G) \in V$ be the source and $d(G) \in V$, $s(G) \neq d(G)$ be the destination of G. Graph G models a system which is operating if and only if at least one path exists from the source to the destination that contains only operating edges.

A pair of failure probability vectors are defined for G. Both vectors are of dimension $K+1$ and are defined similarly to $\mathbf{q}_K^{(A)}$. The first vector ($\mathbf{m}_K^{(G)}$) contains the probabilities of all the failure events that may affect the components of the system modelled by G. Vector element $m_i^{(G)}$ is the sum of the probabilities of all failure events that involve i components from G, irrespective of whether system G as a whole remains operating after the failure event[1].

The second vector ($\mathbf{f}_K^{(G)}$) contains only the probabilities of the stratified failure events that cause the failure of the system modelled by G, i.e., that leave no path of operating edges from $s(G)$ to $d(G)$. By construction, $f_0^{(G)} = 0$, that is, with zero failed components the system is operating.

When combining such structured systems in series or parallel configurations, the following lemmas help derive the failure probability vectors of the resulting structured system. Let $\mathbf{r}_K^{(G)} = \mathbf{m}_K^{(G)} - \mathbf{f}_K^{(G)}$ contain the probabilities of the stratified failure events that leave the system operating.

3.3.1 Series configuration

Let two structured component sets $A = (V^{(A)}, E^{(A)})$ and $B = (V^{(B)}, E^{(B)})$ be combined to form a series configuration. Assume that $E^{(A)} \cap E^{(B)} = \emptyset$ and either

(a) $V^{(A)} \cap V^{(B)} = d(A) = s(B)$, or

(b) $V^{(A)} \cap V^{(B)} = d(B) = s(A)$.

The combined set is $AB = (V^{(A)} \cup V^{(B)}, E^{(A)} \cup E^{(B)})$. In case (a), $s(AB) = s(A)$ and $d(AB) = d(B)$. In case (b), $s(AB) = s(B)$ and $d(AB) = d(A)$.

Lemma 3. *The failure probability vector pair for AB is:*

$$\mathbf{m}_K^{(AB)} = M(\mathbf{m}_K^{(A)})\, \mathbf{m}_K^{(B)}, \tag{3.5}$$

$$\mathbf{f}_K^{(AB)} = \mathbf{m}_K^{(AB)} - M(\mathbf{r}_K^{(A)})\, \mathbf{r}_K^{(B)}. \tag{3.6}$$

[1] $m_i^{(G)} = q_i^{(E)}$ according to the notation used in Section 3.2.

Proof. (3.5) follows directly from (3.4). For (3.6) it must be observed that $f_i^{(AB)} \leq m_i^{(AB)}$. The probability of failures of multiplicity i that leave a path of operating edges both between $s(A)$ and $d(A)$ and between $s(B)$ and $d(B)$ is $r_i^{(AB)} = m_i^{(AB)} - f_i^{(AB)}$. Since $E^{(A)} \cap E^{(B)} = \emptyset$, the following holds true: $r_i^{(AB)} = \{M(\mathbf{r}_K^{(A)})\ \mathbf{r}_K^{(B)}\}_i$, which can be explained similarly as (3.1). □

3.3.2 Parallel configuration

Let two structured component sets $A = (V^{(A)}, E^{(A)})$ and $B = (V^{(B)}, E^{(B)})$ be combined to form a parallel configuration. Assume that $E^{(A)} \cap E^{(B)} = \emptyset$ and $V^{(A)} \cap V^{(B)} = \{s(A) = s(B), d(A) = d(B)\}$. The combined set is $A + B = (V^{(A)} \cup V^{(B)}, E^{(A)} \cup E^{(B)})$. $s(A + B) = s(A)$ and $s(A + B) = d(A)$.

Lemma 4. *The failure probability vector pair for $A + B$ is:*

$$\mathbf{m}_K^{(A+B)} = M(\mathbf{m}_K^{(A)})\ \mathbf{m}_K^{(B)}, \tag{3.7}$$

$$\begin{aligned}\mathbf{f}_K^{(A+B)} &= \mathbf{m}_K^{(A+B)} - M(\mathbf{r}_K^{(A)})\ \mathbf{r}_K^{(B)} \\ &\quad -M(\mathbf{r}_K^{(A)})\ \mathbf{f}_K^{(B)} - M(\mathbf{f}_K^{(A)})\ \mathbf{r}_K^{(B)}.\end{aligned} \tag{3.8}$$

Proof. (3.7) follows directly from (3.4). For (3.8) it must be remembered that $f_i^{(A+B)} \leq m_i^{(A+B)}$. The probability of failures of multiplicity i that leave a path of operating edges either between $s(A)$ and $d(A)$ or between $s(B)$ and $d(B)$ or both of them is given by $r_i^{(AB)} = m_i^{(AB)} - f_i^{(AB)}$.

Since $E^{(A)} \cap E^{(B)} = \emptyset$, the following holds true: $r_i^{(A+B)} = \{M(\mathbf{r}_K^{(A)})\ \mathbf{f}_K^{(B)}\}_i + \{M(\mathbf{f}_K^{(A)})\ \mathbf{r}_K^{(B)}\}_i + \{M(\mathbf{r}_K^{(A)})\ \mathbf{r}_K^{(B)}\}_i$, where the first term is the probability that system A remains operating and system B fails, the second term is the probability that system A fails and system B remains operating, and the third term is the probability that both systems remain operating, in case of a multi-component failure of multiplicity i. □

With lemmas 3 and 4 it is possible to estimate the effect of multi-component failures on structured systems whenever the graph that represents the system can be reduced to a single edge, by using the series and parallel reductions (see [93]).

3.4 Summary

This chapter introduced a simple method to compute the probabilities of multi-component failures with any multiplicity. The proposed method offers a straighforward way to compute the availability of structured systems, of which the series and parallel configurations are just two examples illustrated here. When applied to the computation

of stratum probabilities in stratified sampling, the proposed solution has a computational complexity that is linear in the number of components and in the number of strata. This is an improvement over previously proposed methods.

Chapter 4

DiR with Node Failures and Absolute Probabilistic Guarantees

This chapter investigates what impact optical node failures may have on WDM networks, in which reliable end-to-end optical circuits are provisioned dynamically [78]. At the node level, the optical cross-connect (OXC) equipment availability measure is estimated using proven component level availability models. At the network level, end-to-end optical circuits are provisioned only when the level of connection availability required by the application can be guaranteed.

With the objective of yielding efficient utilization of the network resources, i.e., fibers and OXC's, circuit redundancy is achieved by means of shared (backup) path protection (S(B)PP) switching, in combination with Differentiated Reliability (DiR). The resulting optimal routing and wavelength assignment problem is proven to be NP-Complete. To produce suboptimal solutions in polynomial time, a heuristic technique is presented, which makes use of a time-efficient method to estimate the end-to-end circuit availability in the presence of multiple (link and node) failures.

Using the proposed heuristic, a selection of representative OXC architectures and optical switching technologies is examined to assess the influence of node equipment choice on the overall network performance.

4.1 Related work

Quality of service awareness has gained vital importance in service provisioning with the roll-out of applications that impose quality requirements on data transfer. Among other things, fulfilling these requirements necessitates that the underlying networking technology be capable of offering end-to-end transport services at satisfactory availability levels. To meet end-to-end availability requirements, the impact of employed network components on service quality must be evaluated. This impact must be quantified during the service provisioning process, to help determine efficient assignments of network resources to traffic demands.

Most of the literature on optical circuit-switched network availability, e.g. [22, 69, 81, 87, 89], makes the fundamental assumption that the failure probability of optical node equipment is negligible, when compared to link failure probability. While this may be true in a number of cases, a comprehensive analysis must take into consideration optical cross-connect (OXC) architecture and switching technology, as well [61]. Given the wide range of availability and cost options available today, the selection of the OXC architecture may play an important role in certain networks.

Stand-alone availability analysis of most optical switching components is already available [106, 107]. However, few studies address jointly network level and switching equipment level availability. A notable contribution in this direction focuses on the interworking of the optical and the IP layers to achieve protection switching in bidirectional circuit switched rings [35]. Another interesting contribution considers various protection schemes to be applied for circumventing node failures in all-optical networks [50]. In the latter work, network efficiency and failure impact are investigated, producing valuable insight into end-to-end circuit availability. [61] also considers end-to-end connection availability with a detailed reliability model of nodes and demonstrates the relation of connection length and feasible availability guarantees.

The scope of this chapter is to investigate the impact of OXC equipment failure on end-to-end optical circuit provisioning in WDM networks with arbitrary topologies.

It is assumed that the objective to accomplish is to maximize network utilization by providing the minimal level of circuit redundancy that satisfies the application's availability constraint, e.g., as specified by the Service Level Agreement (SLA). To achieve this objective, both node level and network wide availability calculations are combined into a single availability provisioning framework. At the node level, the equipment availability measure is calculated using proven component-level availability models [108], that are tailored to represent a number of OXC architectures and switching technologies. Node level availability calculations are then used to study the availability of end-to-end optical circuits using probability-based requirements.

Circuit redundancy is achieved by means of shared (backup) path protection (S(B)PP) switching, combined with the Differentiated Reliability technique, or SPP-DiR for short [64].

Taking into consideration the increasing desire among network operators to provide solutions that cope with multiple simultaneous failures, the study is carried out assuming failures of any multiplicity.

Dealing with multiple failures represents a double challenge. First, the exact computation of *two-terminal reliability* is, in general, an NP-hard problem [93] irrespective of whether the nodes are failure prone. Second — as proven later — solving the optimal Routing and Wavelength Assignment (RWA) problem for SPP-DiR is NP-Complete. To circumvent these two challenges, an existing suboptimal RWA algorithm is com-

bined with a polynomial-time heuristic method to bound end-to-end circuit availability in the presence of both link and node failures. The heuristic method overcomes some drawbacks of the existing approximation methods that are based on either bounding techniques, (e.g., [21]) or state-space sampling (e.g., [10]). Bounding techniques usually remain general and do not address specific characteristics of the protection scheme being used. By their nature, they tend to offer less efficient solutions. Sampling methods, on the other hand, carefully take into account network behavior and the protection scheme being used. Their drawback is the exponential size of the state-space to be sampled, which restricts their applicability to off-line analysis of problems with modest size.

The proposal of the heuristic method is also justified by the fact that, to the best of the author's knowledge, all methods available in the literature, which are suitable for on-line connection provisioning, e.g. [40, 62, 94, 112, 113], provide non-conservative bounds.

With the aid of the proposed availability framework and the polynomial-time heuristic method, a study is carried out assuming dynamic arrivals of call requests. The network topology, the link capacities, and the OXC equipment type used at the nodes are given as input parameters. Incoming call requests arrive with predetermined connection availability requirements, and are admitted only when there are enough available resources in the network to meet their connection availability constraints. A number of OXC architectures and technologies is investigated, which offers a range of options in terms of both equipment cost and network utilization. Conclusions are drawn regarding favorable OXC choices in a number of network scenarios and sizes.

4.2 Network level availability

This section describes the fundamental assumptions concerning the network level availability model.

A wavelength on a link constitutes a transport channel. One such channel or the concatenation of two or more channels constitutes an optical circuit, or *lightpath*. It is assumed that wavelength converters are not available, that is, a lightpath must be assigned the same wavelength on all the links it traverses. It is assumed that a network with an appropriate description of the availability characteristics of its components is given.

The problem of interest is to determine for each incoming call (circuit) request the routing and wavelength assignment from the source to the destination using available wavelengths, i.e., the RWA problem [67, 82]. The solution to the RWA problem must use wavelengths in an efficient way, while yielding the availability level requested by the incoming call, that is, satisfying the specified *availability constraints*. In simple terms, a good solution must find a tradeoff between efficient wavelength usage and availability

requirements.

When solving the RWA problem with availability constraints, assigning resources along a single (*working*) lightpath to a call request might not be enough to meet the specified connection availability requirement. Therefore, the following modified version of the SPP scheme [67, 82, 109] is used to meet the connection availability requirement.

When using the conventional SPP scheme, each working lightpath is assigned a route-disjoint — i.e., link and node disjoint — (*backup*) lightpath ready to be used if the working lightpath is affected by some failure. Each call request is thus survivable against any single network element, i.e., node or link, failure, except for the failure of its end nodes. Working and backup lightpaths of the same call request do not need to have the same wavelength assigned. Distinct backup lightpaths, whose corresponding working lightpaths are route-disjoint, are allowed to share the same link and wavelength.

The concept of Differentiated Reliability (DiR) [28, 96] is applied to improve the efficiency of the conventional SPP scheme, thus obtaining the so called SPP-DiR protection scheme. The basic notion of DiR is to provide just the desired level of protection for each call request — often referred to as demand —, which will eventually lead to more efficient resource usage.

The SPP-DiR scheme is derived from the SPP scheme as follows. For a demand with a less stringent availability requirement, the backup lightpath does not need to be always available for every possible network element failure scenario. In other words, it is possible to select a subset of network elements along the working lightpath, whose failure will disrupt the demand traffic, i.e., the backup lightpath will not be activated. This subset is carefully chosen to still satisfy the demand's specified availability constraint. Notice that two (or more) demands whose working lightpaths have a common link may also share a link and a wavelength along their respective backup lightpaths. This option is available when at least one of the two demands can afford not to resort to the backup lightpath upon the failure of the link that is shared by the working lightpaths. By the same reasoning, it is also possible to have a working lightpath completely unprotected if the working lightpath failure probability still satisfies the availability requirement.

Formally, let directed graph $G(V, E)$ represent the network, where V is the set of nodes and E is the set of edges, which represent network links. Call requests enter and leave the network dynamically. Let a call request or demand be defined as $d = (n_{\mathrm{s}}^{(d)}, n_{\mathrm{d}}^{(d)}, t_{\mathrm{a}}^{(d)}, t_{\mathrm{h}}^{(d)}, r^{(d)})$, where $n_{\mathrm{s}}^{(d)} \in V$ is the source node and $n_{\mathrm{d}}^{(d)} \in V$ is the destination node, $t_{\mathrm{a}}^{(d)} > 0$ and $t_{\mathrm{h}}^{(d)} > 0$ are the time of arrival of the demand and the holding time of the connection, respectively, while the probability that the connection is interrupted by any failure in the network should be at most $r^{(d)}$. Note that parameters $t_{\mathrm{a}}^{(d)}$ and $t_{\mathrm{h}}^{(d)}$ appear here in order to adjust the definition to that of the dynamic shared-path protected lightpath provisioning problem [69], referred to later.

Let $L_{\mathrm{w}}^{(d)} \subset E$ and $N_{\mathrm{w}}^{(d)} \subset V$ be the set of links and nodes used by the working

lightpath of demand d, respectively. Let $L_{\mathrm{b}}^{(d)} \subset E$ and $N_{\mathrm{b}}^{(d)} \subset V$ be the set of links and nodes used by the backup lightpath of demand d, respectively. Let $L_{\mathrm{p}}^{(d)} \subseteq L_{\mathrm{w}}^{(d)}$ and $N_{\mathrm{p}}^{(d)} \subseteq N_{\mathrm{w}}^{(d)}$ be the set of protected links and nodes along the working lightpath of demand d, that is, the set of links and nodes whose failure leads to the activation of the backup lightpath(s). Note that $n_{\mathrm{s}}^{(d)} \notin N_{\mathrm{p}}^{(d)}$ and $n_{\mathrm{d}}^{(d)} \notin N_{\mathrm{p}}^{(d)}$ are always true. For the sake of compactness let $\mathcal{W}^{(d)} = L_{\mathrm{w}}^{(d)} \cup N_{\mathrm{w}}^{(d)}$ be the set of network elements, i.e., link and nodes, used to route the working lightpath of demand d. Let $\mathcal{P}^{(d)} = L_{\mathrm{p}}^{(d)} \cup N_{\mathrm{p}}^{(d)}$ be the set of protected network elements along the working lightpath of demand d. Let $\mathcal{U}^{(d)} = \mathcal{W}^{(d)} \setminus \mathcal{P}^{(d)}$ be the set of unprotected network elements along the working lightpath of demand d. Let $\mathcal{B}^{(d)} = L_{\mathrm{b}}^{(d)} \cup N_{\mathrm{b}}^{(d)}$ be the set of network elements used to route the backup lightpath of demand d. The resource sharing rule in SPP-DiR for the backup lightpaths of demands d_1 and d_2 can be formally defined as follows: $L_{\mathrm{b}}^{(d_1)}$ and $L_{\mathrm{b}}^{(d_2)}$ may share a wavelength on a link if

$$(\mathcal{W}^{(d_1)} \cap \mathcal{W}^{(d_2)}) \subseteq (\mathcal{U}^{(d_1)} \cup \mathcal{U}^{(d_2)}). \tag{4.1}$$

4.2.1 Estimation of end-to-end availability

Motivated by the need to compute availability in polynomial time the following lower bound is proposed.

It is assumed that the network consists of atomic nodes and links that connect the nodes. Every component in the network can have two states only: it is either *operating* or *failed*. It is assumed that the availability characteristics of the components in the network are independent of each other, and are known in advance.

Let q_e be the asymptotic unavailability of link $e \in E$, and let q_v be the asymptotic unavailability of node $v \in V$. Let

$$q_{1+}^{(C)} = 1 - \prod_{c \in C}(1 - q_c) \tag{4.2}$$

be the probability that there is at least one network element that is failed in set C of network elements — i.e., links and nodes.

Using the notations introduced in the previous section, the $p^{(d)}$ availability for demand d is given by:

$$p^{(d)} = (1 - q_{1+}^{(\mathcal{U}^{(d)})})(1 - q_{1+}^{(\mathcal{P}^{(d)})} q_{\mathcal{B}}^{(d)}), \tag{4.3}$$

where $q_{\mathcal{B}}^{(d)}$ is the probability that the backup lightpath is not available. Two factors contribute to this probability:

1. $q_{1+}^{(\mathcal{B}^{(d)})}$, i.e., the probability that at least one network element in set $\mathcal{B}^{(d)}$ is failed, and

2. the probability that another demand, e.g., d_2, that shares backup resources with demand d, has already activated its backup lightpath due to a failure. Note that in this case the failure that prevents demand d from using its backup lightpath is not the failure of a network element in set $\mathcal{B}^{(d)}$.

The second factor makes the exact calculation difficult and time consuming. In order to quickly evaluate the availability of a circuit, a conservative bound is calculated instead of the exact value. The bound is obtained as follows:

$$\check{p}^{(d)} = (1 - q_{1+}^{(\mathcal{U}^{(d)})})(1 - q_{1+}^{(\mathcal{P}^{(d)})} q_{1+}^{\left((E \cup V) \setminus \mathcal{W}^{(d)}\right)}). \tag{4.4}$$

In other words, the circuit is considered to be disrupted if either any unprotected component on the working lightpath fails or there is at least one failure among the protected components of the working lightpath and at least one failure among all the components not used by the working lightpath.

Lemma 5. *$p^{(d)}$ is always lower bounded by $\check{p}^{(d)}$.*

Proof. $(E \cup V) \setminus \mathcal{W}^{(d)}$ contains all the components that potentially contribute to $q_{\mathcal{B}}^{(d)}$, therefore $q_{1+}^{\left((E \cup V) \setminus \mathcal{W}^{(d)}\right)} \geq q_{\mathcal{B}}^{(d)}$ always holds true. Hence $\check{p}^{(d)}$ is always a lower bound. □

The rationale behind this estimate is the worst case assumption that any failure scenario that affects at least one element of $\mathcal{P}^{(d)}$ and at least one element that does not belong to the working lightpath of demand d leads to the unavailability of the shared backup resources of the circuit. This pessimistic assumption always yields a conservative bound, as shown by lemma 5 and makes it simple to verify if the availability requirement $r^{(d)}$ of a demand is satisfied, i.e., only

$$\hat{q}^{(d)} = 1 - \check{p}^{(d)} \leq r^{(d)} \tag{4.5}$$

has to be checked.

Albeit a more accurate conservative bound is available [5], the bound in (4.4) proves to be sufficiently accurate with the component failure statistics considered in section 4.3.3 and in section 4.4.1. Moreover, by means of using (4.4) it is not necessary to check whether the availability requirements of active demands are violated by the newly arrived one, which is a major decrease in computation complexity when compared to [5]. Without solving this problem the approximation method presented in [5] cannot be directly applied in an on-line RWA algorithm.

Algorithm 2 shows the pseudo code of the outlined availability bounding technique applied to the call admission control of SPP-DiR. In order to make it more concise, $D^{(b)}(t)$ is defined as the set of active demands whose backup lightpath uses backup

Algorithm 2 Extended DiR RWA check algorithm

procedure EDIR_RWA_CHECK($d, D^{(\cdot)}(t_{\mathrm{a}}^{(d)}), \mathcal{W}^{(d)}, w_{\mathrm{w}}, \mathcal{P}^{(d)}, \mathcal{B}^{(d)}, w_{\mathrm{b}}$)
  $accept = true$
  for all $e \in L_{\mathrm{w}}^{(d)}$ **do**
    if (e, w_{w}) is already used **then**
      $accept = false$
    end if
  end for
  for all $e \in L_{\mathrm{b}}^{(d)}$ **do**
    for all $d_i \in D^{(e,w_{\mathrm{b}})}(t_{\mathrm{a}}^{(d)}, d)$ **do**
      if $(\mathcal{W}^{(d_i)} \cap \mathcal{W}^{(d)}) \nsubseteq (\mathcal{U}^{(d_i)} \cup \mathcal{U}^{(d)})$ **then**
        $accept = false$
      end if
    end for
  end for
  for all $n \in N_{\mathrm{b}}^{(d)}$ **do**
    for all $d_i \in D^{(n)}(t_{\mathrm{a}}^{(d)}, d)$ **do**
      if $(\mathcal{W}^{(d_i)} \cap \mathcal{W}^{(d)}) \nsubseteq (\mathcal{U}^{(d_i)} \cup \mathcal{U}^{(d)})$ **then**
        $accept = false$
      end if
    end for
  end for
  if $\hat{q}^{(d)} > r^{(d)}$ **then**
    $accept = false$
  end if
  return *accept*
end procedure

resource $b = (e, w)$, i.e., wavelength w on link e, at time t. Note that $D^{(n)}(t),\ n \in V$ is interpreted as the set of active demands whose backup lightpath traverses node n.

The extension of the SPP-DiR method to absolute availability guarantees as described in this section will be referred to as SPP-eDiR.

Further details of the implementation of the provisioning method are discussed in section 5.6.

4.3 Node level availability

In what follows the assumptions made for modeling the availability of equipment deployed at network nodes are introduced.

4.3.1 Switching component technologies

For a certain application one has to consider the required switching speed, crosstalk, insertion loss, etc., as well as reliability performance and price. In this study, the choice of optical switch components is made based upon the basic requirements of -45 dB crosstalk, -50 dB return loss and minimum insertion loss. For protection switching purposes the switching time does not need to be particularly short, e.g. the order of ten milliseconds is acceptable. A large number of ports is not necessary in the network scenarios reported in section 4.4.1.

Optical switches can be classified into two categories: solid state (or integrated optic) and free-space. Switches can be made of a number of materials such as lithium niobate, indium phosphide, silicon, etc., and the switching performance depends on the component technology.

In this study the indium phosphide (InP) semiconductor optical amplifier based technology is considered as a representative of the first category. Even though integrated optical switch matrices are typically characterized by a small number of input/output ports, they could be a viable alternative up to 8×8, which suffices for the network scenarios to be studied. Semiconductor optical amplifier based switches offer high switching speed, loss compensation and high extinction ratio. However, amplifier noise accumulation and ring lasing may represent a problem for these switches in WDM networks. Additionally, this technology has relatively high costs and poor reliability performance.

From the second category the planar tilt mirror type Micro-Opto-Electro-Mechanical Systems (MOEMS) switch [39] is selected here, being a representative of the most promising free-space switching technology, optical MEMS. These switches are characterized by a lower switching speed (typically several milliseconds) than that of integrated optical switches. As opposed to sliding or pop-up vertical mirror MOEMS switches [54, 65] in planar tilt mirror type switches a servo controlled tilt adjustment

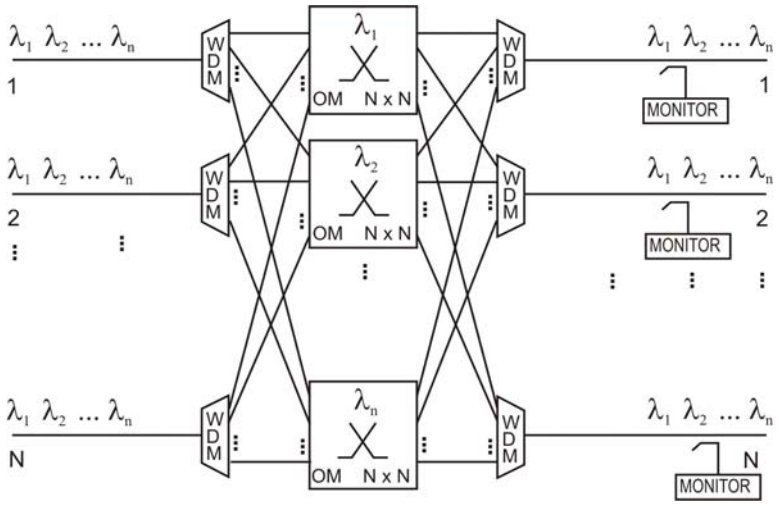


Figure 4.1: Node architecture A (without redundancy)

minimizes the insertion loss. It also compensates for spatial drift due to thermal expansion, creep, or other such phenomenon; therefore, these switches have superior properties. Other reasons to justify the selection of the MOEMS technology are the high reliability performance of its components, and the relatively low implementation cost due to the mature planar silicon technology.

4.3.2 Node architectures

A representative internally non-blocking node architecture (see node architecture A shown in Figure 4.1) is considered. The node has N input and output fibers and W wavelength channels carried on each fiber. The node degree N varies from 2 to 7, and the number of wavelengths per fiber W takes the values of 4, 32 and 64 in the network examples studied in section 4.4. This node architecture is selected because of its low complexity, which allows for high availability and relatively low price.

In order to improve the node availability and provide additional functionality, the inherent node redundancy proposed in [107] is considered. The modified node architecture (node architecture B) is illustrated in Figure 4.2 [107]. An optical tap is used to split the signals prior to wavelength demultiplexing. One part of the signal enters an optical space switch that serves as the primary switch core. The tapped portion is connected to a tunable receiver at the protection switch and switched to appropriate tunable transmitter by the electronic cross-point switch. The switched wavelength channels are multiplexed to the output fibers. In the case of the failure of one $N \times N$ switch matrix, the corresponding wavelength channels are guided through the protective electronic circuit, and added to the output links by the optical couplers.

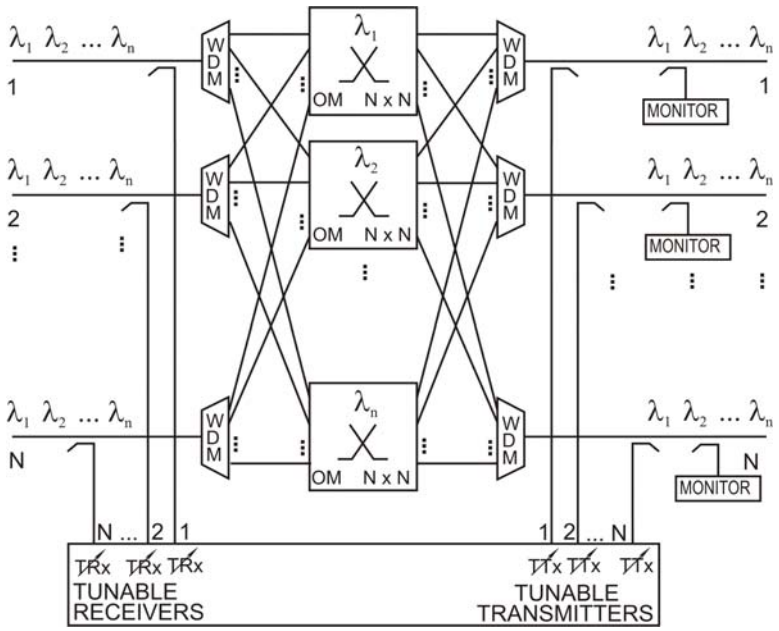


Figure 4.2: Node architecture B (with built-in 1:N redundancy)

4.3.3 Node level availability model

A component-based approach for node availability calculations proposed in [108] is adopted and improved here. All equipment is assumed to be composed of independent components that have two states: they are either operating or failed, and the whole equipment is also assumed to have only these two states. This node availability model is in compliance with the requirements of the model proposed in section 4.2. One particular difference with respect to the model persented in [108] is that components used at link terminations are included in the availability model of the respective link as they are required for the operation of that link only. The availability measure used in the network availability model is thus derived considering the switching matrices only.

Reliability block diagrams [57] are constructed based on the node architecture and functionality by applying the basic principles of availability performance evaluation. Availability measures are then assigned to the individual components of these block diagrams, which can be derived from measurement statistics. Based on component level data, estimated node level availability measures are derived, taking into account the interconnections represented by the block diagram. The block diagram may consist of well-known configurations, such as series, parallel, or k-out-of-n configurations. The derived node level availability measures include the asymptotic availability and unavailability: the former is defined as the probability that the system (or component) is operating at any time, while the latter is the complementary probability.

In addition to the used technologies and redundancy of the architecture, the OXC

Node architecture, technology, size and W	**Unavailability** $[*10^{-6}]$	**Mean downtime per year [h]**
A / InP / 2×2 / 4	192.0	1.65888
A / InP / 4×4 / 4	576.0	4.97664
A / InP / 7×7 / 4	1872.0	16.17408
A / MOEMS / $N \times N$ / 4	24.0	0.20736
B / InP / 2×2 / 4	0.0230378	0.000199
B / InP / 4×4 / 4	0.2073	0.001791
B / InP / 7×7 / 4	2.18819	0.0189
B / MOEMS / $N \times N$ / 4	0.00036	0.000003
A / InP / 2×2 / 32	1536.0	13.27106
A / InP / 4×4 / 32	4608.0	39.81312
A / InP / 7×7 / 32	14976.0	129.39264
A / MOEMS / $N \times N$ / 32	192.0	1.65888
B / InP / 2×2 / 32	1.21531	0.0105
B / InP / 4×4 / 32	10.9161	0.094315
B / InP / 7×7 / 32	114.532	0.98956
B / MOEMS / $N \times N$ / 32	0.0178539	0.0001543
A / InP / 2×2 / 64	3072.0	26.54208
A / InP / 4×4 / 64	9216.0	79.62624
A / InP / 7×7 / 64	29952.0	258.78528
A / MOEMS / $N \times N$ / 64	384.0	3.31776
B / InP / 2×2 / 64	4.63566	0.04
B / InP / 4×4 / 64	41.5558	0.359
B / InP / 7×7 / 64	433.102	3.742
B / MOEMS / $N \times N$ / 64	0.072558	0.000627

Table 4.1: Node availability measures

	Unavailability $[*10^{-6}]$	
W	**Node architecture A**	**Node architecture B**
4	0.6	0.9
32	2.4	2.7
64	4.8	5.1

Table 4.2: Unavailability of node components at link termination points

availability measure depends on complexity, i.e., on switching capacity. Throughout the availability calculations and the traffic simulations later to be discussed, each link is assumed to have a capacity of W wavelengths, which thus determines the necessary switching capacity: each OXC has to support W wavelengths on all of its interfaces.

Table 4.1 presents node level availability results expressed by the asymptotic node unavailability, and mean downtime per year. Calculations are based on previously published component availability figures [107, 108]. When calculating the availability characteristics, the mean repair time is assumed to be 6 hours.

Devices used at nodes for link termination, i.e., splitters, transmitters, receivers, couplers and multiplexers are treated as components connected to links in series configuration. That is, when determining the availability measure of links the failure rate of these optical components is also considered in addition to the inherent failure characteristics of fibers. Table 4.2 displays the asymptotic unavailability of link terminations at nodes, derived using the same technique as in case of node level availability results.

4.4 Results

This section documents the impact of OXC failures on end-to-end availability by applying the proposed model in a dynamic call admission control scenario in WDM networks without wavelength converters.

4.4.1 Assumptions on the studied scenarios

First, the common assumptions used throughout the experiments are introduced: the derivation of availability parameters, the applied simulator and the studied network scenarios are overviewed.

Availability characteristics of components

The asymptotic unavailability of each component has to be determined in order for the proposed model to be applied. The output of the method described in section 4.3.3,

which considers the full structure of the equipment deployed at nodes, gives the unavailability for the switching part of the nodes (q_v) and the link terminations at nodes.

The inherent asymptotic unavailability of links is expressed using (2.2) as discussed in section 2.3. q_e is then obtained by computing the asymptotic unavailability of a series system that contains two link terminations (c.f. section 4.3.3) and a component with the inherent asymptotic unavailability of the link computed using (2.2).

Note that the highest value of $q_{1+}^{(E\cup V)}$ computed during all of the experiments is 0.339976, which applies to the continental topology with 64 wavelengths/link using the least reliable, InP-based nodes with architecture A. On the other hand, the lowest such value is more than two orders of magnitude below, which shows that (4.4) yields estimations that are not overly conservative.

Simulator and RWA with availability guarantees

In order to assess the effect of OXC failures on end-to-end availability, a centralized call admission mechanism based on the SPP-eDiR model described in Section 4.2 is implemented using the simulation framework detailed in Appendix B.

The simulation takes both the network configuration and the availability requirement $r^{(d)} = r$ of the generated calls as input. Call generation is a Poisson process with parameter λ. Source and destination nodes of calls are selected uniformly and the same availability requirement r is considered for all the generated calls. The established circuits release resources after a time $t_{\mathrm{h}}^{(d)}$ that is exponentially distributed with parameter $\mu = 1$.

Assuming that complete information of the state of the network is available, the problem of deciding whether a working and backup lightpath pair exists that satisfies the availability requirement of a call request (SPP-DiR) can be proved to be NP-Complete, even if availability computations only require polynomial time[1]. Therefore, efficient resource allocation for each call is done by running an optimization process based on Simulated Annealing (SA). Only node disjoint working and backup lightpath pairs are considered.

The cost function used for the optimization of the resource allocation for an incoming demand d is

$$\mathrm{cost}(d) = \frac{|\mathcal{W}^{(d)}| + |\mathcal{B}^{(d)}| - |\mathcal{S}^{(d)}|}{2} + r - \hat{q}^{(d)}, \tag{4.6}$$

where $\mathcal{S}^{(d)}$ is the set of links along the backup lightpath in which the spare wavelength is shared by at least one other backup lightpath.

A more comprehensive discussion of the details for near-optimal lightpath selection subject to the constraints of the problem can be found in Appendix B.

[1]For the concise problem statement and the proof the reader is referred to Section 4.6.

The number of rejected and total call requests is recorded for each simulation, and the ratio of the two gives the call blocking probability, which is used to point out the difference between the performance of different network configurations.

Network scenarios

Three network scenarios are considered. The first one is that of networks of continental size. The European WDM network [96], depicted in Figure A.2, is chosen to represent this category. The second category of interest is that of national networks. This category is represented by the Italian WDM network on Figure A.3 [17]. Finally, the third category includes metropolitan area networks. This category is represented by the WDM metropolitan network topology used for simulations in [101].

In order to ease comparison between networks of different scale the differences in blocking due to differing topologies need to be eliminated. Therefore, the European network topology with the link lengths reduced by a factor of 5 and 50 is also used to obtain results for the national and metropolitan scales, respectively. For an overview of the characteristic parameters of these network topologies the reader is referred to Appendix A.

4.4.2 Discussion of results

The following discussion is based on a large number of simulations executed with various combinations of networking scenarios, switching technologies, node architectures, numbers of wavelength channels per fiber and availability requirements.

The presented simulation results have 10% confidence intervals at a 95% confidence level.

Effects of difference in the topology

When making comparisons between results obtained for different network topologies one has to be aware of the differences in blocking probability that may derive from the difference in the topology. Two such reasons for difference may be identified. The effect of the first one, the change of W while keeping the connectivity of the network unchanged, is presented in Figure 4.3. One may observe that the increased capacity of links allows for higher loads to be served with the same blocking probability.

The second reason is the change in the number of nodes and/or links, whose effect is somewhat harder to predict. Figures 4.4 and 4.5 show a comparison between the scaled versions of the continental topology and the representatives of the networks of national and metropolitan scale, respectively. One may conclude that change in the number of nodes and links leads to an earlier/later saturation of the network as a function of load. Since the continental topology is more connected than any of the other two representative

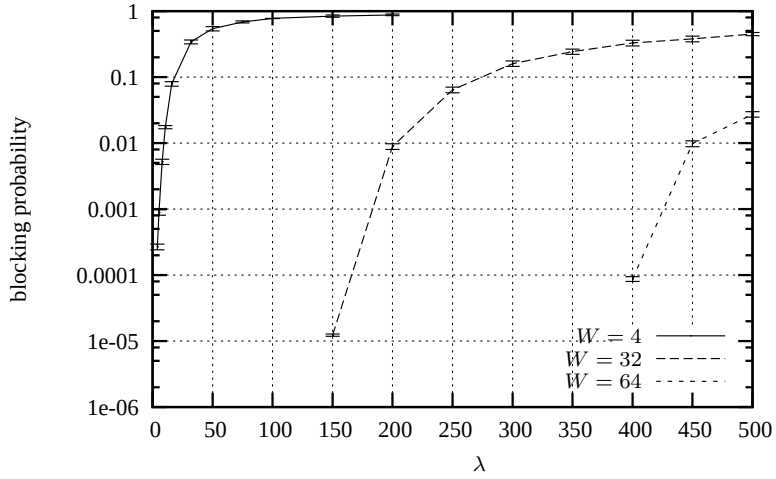


Figure 4.3: Blocking in the continental network using B/MOEMS switches with $r = 0.005$ as a function of network load and W

topologies, more resources are available for call requests and thus blocking is lower. The difference is larger in case of the metropolitan scenario, as suggested by the higher difference between the average node degrees.

Note that even though these differences exist, *qualitatively* the same behavior is expected in all of the studied topologies, which is confirmed by Figures 4.4 and 4.5.

Effect of the availability constraint

Figures 4.4, 4.5, and 4.6 report results obtained when the node equipment uses MOEMS technology in a redundant architecture (node architecture B) when each fiber carries 32 wavelengths. Figure 4.4 shows how availability constraints translate to blocking of call requests in the national network. The difference between the curves for $r = 0.004$ and $r = 0.0005$ increases with the decreasing values of λ. Figure 4.7 shows the blocking of call requests as a function of the availability constraint taken from Figure 4.4 at $\lambda = 250$. The range of values for which network performance improves due to the additional sharing made possible by the SPP-eDiR scheme is clearly visible. When the availability constraint r is in the 10^{-4} range, all call requests have a backup lightpath and the majority of call requests has set $\mathcal{U}^{(\cdot)}$ empty, i.e., SPP-eDiR coincides with the conventional SPP. On the other hand, when the availability constraint r is greater than 0.006, then set $\mathcal{U}^{(\cdot)}$ coincides with set $\mathcal{W}^{(\cdot)}$, i.e., all the network elements used by the working lightpaths are unprotected for all demands. In the latter case, all demands meet the availability requirement without a backup lightpath.

Qualitatively, the same phenomenon is experienced in all the examined network

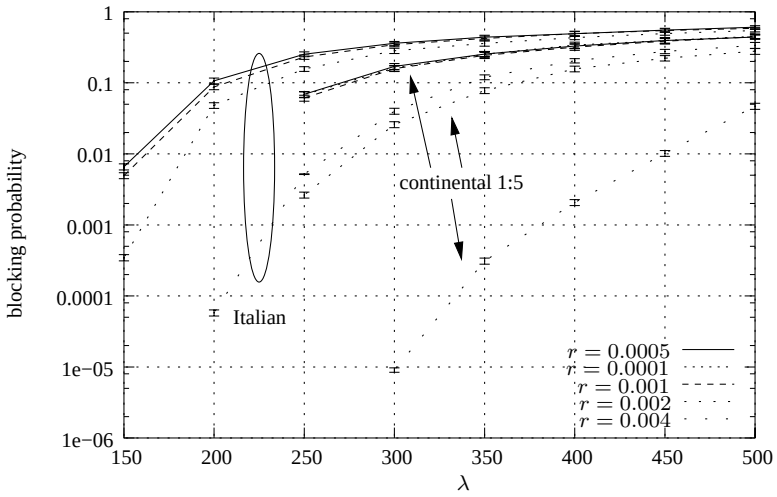


Figure 4.4: Blocking in the national networks using B/MOEMS switches with $W = 32$ as a function of network load and r

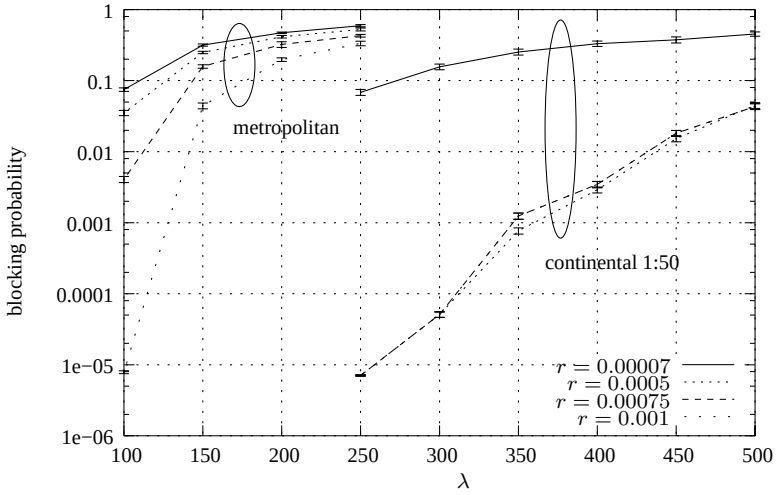


Figure 4.5: Blocking in the metropolitan networks using B/MOEMS switches with $W = 32$ as a function of network load and r

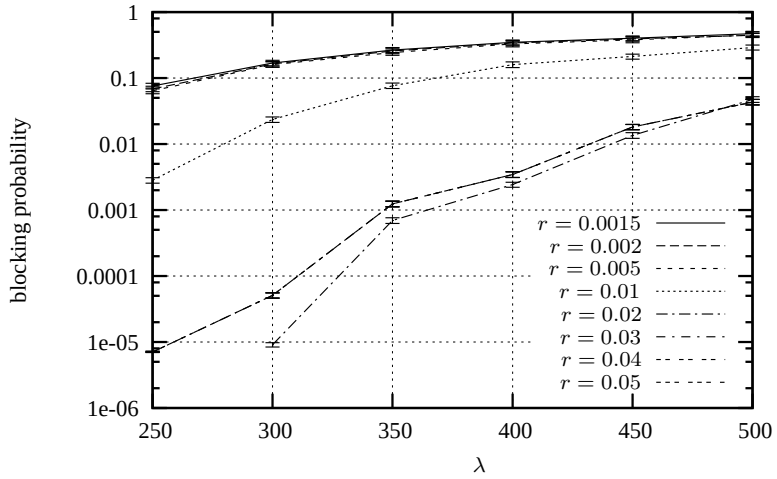


Figure 4.6: Blocking in the continental network using B/MOEMS switches with $W = 32$ wavelengths as a function of network load

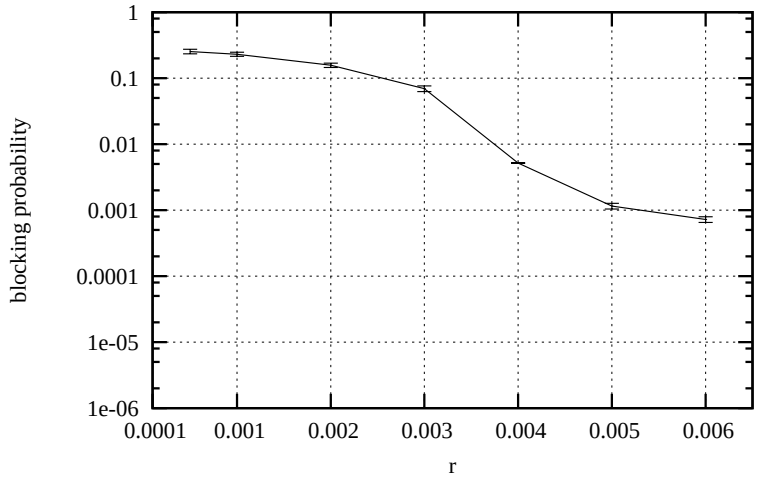


Figure 4.7: Blocking as a function of r at $\lambda = 250$ in the national network

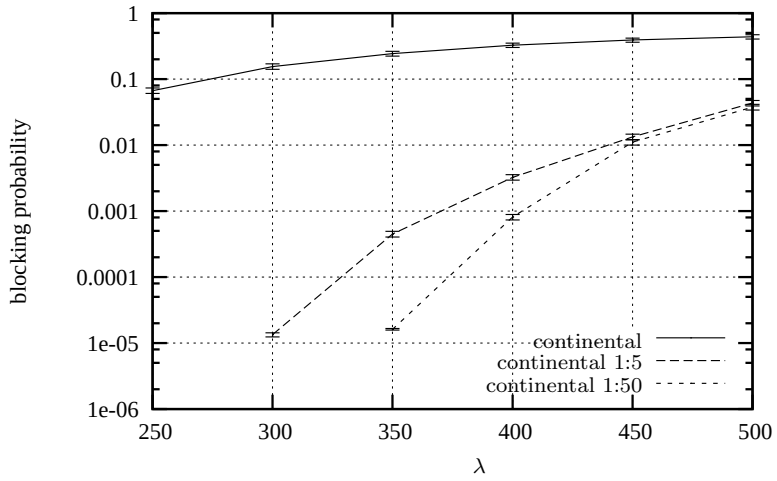


Figure 4.8: Blocking at $r = 0.005$ in different networks using B/InP switches with $W = 32$ wavelengths as a function of network load

scenarios. Note, however, that the respective ranges of r are different, as the value of the asymptotic availability of links is a function of the link length.

Effect of link availability

Figures 4.8 illustrates the performance of the same switching equipment in the three different network scenarios with the same maximum asymptotic connection unavailability r, when the node equipment uses InP technology in a redundant architecture (node architecture B). Note that in order to ease comparison over the same load range Figure 4.8 displays results for the scaled versions of the continental topology.

The results confirm the claim that fulfilling the same availability requirement is harder when component availability is lower. The larger the span lengths are, the less available the links are. Consequently, a higher amount of call requests needs a backup lightpath, which results in a higher consumption of network resources, and higher blocking is experienced. The difference in blocking probability may reach more than four orders of magnitude. This also indicates that while for networks of smaller scale SPP-eDiR could provide reasonably high connection availability, in networks of continental scale different resilience mechanisms should be considered when high connection availability must be guaranteed, e.g. [36, 97].

Effect of node equipment choice

Figure 4.9 reports results obtained for the continental network using various technologies and 4 wavelengths per fiber, when the availability requirement is set to $r = 0.002$. The

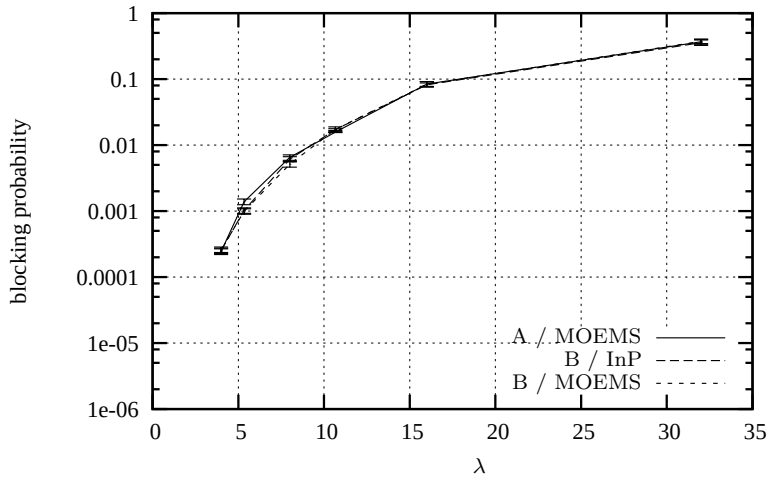


Figure 4.9: Blocking at $r = 0.002$ in the continental network using various switches with $W = 4$ wavelengths as a function of network load

results reveal that there is minimal difference in terms of blocking probability among the different optical switching element technologies and node architectures. This is due to the fact that the link unavailability dominates end-to-end call request unavailability in this network.

Figure 4.10 reports results obtained for the continental 1:5 network using various optical switching element technologies and 32 wavelengths per fiber, when the availability requirement is set to $r = 0.002$. The results demonstrate that in this network link and node availabilities are in comparable orders of magnitude. Therefore, using component technologies and node architectures characterized by higher availability significantly affects network performance in terms of blocking probability.

Figure 4.11 shows the results obtained for the continental 1:50 network using various technologies and 32 wavelengths per fiber, when the availability requirement is set to $r = 0.0005$. In this set of results end-to-end connection unavailability is dominated by node unavailability. For this reason the gain in blocking probability can be as high as two orders of magnitude when switching equipment technologies and node architectures characterized by higher availability are used.

Note that in Figures 4.9, 4.10, and 4.11 the curves for switching equipment based on InP technology in a non-redundant architecture is not shown. The reason for this is that the requested availability cannot be fulfilled with this type of node equipment.

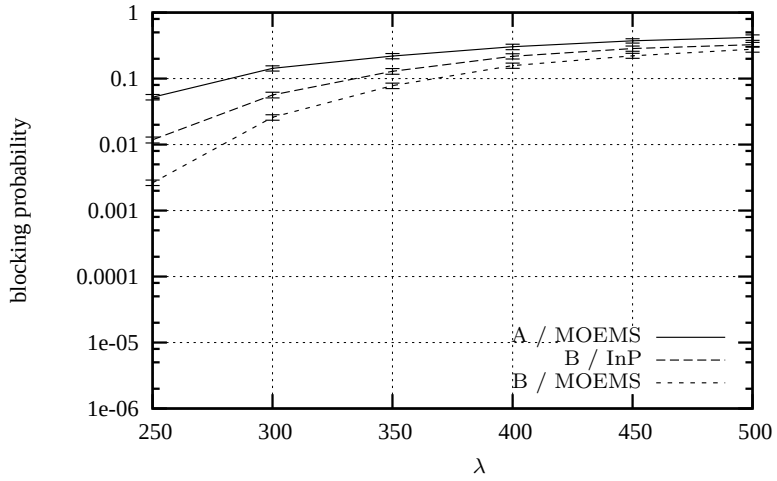


Figure 4.10: Blocking at $r = 0.002$ in the continental 1:5 network using various switches with $W = 32$ wavelengths as a function of network load

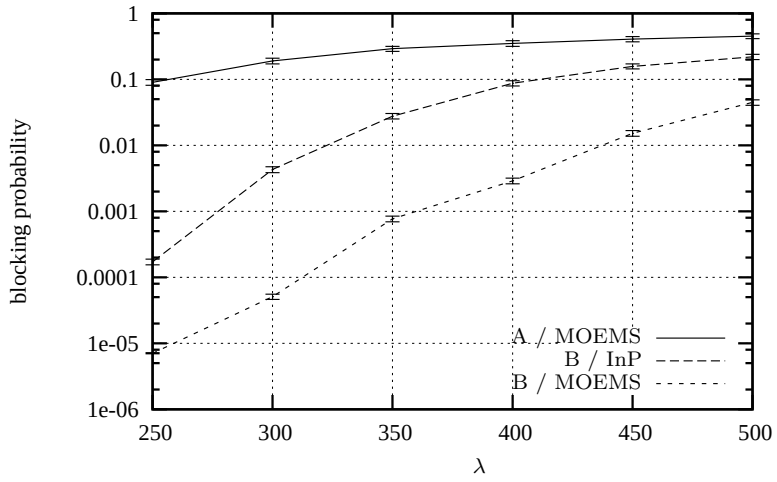


Figure 4.11: Blocking at $r = 0.0005$ in the continental 1:50 network using various switches with $W = 32$ wavelengths as a function of network load

Networking scenario	**Node architecture, and technology**	W	$r_{\min}$ **SPP-eDiR**	**DPP**
continental	A / InP	4	0.00451931	0.00379703
national	A / InP	4	0.00388749	0.00376362
metropolitan	A / InP	4	0.00116495	0.00115868
continental	A / MOEMS	4	0.00150328	0.00035469
national	A / MOEMS	4	0.00015567	0.00007596
metropolitan	A / MOEMS	4	0.00005305	0.00005022
continental	B / InP	4	0.00144367	0.00030381
national	B / InP	4	0.00010381	0.00002668
metropolitan	B / InP	4	0.00000467	0.00000200
continental	B / MOEMS	4	0.00144267	0.00030357
national	B / MOEMS	4	0.00010361	0.00002640
metropolitan	B / MOEMS	4	0.00000462	0.00000195
continental	A / InP	32	0.033478	0.0301384
national	A / InP	32	0.0318543	0.0302193
metropolitan	A / InP	32	0.00944613	0.00938901
continental	A / MOEMS	32	0.00193676	0.00071418
national	A / MOEMS	32	0.00052464	0.00042491
metropolitan	A / MOEMS	32	0.00039289	0.00038871
continental	B / InP	32	0.00149888	0.00031682
national	B / InP	32	0.00027903	0.00023466
metropolitan	B / InP	32	0.00002545	0.00002285
continental	B / MOEMS	32	0.00144629	0.00030422
national	B / MOEMS	32	0.00010471	0.00002671
metropolitan	B / MOEMS	32	0.00000478	0.00000203
continental	A / InP	64	0.0681643	0.0601911
national	A / InP	64	0.0657363	0.0607118
metropolitan	A / InP	64	0.0192178	0.0190553
continental	A / MOEMS	64	0.00243609	0.00112693
national	A / MOEMS	64	0.00095150	0.00082669
metropolitan	A / MOEMS	64	0.00078244	0.00077666
continental	B / InP	64	0.00165265	0.00090066
national	B / InP	64	0.00092732	0.00087357
metropolitan	B / InP	64	0.00008723	0.00008437
continental	B / MOEMS	64	0.00145119	0.00030515
national	B / MOEMS	64	0.00010624	0.00002719
metropolitan	B / MOEMS	64	0.00000506	0.00000221

Table 4.3: Best feasible availability constraints

Comparison with dedicated path protection

In order to further demonstrate that the applied polynomial availability estimation method yields satisfactory bounds the best feasible availability guarantees of SPP-eDiR are compared against those of another well-known protection scheme, dedicated path protection (DPP).

Differently from SPP, DPP does not allow sharing of backup resources, that is, backup resources along a backup lightpath are dedicated to the corresponding working lightpath. With DPP the evaluation of equation (4.3) is simple, as the term associated with the probability that the backup lightpath is not available is affected only by the unavailability of network elements used to route the backup lightpath ($q_{\mathcal{B}}^{(d)} \equiv q_{1+}^{(\mathcal{B}^{(d)})}$). Once equation (4.3) is evaluated, $r_{\min}$ can be calculated the same way as in case of SPP-eDiR described next.

Taking advantage of the fact that the optimization method applied in the experiments makes use of a predetermined set of candidate lightpaths (c.f. Appendix B) it is possible to compute a bound on the best availability constraint that can be offered by SPP-eDiR to all of the call-requests using equation (4.4). The bound is obtained by setting $\mathcal{U}^{(d)} = \emptyset$. Note that with this assumption SPP-eDiR coincides with conventional SPP. Next, for each source-destination pair the working/backup lightpath pair with the highest availability is found. Then, among all source-destination pairs the one with the least available working/backup lightpath pair is the one that gives the value of the bound. Formally:

$$r_{\min} = \max_{d} \big(1 - \max_{L_{\mathrm{w}}^{(d)}, N_{\mathrm{w}}^{(d)}, L_{\mathrm{b}}^{(d)}, N_{\mathrm{b}}^{(d)}} \check{p}^{(d)}\big). \tag{4.7}$$

Table 4.3 contains the values of $r_{\min}$ for numerous combinations of network size and node equipment type.

The value of $r_{\min}$ of DPP is obviously a lower bound on $r_{\min}$ for all the protection schemes relying on a single backup lightpath per connection. When comparing it to $r_{\min}$ of SPP the surprising observation can be made that by means of the proposed availability approximation the lowest feasible availability constraint is at most five times the absolute lower bound. In general, the values of $r_{\min}$ for DPP and SPP are the closest to each other in the metropolitan scenarios for a given node equipment.

Thus it is confirmed that the proposed availability approximation method yields adequate results, while the computation complexity required to verify and maintain the availability guarantees is acceptable.

Also note that the values of blocking probability obtained for DPP will always be as high as that of either SPP or SPP-eDiR for the same network load and availability requirement.

Not surprisingly, the MOEMS technology, being built on components with higher

reliability performance, always performs better in terms of achievable guarantees than its counterpart. The difference measured in terms of $r_{\min}$ can be as large as two orders of magnitude when the non-redundant node architecture is applied (Figure 4.1). This gap can be decreased significantly by introducing redundancy (Figure 4.2). In the case of the continental network the unavailability of links is so high that it diminishes the difference among the various node equipment types. As a consequence, non-redundant MOEMS based switching equipment may deliver almost the same performance as that of the equipment of much higher reliability performance with built-in redundancy. It appears that at this network scale it is not reasonable to invest in making node equipment of extremely high reliability performance. On the other hand, in a metropolitan scenario a redundant node architecture might increase feasible availability guarantees by as much as two orders of magnitude with respect to the non-redundant case. Here, the gain is significant even for the MOEMS based switching equipment of higher availability.

Note that even though increasing the number of wavelengths per fiber (W) decreases the achievable guarantees, it accommodates a higher number of simultaneous call requests (higher values of λ). It is also interesting to observe how the values of $r_{\min}$ change as the network scenario changes. For highly reliable node equipment sharp drops are experienced when link spans are of smaller scale. Nevertheless, the least reliable equipment perform almost "equally badly" under all studied circumstances.

4.5 Summary

The availability requirements of applications must be met with sufficient circuit redundancy in the network. In general, a good objective is to guarantee the required availability with the minimum redundancy possible. This objective is likely to yield efficient utilization of network resources and increased network throughput.

With this objective in mind, the chapter investigated the impact of the OXC equipment choice on the blocking probability in dynamic circuit-switched WDM networks that do not make use of wavelength converters. A framework was presented in which state-of-the-art results on optical switching fabric reliability performance are combined with routing and wavelength assignment optimization. The application's availability requirement is taken into account and multiple simultaneous network failures are assumed to be possible. The framework permits to bound circuit availability in the WDM network when shared path protection switching and differentiated availability techniques are jointly used at the optical layer.

The obtained results confirm that in some cases the widely used assumption of negligible node failures is acceptable. In other cases, however, they reveal that this assumption is not acceptable at all. In the latter case the selection of both the OXC architecture and switching fabric technology is driven by the availability requirements, in addition

to other conventional metrics, e.g., cost, scalability, etc.

For this reason, a number of OXC architectures with and without built-in redundancy was considered in the study. The considered OXC's are representative of two switching technologies: planar tilt mirror type MOEMS, and indium phosphide semiconductor optical amplifier. The continental, national and metropolitan network scales considered in the study provide some general guidelines as to when the choice of the OXC may be driven by availability metrics.

4.6 Proof of NP-Completeness of SPP-DiR

The authors in [69] define the dynamic shared-path protected lightpath provisioning (DSPLP) problem and prove that it is NP-Complete. Based on this proof it is easy to see that introducing DiR and node failures do not make the problem easier. The basic idea of the proof presented here is the following. First, the SPP-DiR problem is shown to be in NP. The DSPLP problem is then reduced to the SPP-DiR problem, that is, a polynomial time transformation is shown that assigns a problem instance of SPP-DiR to any arbitrary instance of DSPLP. Next, the DSPLP problem instance is shown to have a solution if and only if the SPP-DiR problem instance has one. The NP-Completeness of SPP-DiR immediately follows.

Let us define the problem of dynamic shared-path protected lightpath provisioning with DiR (SPP-DiR) as follows. Let $\mathcal{L} = \{l^{(i)}\}$ be the set of existing lightpaths, where the i^{th} lightpath $l^{(i)}$ is characterized by $(L_{\text{w}}^{(i)}, N_{\text{w}}^{(i)}, L_{\text{b}}^{(i)}, N_{\text{b}}^{(i)}, L_{\text{p}}^{(i)}, N_{\text{p}}^{(i)}, t_{\text{a}}^{(i)}, t_{\text{h}}^{(i)})$ as defined in Section 4.2.

Let $F(r, l, \mathcal{L})$ be the availability check function that returns 1 if the working and backup lightpaths meet the availability requirement r of the lightpath request with respect to the set of protected links and the existing lightpaths and 0 if they fail to meet it. We require that $F(r, l, \mathcal{L})$ be computed in polynomial time[2].

Let $l^{(*)}$ be an incoming lightpath request. The RWA solution of the SPP-DiR problem must satisfy the following constraints:

C.1 $L_{\text{w}}^{(*)}$ and $L_{\text{b}}^{(*)}$ are disjoint, and $N_{\text{p}}^{(*)}$ and $N_{\text{b}}^{(*)}$ are disjoint,

C.2 $L_{\text{w}}^{(*)}$ and $L_{\text{w}}^{(i)}$, $1 \leq i \leq |\mathcal{L}|$, do not utilize the same wavelength on any common link they traverse,

C.3 $L_{\text{w}}^{(*)}$ does not share any wavelength with $L_{\text{b}}^{(i)}$, $1 \leq i \leq |\mathcal{L}|$ on any common link they traverse,

C.4 $L_{\text{b}}^{(*)}$ and $L_{\text{b}}^{(i)}$ can share a wavelength on a common link only if (4.1) is satisfied,

[2]Note that $F(\cdot)$ is not required to give exact results, e.g., a decision based on an approximate but conservative bound suffices.

C.5 $F(r^{(i)}, l^{(i)}, \mathcal{L} \cup \{l^{(*)}\}) = 1$, $1 \leq i \leq |\mathcal{L}|$ and $F(r^{(*)}, l^{(*)}, \mathcal{L} \cup \{l^{(*)}\}) = 1$[3].

Instance: graph $G(V, E, W)$, where $W(e)$ is the number of wavelengths on link $e \in E$, the set of existing lightpaths $\mathcal{L}$, the availability check function $F(r, l, \mathcal{L})$ and a lightpath request $l^{(*)}$ from $n_{\mathrm{s}}^{(*)} \in V$ to $n_{\mathrm{d}}^{(*)} \in V$ with an availability requirement $r^{(*)}$.

Question: Do there exist from $n_{\mathrm{s}}^{(*)}$ to $n_{\mathrm{d}}^{(*)}$ two lightpaths $(L_{\mathrm{w}}^{(*)}, N_{\mathrm{w}}^{(*)})$ and $(L_{\mathrm{b}}^{(*)}, N_{\mathrm{b}}^{(*)})$ and a protection assignment $(L_{\mathrm{p}}^{(*)}, N_{\mathrm{p}}^{(*)})$ such that they satisfy the SPP-DiR constraints with respect to the existing lightpaths?

Theorem 1. *SPP-DiR is NP-Complete.*

Proof. SPP-DiR $\in$ NP, since a nondeterministic algorithm can guess $(L_{\mathrm{w}}^{(*)}, N_{\mathrm{w}}^{(*)})$, $(L_{\mathrm{b}}^{(*)}, N_{\mathrm{b}}^{(*)})$, and $(L_{\mathrm{p}}^{(*)}, N_{\mathrm{p}}^{(*)})$ and verify in polynomial time whether constraints C.1-C.5 are satisfied.

The DSPLP problem was shown to be NP-Complete in [69]; therefore, it is enough to show a polynomial time algorithm to transform an arbitrary instance of DSPLP to an instance of SPP-DiR such that the DSPLP problem can be satisfied if and only if the corresponding SPP-DiR problem can be satisfied.

A DSPLP problem is defined as $G = (V, E, C, W)$, $\mathcal{L}$ and $s, d \in V$. Note that the C cost function of the DSPLP problem is irrelevant with respect to the constraints that define the problem, hence it will be omitted in this proof. To an instance of the DSPLP problem an SPP-DiR problem defined as $G' = (V, E, W)$, $n_{\mathrm{s}}^{(*)} = s$, $n_{\mathrm{d}}^{(*)} = d$ is assigned.

The elements of $\mathcal{L}'$ can be obtained as follows: $((L_{\mathrm{w}}^{(i)}, n(L_{\mathrm{w}}^{(i)}), L_{\mathrm{b}}^{(i)}, n(L_{\mathrm{b}}^{(i)}), L_{\mathrm{p}}^{(i)}, \emptyset, t_{\mathrm{a}}^{(i)}, t_{\mathrm{h}}^{(i)}) | (L_{\mathrm{w}}^{(i)}, L_{\mathrm{b}}^{(i)}, t_{\mathrm{a}}^{(i)}, t_{\mathrm{h}}^{(i)}) \in \mathcal{L})$, where $n(L_{\mathrm{w}}^{(i)})$ and $n(L_{\mathrm{b}}^{(i)})$ are the set of nodes incident to the edges of $L_{\mathrm{w}}^{(i)}$ and $L_{\mathrm{b}}^{(i)}$, respectively. That is, each element of $\mathcal{L}'$ corresponds to an element in $\mathcal{L}$ so that the link related parameters are identical and the node related ones are derived from the link related ones except for the set of protected nodes, which is empty. Let $F(r^{(i)}, l^{(i)}, \mathcal{L})$ return 1 if and only if $L_{\mathrm{p}}^{(i)} = L_{\mathrm{w}}^{(i)}$ and $N_{\mathrm{p}} = \emptyset$ and 0 otherwise.

It is trivial to see that whenever there exists an $L_{\mathrm{w}}^{(*)}, L_{\mathrm{b}}^{(*)}$ lightpath pair between s and d, then lightpath pair $(L_{\mathrm{w}}^{(*)}, n(L_{\mathrm{w}}^{(*)}))$, $(L_{\mathrm{b}}^{(*)}, n(L_{\mathrm{b}}^{(*)}))$ with $L_{\mathrm{p}}^{(*)} = L_{\mathrm{w}}^{(*)}$ and $N_{\mathrm{p}}^{(*)} = \emptyset$ will satisfy the constraints of the corresponding SPP-DiR problem as well.

It is also trivial to see that if there exists an $(L_{\mathrm{w}}^{(*)}, N_{\mathrm{w}}^{(*)})$, $(L_{\mathrm{b}}^{(*)}, N_{\mathrm{b}}^{(*)})$ lightpath pair with a protection assignment $(L_{\mathrm{p}}^{(*)}, N_{\mathrm{p}}^{(*)})$ between $n_{\mathrm{s}}^{(*)}$ and $n_{\mathrm{d}}^{(*)}$ then lightpath pair $(L_{\mathrm{w}}^{(*)}, L_{\mathrm{b}}^{(*)})$ will satisfy the constraints of the corresponding DSPLP problem, as well.

Consequently, the SPP-DiR problem is NP-Complete. □

[3]Note that if (4.4) is used for the estimation in $F(\cdot)$, then it suffices to check $F(r^{(*)}, l^{(*)}, \mathcal{L} \cup \{l^{(*)}\}) = 1$ only.

Chapter 5

A Threshold Based Algorithm for Higher Availability Guarantees

When availability is a major concern in optical WDM networks, shared (backup) path protection (S(B)PP) schemes offer the potentially appealing feature of requiring fewer network resources than their counterpart dedicated path protection (DPP) schemes. However, while backup resource sharing increases resource utilization, it reduces end-to-end availability of connections. Moreover, the evaluation of connection demand availability is easy with DPP, i.e., it depends only on the components used to route the connection demand. In contrast to this, it is more complex with SPP, as sharing of protection resources introduces dependencies among different connection demands, and the availability of a connection depends also on the failure of components that are not used to route the connection demand. The possible occurrence of multiple failures exacerbates this problem significantly. For this reason, the simpler DPP schemes are often chosen, even though, as a result, network resources are not efficiently used.

This chapter presents a simple way to determine routing and wavelength assignment for connection demands subject to availability constraints using the SPP scheme [77]. The solution is based on controlling the amount of sharing that can be done on spare resources. Controlling the amount of sharing is based on a threshold, assigned to the spare resource, which allows one to quickly determine whether or not (additional) sharing is permitted, i.e., the required availability level of both the newly arrived demand and the already established demands are all guaranteed. When adopting the proposed solution the challenge is to select the threshold value that yields a good utilization of the network resources. When such value can be found, the proposed solution offers the possibility to achieve network utilization values that are possible only by means of SPP schemes, while at the same time requiring a simple computational technique that is comparable to that of DPP schemes.

For the ease of discussion this chapter assumes that only links in the network are failure-prone. This assumption may be dropped and the introduced method may be extended to include node failures, as well.

5.1 Related work

Quality of Service (QoS) awareness gained vital importance in service provisioning with the roll-out of applications that impose quality requirements on data transfer. Fulfilling these requirements also necessitates that the underlying networking technology is capable of offering end-to-end transport service at different availability levels. Availability can be improved by providing protection, that is, providing connections that require high availability with additional standby resources that can be readily activated and used in case of a failure of one or more components. Protection schemes can utilize dedicated resources, e.g., dedicated path protection (DPP) [109], or shared resources, e.g., shared (backup) path protection (SPP) [109]. There is a clear tradeoff between dedicated protection schemes and shared protection schemes. As a general rule of thumb, dedicated protection requires more resources, but in turn it provides higher availability. On the other hand, shared protection is more resource efficient but it provides lower availability [12].

Both dedicated and shared protection schemes, have been an active area of research. [86] discusses the problem of dual-failure resilient design for p-cycles and presents an approximation technique that improves survivability with respect to double failures. However, possibly different end-to-end availability requirements are not taken into account during the computations. [22] argues that in some cases single-failure protection designs provide adequate protection even against double failures. Moreover, the paper also shows that by limiting the extent of resource sharing dual-failure resilience can be improved. An intuitive explanation for the first claim is that a connection is only disrupted by a small fraction of the total possible failures of higher multiplicity. The second claim can be explained as follows: limiting sharing decreases the interference among different connections, so they will not block each other when trying to survive failures. The observations in [22] are made based on an analysis executed on designs optimized for single-failure and dual-failure robustness; however, quantitative availability guarantees are not addressed on a per connection basis.

The papers cited above, like most of the related literature, address availability issues from the failure state-based requirement point of view, for example, the discussed solutions provide protection against any single/double link failure. [97] introduces a technique that is, in principle, capable of providing a design that is robust in the presence of failures of any multiplicity and also addresses probabilistic demand survivability requirements, in other words, satisfies each demand's Maximum Downtime Ratio (MDR). The drawback is that the complexity of the necessary computations increases with the failure multiplicity. Another technique that is capable of addressing connection availability guarantees is presented in [5, 102], also assuming a network dimensioning context.

The amount of necessary computations is especially critical if an on-line Routing and

Wavelength Assignment (RWA) algorithm has to be operated. The exact computation of connection availability is difficult in case of protection schemes with backup resource sharing due to the dependencies among multiple connections that the sharing introduces (c.f. section 2.4.3). In most cases, when the knowledge and control of the exact connection availability is required, dedicated protection schemes are used due to the much simpler computation of the connection availability, i.e., a connection is failed if both the working lightpath and the protection lightpath are non-operational, which corresponds to the failure of at least one link on the working lightpath and at least one link on the protection lightpath. In case of DPP, links, or more generally network elements that are associated neither with the working nor with the protection lightpath do not affect connection availability.

In what follows a simple method is proposed to use SPP, when the connection availability is a requirement that needs to be satisfied. To reduce the amount of computation that an on-line RWA algorithm has to perform, instead of calculating the exact value of the availability for each connection, a conservative bound is used. The bound is based on a threshold value that controls the amount of sharing on each protection resource. Thus connection availability is guaranteed to individual connections according to their requirements. As demonstrated by the simulation results, the presented approach greatly simplifies computations done by the on-line RWA algorithm while keeping resource utilization highly efficient.

5.2 Problem statement

The following assumptions and notations are adhered to. The $G(V, E)$ network consists of failure prone links and always operating nodes[1]. Links can have two states, they are either operating or failed. Link failures are assumed to be statistically independent, and the link failure probability q_e is known for all links $e \in E$. Each link $e \in E$ corresponds to a pair of fibers, one for each direction of propagation. Without loss of generality, links are assumed to have W wavelength channels on each fiber.

Demands, denoted by $d = (n_{\mathrm{s}}^{(d)}, n_{\mathrm{d}}^{(d)}, t_{\mathrm{a}}^{(d)}, t_{\mathrm{h}}^{(d)}, r^{(d)})$ dynamically enter and leave the network. Each demand d requires one wavelength channel between a source node $n_{\mathrm{s}}^{(d)} \in V$ and a destination node $n_{\mathrm{d}}^{(d)} \in V$. $0 < t_{\mathrm{a}}$ is the time of arrival of the call request, while $0 < t_{\mathrm{h}}^{(d)}$ is the duration of time for which the connection needs network resources. The fifth parameter, $0 < r^{(d)} < 1$ is the availability requirement, i.e., the maximum accepted probability that demand d is disrupted by a failure of any multiplicity in the network.

The objective of an on-line RWA algorithm is to assign resources to the arriving demands so that their availability requirement $r^{(\cdot)}$ is satisfied. For an incoming demand

[1]Extension to the case where nodes are also failure prone is straightforward [75].

d first a wavelength channel, or lightpath has to be established between $n_{\mathrm{s}}^{(d)}$ and $n_{\mathrm{d}}^{(d)}$. This is called the working lightpath for demand d. Let $L_{\mathrm{w}}^{(d)}$ be the ordered set of links used by the working lightpath of demand d. Since link failures are assumed to be independent, the probability that the working lightpath is disrupted ($q_{L_{\mathrm{w}}}^{(d)}$) can be expressed as:

$$q_{L_{\mathrm{w}}}^{(d)} = 1 - \prod_{e \in L_{\mathrm{w}}^{(d)}} (1 - q_e). \tag{5.1}$$

If $q_{L_{\mathrm{w}}}^{(d)} \leq r^{(d)}$, the working lightpath itself provides a connection that is of sufficient availability. Otherwise, additional resources have to be allocated to provide one (or more) alternative path to the destination in case the working lightpath fails. Backup resources are not used when the working lightpath is operating. Therefore, resource efficiency can be greatly increased by sharing backup resources among multiple demands.

In particular, the SPP scheme has been selected to take advantage of the increased resource efficiency, while guaranteeing availability. With SPP, if $q_{L_{\mathrm{w}}}^{(d)} > r^{(d)}$, another lightpath is assigned to the connection from $n_{\mathrm{s}}^{(d)}$ to $n_{\mathrm{d}}^{(d)}$ called the backup lightpath. Let $L_{\mathrm{b}}^{(d)}$ be the ordered set of links used by the backup lightpath of demand d. The backup lightpath is link disjoint from the working lightpath and is activated only if the working lightpath is disrupted by a failure. Under the often used single failure assumption it can be easily seen, that among all demands sharing a common protection resource at most one demand will attempt to activate it, if the corresponding working paths are route disjoint. Formally, backup lightpaths of demands d_1 and d_2 may share the same protection resource, i.e., wavelegth w on link e, if $L_{\mathrm{w}}^{(d_1)} \cap L_{\mathrm{w}}^{(d_2)} = \emptyset$.

When the probability of failures affecting multiple components is not negligible, the above condition does not lead to the guarantee that at most one connection will attempt to activate a backup resource. As a matter of fact, it is not possible to enforce a condition that guarantees that, when considering all possible failure scenarios, two connections sharing a common resource will not attempt to activate the same shared backup resource as a result of a failure. As a consequence, the calculation of the availability of each active demand becomes more difficult as illustrated in the example on Figure 2.2 in section 2.4.3.

The presented example can be generalized and stated formally using the following notation. Let $b = (e, w)$ denote a shared backup resource — i.e., wavelength $1 \leq w \leq W$ on link $e \in E$. Let $D^{(b)}(t)$ be the set of active demands d_i whose backup path $L_{\mathrm{b}}^{(d_i)}$ is routed on link $e \in E$ and uses wavelength w at time t. When deciding whether a new demand d may share backup resources with other demands the on-line RWA algorithm using SPP has to check the following three criteria:

1. $L_{\mathrm{w}}^{(d)} \cap L_{\mathrm{w}}^{(d_i)} = \emptyset$ for each $d_i \in D^{(b)}(t_{\mathrm{a}}^{(d)})$ for each link e used by $L_{\mathrm{b}}^{(d)}$,

2. admitting demand d does not violate the availability requirements of other connections, i.e., $q^{(d_i)} \leq r^{(d_i)}$ for each d_i in $D^{(b)}(t_{\mathrm{a}}^{(d)})$ and for each link e used by $L_{\mathrm{b}}^{(d)}$,

3. the availability requirement for connection demand d is satisfied, that is, $q^{(d)} \leq r^{(d)}$ with respect to the effect of sharing b with connections in $D^{(b)}(t_{\mathrm{a}}^{(d)})$ for each link e used by $L_{\mathrm{b}}^{(d)}$.

When failures of multiple components are not negligible, the exact computation of 2) and 3) is time consuming hence not feasible for a fast on-line RWA algorithm. The next section discusses a conservative bounding technique, which facilitates quick assessment of connection availability and is, therefore, suitable for on-line computations.

5.2.1 Sharing unavailability

We define the concept of *sharing unavailability* as follows. A shared protection resource, i.e., a wavelength channel on a link will be activated when certain failures occur. The sharing unavailability $u^{(e,w)}(t,d)$ of wavelength channel $1 \leq w \leq W$ on link $e \in E$ with respect to demand d is the total probability of having a failure affecting demands other than demand d that leads to the activation of resource $b = (e,w)$ at time t. In other words, it is the increase in unavailability of resource b from the viewpoint of demand d that derives from the fact that b is shared with connection demands $d_i \in D^{(b)}(t)$.

Using $u^{(b)}(t,d)$, the failure probability of the backup path of demand d, $q_{L_{\mathrm{b}}}^{(d)}(t)$ can be upper bounded by:

$$q_{L_{\mathrm{b}}}^{(d)}(t) \leq 1 - \prod_{e \in L_{\mathrm{b}}^{(d)}} \big(1 - (q_e + u^{(e,w)}(d,t))\big). \tag{5.2}$$

Note that formula (5.2) is an upper bound on the failure probability because it considers the sharing unavailabilities $u^{(e,w)}(t,d)$ on different links to be independent. If two backup lightpaths share more than one backup resource, the sharing unavailability $u^{(e,w)}(t,d)$ is taken into account more than once at the demand level, thus overestimating the probability that backup resources are activated by other connections.

The exact computation of the sharing unavailability is still a difficult task, but an upper bound can be easily obtained as follows. A backup resource b is shared among demands in $D^{(b)}(t)$. Let

$$S_{\mathrm{w}}^{(b)}(t) = \bigcup_{d_i \in D^{(b)}(t)} L_{\mathrm{w}}^{(d_i)}$$

be the set of links that are used by the working lightpaths of all connection demands $d_i \in D^{(b)}(t)$. Let $S_{\mathrm{w}}^{(b)}(t,d)$ be $S_{\mathrm{w}}^{(b)}(t) \setminus L_{\mathrm{w}}^{(d)}$, which is the set of links used by the working lightpaths of all connection demands $d_i \in D^{(b)}(t)$ minus the links used by the working

lightpath $L_{\mathrm{w}}^{(d)}$ of connection demand d. Let $D^{(b)}(t,d)$ be $D^{(b)}(t) \setminus \{d\}$, i.e., set $D^{(b)}(t)$ minus connection demand d.

Lemma 6. *$u^{(b)}(t,d)$ is always upper bounded by the probability that any link in $S_w^{(b)}(t,d)$ fails.*

Proof. Since link failures are assumed to be independent, and only links in set $S_{\mathrm{w}}^{(b)}(t,d)$ may trigger the activation of b, the probability that at least one link is failed in set $S_{\mathrm{w}}^{(b)}(t,d)$ might equal to $u^{(b)}(t,d)$. This is the case when, for example, only demands d_1 and d_2 are present in the network, and the single-link backup lightpath of demand d_1 only shares b with the backup lightpath of demand d_2.

However, the failure of links not in $S_{\mathrm{w}}^{(b)}(t,d)$ might prevent the activation of backup resources of connection demands in set $D^{(b)}(t,d)$, due to possible sharing conflicts. As an example, consider Figure 2.2 with $d = A$. The activation of link a is not only a function of $S_{\mathrm{w}}^{(a,w)}(t,A)$ — i.e. the links on the working lightpath of demand C. If a link on the working lightpath of demand B fails first, then due to sharing on link b the protection lightpath of demand C will not be activated regardless of whether there is a subsequent failure on its working lightpath.

As a consequence, the probability that any link in $S_{\mathrm{w}}^{(b)}(t,d)$ is failed becomes an upper bound for the value of $u^{(b)}(t,d)$. □

An additional remark that strengthens the claim of the lemma may be made as follows. Note that, for example, b will only be activated by demand $d_2 \in D^{(b)}(t,d)$ if the rest of the resources used by $L_{\mathrm{b}}^{(d_2)}$ are available. Therefore, the set of links that may prevent the activation of backup resources of connection demands in set $D^{(b)}(t,d)$ may be even larger.

Technically, a set $S_{\mathrm{w}}^{(b)}(t,d_i)$ should be maintained for each shared backup resource b and for each connection $d_i \in D^{(b)}(t)$. The upper bound on the sharing unavailability is then given as

$$u^{(e,w)}(t,d) \leq \hat{u}^{(e,w)}(t,d) = 1 - \prod_{f \in S_{\mathrm{w}}^{(e,w)}(t,d)} (1 - q_f). \tag{5.3}$$

If the value $\hat{u}^{(e,w)}(t,d)$, provided by inequality (5.3), is used in place of the sharing unavailability in equation (5.2) then the expression of the probability that the protection lightpath fails still remains an upper bound.

Note that the presented method is conservative also with respect to failure sequences. The reason for this is that it is implicitly assumed that if a shared protection resource needs to be activated by another demand, that particular demand will always be the first one to activate it.

5.3 On-line RWA algorithm with availability guarantees

If an on-line RWA algorithm can control $u^{(e,w)}(t,d)$, the sharing unavailability for every demand then availability computations become much easier. To facilitate this, a threshold $q_{\mathrm{s}}^{(e,w)}(d)$ is introduced for each link $e \in E$, wavelength w and connection d. If the on-line RWA algorithm enforces that $u^{(e,w)}(t,d) \leq q_{\mathrm{s}}^{(e,w)}(d)$, then

1. $q_{\mathrm{s}}^{(e,w)}(d)$ appears to be a limit that determines the extent to which different demands may re-use the same resource, and
2. $q_{\mathrm{s}}^{(e,w)}(d)$ also guarantees that sharing unavailability of resources is always upper bounded, and, therefore, future connection demands will not decrease availability of already admitted connection demands below their requirements. In other words, time dependence of backup resource availability is removed from the computations.

This threshold will then appear as an addition to the failure probability of backup links in the computations so that the probability that the connection of demand d is disrupted ($q^{(d)}$) is upper bounded as:

$$\begin{aligned} q^{(d)} &= q_{L_{\mathrm{w}}}^{(d)} q_{L_{\mathrm{b}}}^{(d)} \leq \hat{q}^{(d)}, \\ \hat{q}^{(d)} &= q_{L_{\mathrm{w}}}^{(d)} \hat{q}_{L_{\mathrm{b}}}^{(d)}, \\ \hat{q}_{L_{\mathrm{b}}}^{(d)} &= 1 - \prod_{e \in L_{\mathrm{b}}^{(d)}} \left(1 - (q_e + q_{\mathrm{s}}^{(e,w)}(d))\right). \end{aligned} \tag{5.4}$$

An on-line RWA algorithm may then decide whether demand d can be served using $L_{\mathrm{w}}^{(d)}$ and $L_{\mathrm{b}}^{(d)}$ (if necessary) by checking the following criteria:

1. $L_{\mathrm{w}}^{(d)} \cap S_{\mathrm{w}}^{(e,w)}(t_{\mathrm{a}}^{(d)}) = \emptyset$ for each $e \in L_{\mathrm{b}}^{(d)}$ and the assigned protection wavelength w,
2. $\hat{q}^{(d)} \leq r^{(d)}$,
3. $\hat{u}^{(e,w)}(t_{\mathrm{a}}^{(d)}, d_i) \leq q_{\mathrm{s}}^{(e,w)}(d_i)$ for each shared resource (e,w) used by the protection lightpath $L_{\mathrm{b}}^{(d)}$ of connection demand d and each demand $d_i \in D^{(e,w)}(t_{\mathrm{a}}^{(d)}, d)$.

The first criterion enforces that no shared resource might be used by two different demands for protection against the failure of the same component. The second one verifies that the survivability requirement of the demand to be admitted is met. The third one checks that the unavailability threshold is not violated anywhere. With respect to the latter the upper bounds can be determined based on the current state of the network, while the threshold is considered to be an input to the RWA algorithm. Algorithm 3 shows the pseudocode.

The provisioning method discussed in this section will be referred to as the sharing unavailability threshold based method, or ShUT method for short. Further details of the implementation of this provisioning method are discussed in section 5.6.

Algorithm 3 Sharing unavailability threshold based RWA check algorithm

procedure SHUT_RWA_CHECK($d, q_{\mathrm{s}}^{(\cdot)}(\cdot), S_{\mathrm{w}}^{(\cdot)}(t_{\mathrm{a}}^{(d)}), D^{(\cdot)}(t_{\mathrm{a}}^{(d)}), L_{\mathrm{w}}^{(d)}, w_{\mathrm{w}}, L_{\mathrm{b}}^{(d)}, w_{\mathrm{b}}$)
  $accept = true$
  for all $e \in L_{\mathrm{w}}^{(d)}$ **do**
    if (e, w_{w}) is already used **then**
      $accept = false$
    end if
  end for
  for all $e \in L_{\mathrm{b}}^{(d)}$ **do**
    if $L_{\mathrm{w}}^{(d)} \cap S_{\mathrm{w}}^{(e,w_{\mathrm{b}})}(t_{\mathrm{a}}^{(d)}) \neq \emptyset$ **then**
      $accept = false$
    end if
  end for
  if $\hat{q}^{(d)} > r^{(d)}$ **then**
    $accept = false$
  end if
  for all $e \in L_{\mathrm{b}}^{(d)}$ **do**
    for all $d_i \in D^{(e,w_{\mathrm{b}})}(t_{\mathrm{a}}^{(d)}, d)$ **do**
      if $\hat{u}^{(e,w)}(t_{\mathrm{a}}^{(d)}, d_i) > q_{\mathrm{s}}^{(e,w_{\mathrm{b}})}(d_i)$ **then**
        $accept = false$
      end if
    end for
  end for
  return *accept*
end procedure

5.4 Determining a good value for the threshold

In general, it is quite difficult to predict a good value for the threshold $q_{\mathrm{s}}^{(e,w)}(d)$, especially because optimal network performance probably requires careful selection of each $q_{\mathrm{s}}^{(e,w)}(d)$ considering the network topology, the candidate path sets and availability requirements of demands. However, if the general problem is restricted to finding a single value $q_{\mathrm{s}} = q_{\mathrm{s}}^{(e,w)}(d)$ for all backup resources (e, w), and demands d so that $r = r^{(d)}$ for all demands d is also assumed, then it is possible to answer the following questions:

- What is the lowest feasible value of r for a given q_{s} given the set of path candidates?
- What is the highest permitted value of q_{s} given the value of r and the set of path candidates?
- What is the minimal value of q_{s} needed to enable maximal backup sharing given the set of path candidates?

Even though the answers to these questions do not yield a single best value of q_{s}, they provide information of reasonable ranges to pick q_{s} from, without the need to run simulations.

5.4.1 Finding the lowest feasible value of r

The same approach could be followed to determine the value of $r_{\min}$ as detailed in section 4.4.2, except that in (4.7) $N_{\mathrm{w}}^{(d)}$ and $N_{\mathrm{b}}^{(d)}$ are always empty sets. $\check{p}^{(d)} = 1 - \hat{q}^{(d)}$ is given by (5.4) since q_{s} is assumed to be known.

Note that the set of candidate paths to be examined for determining the value of $r_{\min}$ may be restricted, which is useful when differentiated availability guarantees are computed for node pairs in different distance ranges. For example, [26] suggests different levels of availability requirements for connections of different length. This technique is used in section 5.5 to compare different availability estimation methods.

5.4.2 Finding the highest permitted value of q_{s}

If the availability requirement r can be met without using backup paths then there is no point in trying to find the highest possible value for q_{s}, because the performance of the network is then indifferent to the coice of q_{s}. However, even if a backup path is necessary for at least one demand, it may not be the case for the whole set of demands. Even so, the formal assumption used in the beginning is that there is a backup path for every demand regardless of the value of r. Note that r is assumed to be an input parameter for the following discussion.

For the highest permitted value of q_{s}, denoted by $q_{\mathrm{s,max}}$, the working and backup paths assigned to each demand satisfy the availability requirement:

$$r \geq q_{L_{\mathrm{w}}}^{(d)}\Big(1 - \prod_{e \in L_{\mathrm{b}}^{(d)}} \big(1 - (q_e + q_{\mathrm{s,max}})\big)\Big). \tag{5.5}$$

By rearranging these equations one may obtain

$$0 \geq 1 - \frac{r}{q_{L_{\mathrm{w}}}^{(d)}} - \prod_{e \in L_{\mathrm{b}}^{(d)}} \big(1 - (q_e + q_{\mathrm{s,max}})\big), \tag{5.6}$$

in which the only unknown is $q_{\mathrm{s,max}}$.

If a single demand d is considered first, then $q_{\mathrm{s,max}}(d)$ determined by this demand may be found by means of substituting the inequality sign in (5.6) with an equality sign and solving for the unknown $q_{\mathrm{s,max}}$. The order of the equation to be solved equals the number of links used by the backup path of the demand. As closed from solutions only available up to quartic equations, it is better to transform the problem to finding the eigenvalues of the companion matrix. Once the roots are obtained the following lemma helps identify $q_{\mathrm{s,max}}(d)$.

Lemma 7. *Let* $D = 1 - \frac{r}{q_{L_w}^{(d)}}$. *If* $D \leq 0$ *then* $q_{s,max}(d)$ *may be arbitrary in the range* $[0 \ldots 1 - \max_{e \in L_b^{(d)}} q_e)$, *as the unavailability of demand* d *is always less than* r. *If* $0 < D$ *then let* z *be the lowest real root of the equation. If* $z < 0$ *then no choice of* $q_{s,max}(d)$ *guarantees the availability requirement for demand* d. *Otherwise,* $q_{s,max}(d) = z$.

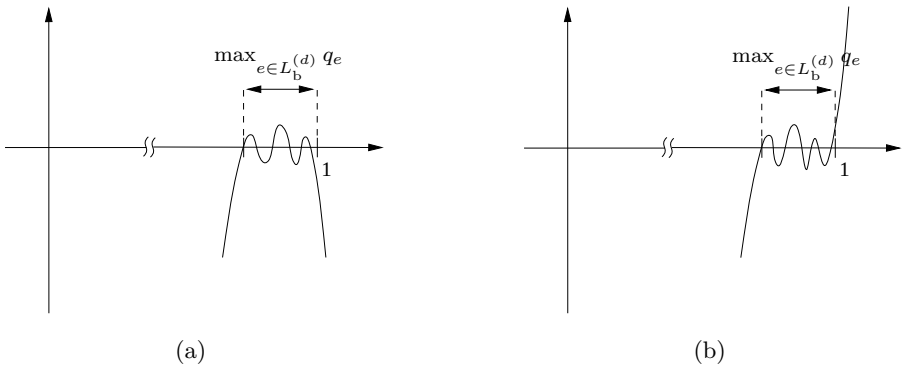
Proof. The roots of interest are in the range $[0 \ldots 1 - \max_{e \in L_{\mathrm{b}}^{(d)}} q_e)$, if any, because this is the parameter range to which a physical meaning may be associated.

Examples of even and odd polynomials to be solved are plotted on Figure 5.1 with the assumption $D = 0$. It is easy to see that if $D \leq 0$ then the curves are shifted downwards and, as a consequence, in the range of interest they remain strictly below the x-axis. In other words, (5.6) holds true as long as the value of $q_{\mathrm{s,max}}$ is in the range $[0 \ldots 1 - \max_{e \in L_{\mathrm{b}}^{(d)}} q_e)$.

Another interpretation of the case $D \leq 0$ is that $r > q_{L_{\mathrm{w}}}^{(d)}$. Obviously, it means that there is no need for a backup path for demand d if the availability requirement is r.

If, however, $D > 0$ then the curves on Figure 5.1 are shifted upwards and thus the lowest real root z is guaranteed to be less than $1 - \max_{e \in L_{\mathrm{b}}^{(d)}} q_e$. If z is below zero, then the curves remain strictly above the x-axis in the range of interest. In other words, $z < 0$ means that (5.6) cannot be satisfied.

Otherwise, picking $q_{\mathrm{s,max}}$ from the range $[0 \ldots z]$ will fulfil (5.6) and the equality holds if $q_{\mathrm{s,max}} = z$. □



(a) (b)

Figure 5.1: Polynomials of (a) even and (b) odd degree

Since the set of demands changes dynamically over time, and computations depend only on the paths used by demands, it is better to use the candidate paths for the computation. In fact, the value of $q_{\mathrm{s,max}}$ may be derived based exclusively on preprocessing the set of candidate paths using lemma 7 as shown by Algorithm 4.

Algorithm 4 returns -1 if there is at least one node pair for which no candidate paths exist that would conform the availability requirement r even if $q_{\mathrm{s}} = 0$. Otherwise, the return value is the highest q_{s} as permitted by the candidate path set.

5.4.3 Finding the minimal value of q_{s} for maximal sharing

If the set of candidate paths is known, a reasonable goal is to find the value of q_{s} that permits backup resources to be shared to the highest possible extent, denoted by $q_{\mathrm{s,maxS}}$. An obvious choice for the value of $q_{\mathrm{s,maxS}}$ is $q_{1+}^{(E)}$, because $S_{\mathrm{w}}^{(b)} \subsetneq E$ always holds true, therefore $\hat{u}^{(b)} < q_{1+}^{(E)}$ also holds true according to (5.3). However, the candidate path set may not always require such a high value.

It is, however, a difficult problem to determine $q_{\mathrm{s,maxS}}$, because it implies finding the maximum number of connections that may share a particular backup resource. The latter problem, called the maximum backup sharing set (BSS) problem, is proven to be NP-Complete in section 5.8. The proof is based on the observation that the task is similar to finding the maximum independent set in graphs (IS).

This observation also helps find a solution, because IS and its potential solutions have been studied for long. [34] shows that a simple greedy algorithm, whose objective is to find a good maximal independent set (MIS) instead of the maximum independent set (IS), performs quite well in case of IS. This algorithm is straightforward to adapt to BSS.

Nevertheless, finding the largest set of connections that may share a single backup resource does not guarantee to give the highest value of $\hat{u}^{(b)}$[2]. Therefore, the heuristic

[2] $\hat{u}^{(b)}$ is computed as in (5.3).

Algorithm 4 An algorithm to determine $q_{\mathrm{s,max}}$

procedure QSMAX(candidate path set,r)
  $q_{\mathrm{s,max}} = 1$
  for all $n_{\mathrm{s}}, n_{\mathrm{d}}$ node pairs **do**
    $q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}}) = -1$
    for all L_{w} and L_{b} path candidates for node pair $n_{\mathrm{s}}, n_{\mathrm{d}}$ **do**
      solve (5.6) using L_{w}, L_{b} and r
      if $D \leq 0$ **then**
        $q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}}) = \max_{e \in L_{\mathrm{b}}} q_e$
      else if $z < 0$ **then**
        $q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}}) = -1$
      end if
    end for
    if $q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}}) = -1$ **then**
      return -1
    else if $q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}}) < q_{\mathrm{s,max}}$ **then**
      $q_{\mathrm{s,max}} = q_{\mathrm{s,max}}(n_{\mathrm{s}}, n_{\mathrm{d}})$
    end if
  end for
  return $q_{\mathrm{s,max}}$
end procedure

assumption is made that the largest set of connections will also yield the largest $\hat{u}^{(b)}$.

Algorithm 5 shows the pseudo code of a procedure that follows the steps described above and returns an estimation for $q_{\mathrm{s,maxS}}$.

Algorithm 5 An algorithm to estimate $q_{\mathrm{s,maxS}}$

```
1: procedure QSMAXS(candidate path set)
2:     maxmis = 0
3:     for all e ∈ E links do
4:         clear nodes, clear edges, numnodes = 0
5:         for all n_s, n_d node pairs do
6:             for all L_w and L_b path candidates for node pair n_s, n_d do
7:                 if e ∈ L_b then
8:                     numnodes = numnodes + 1
9:                     nodes[numnodes] = (L_w, L_b)
10:                end if
11:            end for
12:        end for
13:        i = 1, j = 1
14:        for all i ≤ numnodes do
15:            for all j ≤ numnodes do
16:                if working paths stored in nodes[i] and nodes[j] overlap then
17:                    edges[i, j] = 1
18:                end if
19:            end for
20:        end for
21:        run MIS algorithm of [34] on the graph defined by (nodes, edges)
22:        mis = û^(b) computed using working paths stored at nodes of the MIS
23:        if mis > maxmis then
24:            maxmis = mis
25:        end if
26:    end for
27:    return maxmis
28: end procedure
```

5.5 Results

A series of experiments is carried out in order to assess the applicability of the presented results. First, the applicability of the presented sharing unavailability threshold based

provisioning method is demonstrated. Then, a comparison is made with the extended DiR method discussed in Chapter 4.

Unfortunately, to the best of the author's knowledge no other works exist in the literature that could serve as a fair basis of comparison. The majority of the papers propose provisioning methods based on non-conservative connection availability estimation algorithms [40, 62, 94, 112, 113]. They would obviously perform better in terms of blocking probability, as they are not so strict about backup resource sharing. At the same time their "guaranteed availability" is not a hard guarantee in contrast to the guarantees of the extended DiR and sharing unavailability threshold based methods presented in this thesis. The two conservative estimation mechanisms known to the author are proposed for a network design scenario [5, 97, 102] and they are, therefore, not suitable for on-line application in their current forms. Further discussion on this may be found in section 5.6.

After the performance comparison, the parameter range suggestions provided by the algorithms in section 5.4 are evaluated.

5.5.1 Methodology

The efficiency of the outlined on-line RWA algorithm can be assessed by means of simulation. Therefore, a simulator was implemented using the assumptions in section 5.2.

The simulator takes the network topology, the λ call generation rate and the $r^{(d)} = r$ availability requirement of calls as input, and it generates random calls in the network. Call generation is a Poisson process with parameter λ. Source and destination nodes of calls are selected uniformly and the same availability requirement is considered for all the generated calls. The duration of established connections is exponentially distributed with a mean of one time unit.

The number of rejected and total call requests is recorded, and the ratio of the two gives the call blocking probability, which has been used to characterize network performance.

The definition of $q_{\mathrm{s}}^{(e,w)}(d)$ in section 5.3 permits the use of a measure that depends on the backup resource and the connection. This study, however, attempts to characterize the effect of a uniform constant sharing unavailability threshold only. Therefore, the value of $q_{\mathrm{s}}^{(e,w)}(d) = q_{\mathrm{s}}$ is also an input parameter of the simulator.

The on-line RWA algorithm must find a routing and wavelength assignment for working and — if necessary — protection lightpaths that guarantee the survivability requirement.

Since even the decision problem involved in solving the routing and wavelength assignment problem is NP-complete [69], efficient resource allocation for each call is done by running an optimization process based on Simulated Annealing (SA). Further

details of the optimization and the simulator may be found in Appendix B.

The cost function used during the optimization of resource allocation is defined based on the idea presented in [104]. Capitalizing on the fact that the candidate path sets to be used by demands are known *a priori*, each link may be assigned the number of candidate paths that traverse the link. These numbers are normalized with the total number of candidate paths to obtain link weights. The weight assigned to a certain candidate path equals the highest one of all the weights of links traversed by the path. The cost function of SA assigns to a potential RWA solution for a demand the sum of the weights of working and (if needed) protection path candidates, multiplied by 10. The multiplication is needed to adjust the range of the cost function to that of the one for which SA parameters are optimized[3].

Verification of survivability constraints and the decision about call admission are implemented according to the algorithm outlined in section 5.3. This verification is integrated into the optimization loop of the SA-based RWA algorithm. Thus at each SA iteration the availability constraint must be satisfied in order to accept the solution examined in the current iteration to be a feasible solution.

For the simulations the topology of a European optical network [96] is used, which is shown on Figure A.2. In all simulations the number of wavelengths per fiber W is set to 32. The unavailability of any link is determined using (2.2).

The presented simulation results have 10% confidence intervals at a 95% confidence level.

In order to assess the quality of parameter range estimation algorithms, they are implemented as well, and the results are compared with results of simulations.

5.5.2 Simulations

Figures 5.2-5.5 show the blocking probability plotted as a fuction of network load with different survivability requirements.

In the European network the average link length is about 644 km, which corresponds to an average $q_e = 2.5 * 10^{-3}$. In case of $r = 0.005$ only about 55% of the connections needs a protection path, while in case of $r = 0.0007$ this ratio is around 99%. To facilitate comparison the blocking probability of DPP is also plotted, obtained by setting $q_\mathrm{s} = 0$. For the different availability requirements different values of q_s were tested. The reason for the choices is that lower values of r obviously require lower values of q_s. Otherwise there may be at least one node pair for which no connection request could be served due to the increased "unavailability" of backup resources. A smaller blocking probability indicates more efficient use of total resources.

One may intuitively reason that a resource efficiency gain due to sharing can only be

[3]See the discussion in Appendix B.

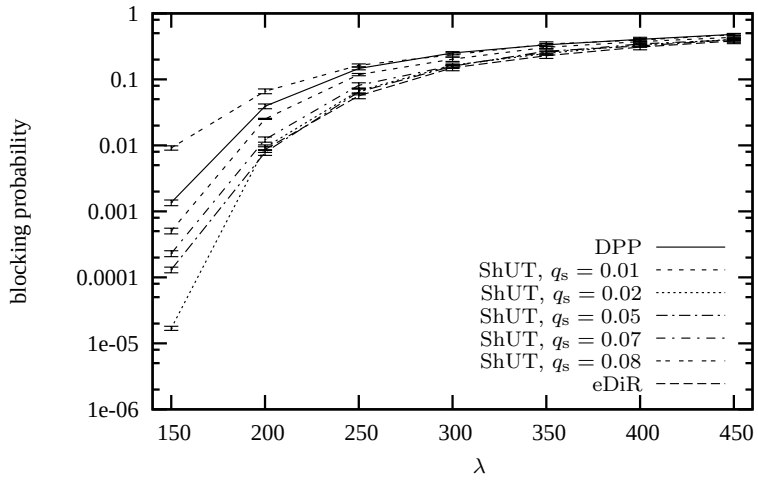


Figure 5.2: Blocking probability as a function of network load at $r = 0.005$

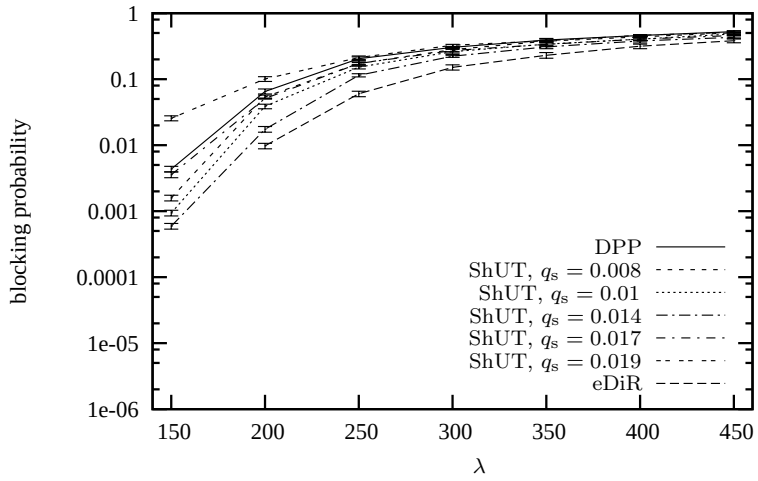


Figure 5.3: Blocking probability as a function of network load at $r = 0.0015$

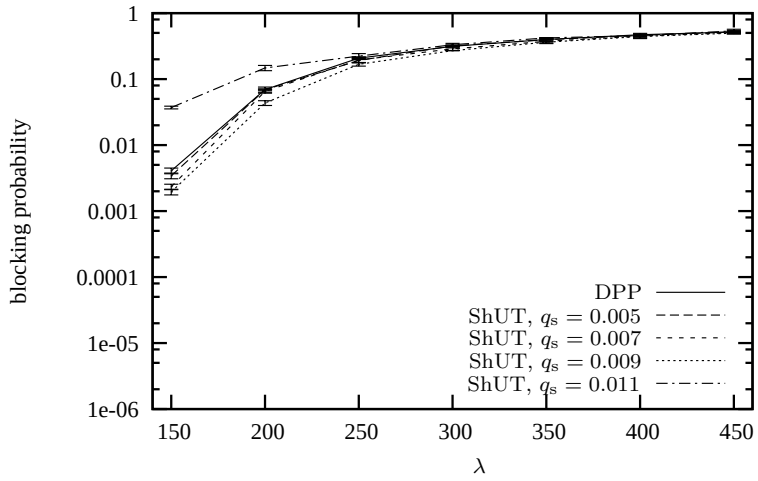


Figure 5.4: Blocking probability as a function of network load at $r = 0.001$

expected if a significant portion of the connections needs protection. However, the more stringent the availability requirement is, the lower the decrease is in terms of blocking that can be accomplished by an appropriate selection of q_s. The reason for this is that a lower value of r also means that backup resource sharing must be limited more strictly in order to ensure high availability of connections. In the examined range of r this latter effect (i.e., the need to limit sharing) is stronger than the former one (i.e., the more connections need backup the higher the resource efficiency gain may be).

It can be concluded from Figure 5.2 that the gain in blocking probability can be as big as an order of magnitude compared to the DPP method. On the other hand, the gain becomes marginal when the availability requirement is very strict (see Figure 5.5).

Figure 5.6 is a "cross section" of Figure 5.2, while Figure 5.7 is that of Figure 5.3, both taken at $\lambda = 200$.

Figures 5.6 and 5.7 shed light on how the achievable blocking probability depends on the choice of q_s. The curves on both figures resemble a bathtub, i.e., there seems to be a range of q_s, where blocking is lower than outside this range. It might be a non-intuitive observation at first.

This phenomenon can be explained as follows. As q_s starts to increase from 0 gradually more and more sharing is possible, which leads to more efficient resource use and lower blocking. If q_s is further increased, some of the paths of longer connections can no longer provide a feasible routing because of the backup resources of seemingly less availability. As a consequence, these connections will have to wait in the single-slot buffer[4] until their path pairs of highest availability become available, which increases blocking. Note that the phenomenon would be perceived even without the single-slot

[4]The single-slot buffer and other details of the simulator are explained in Appendix B.

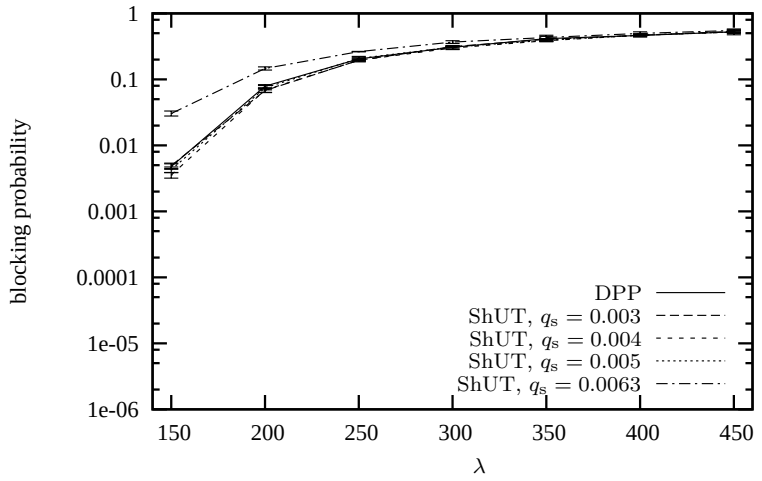


Figure 5.5: Blocking probability as a function of network load at $r = 0.0007$

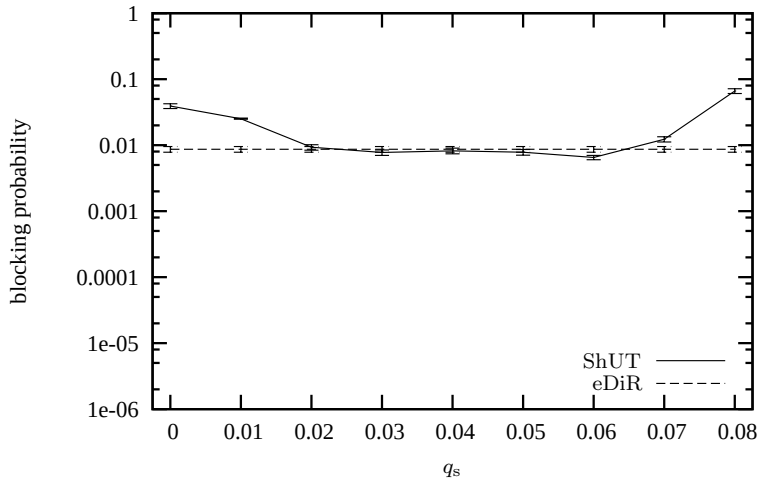


Figure 5.6: Blocking probability as a function of q_s at $\lambda = 200$

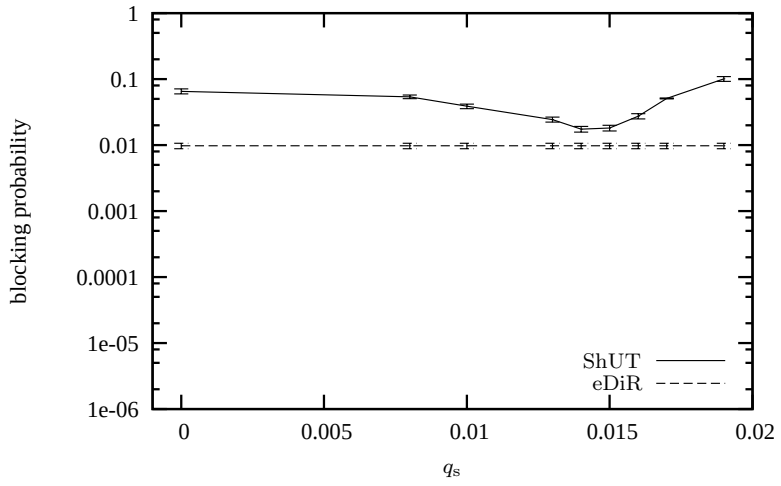


Figure 5.7: Blocking probability as a function of q_s at $\lambda = 200$

	$r_{\min}$	
network	**SPP with eDiR**	**DPP**
US	0.00236885	0.000586236
European	0.00133932	0.000301568
Italian	0.0000754779	0.0000212594
metropolitan	0.00000374488	0.00000192799

Table 5.1: Best feasible availability guarantees for different networks without node failures

central buffer, but the buffer strengthens the effect.

5.5.3 Comparison with the extended DiR method

For comparison, the simulator is also run using the extended DiR method as an estimation of connection availability. The curves obtained are also plotted on Figures 5.2, 5.3, 5.6 and 5.7 marked with "eDiR". The availability requirements used for obtaining results plotted on Figures 5.4 and 5.5 are too strict for the SPP-eDiR method. Table 5.1 shows the best feasible availability guarantees for different networks without node failures obtained as described in section 4.4.2.

It is obvious that — without node failures — the application of the ShUT method may be worth only in networks of continental scale.

If the r availability requirement is not too strict, then the blocking probability of the extended DiR method and that of the ShUT method may be close to each other with an appropriate choice of q_s, as demonstrated by Figures 5.2 and 5.6. If, on the other hand, r is close to the $r_{\min}$ of the extended DiR method, then the blocking of the ShUT

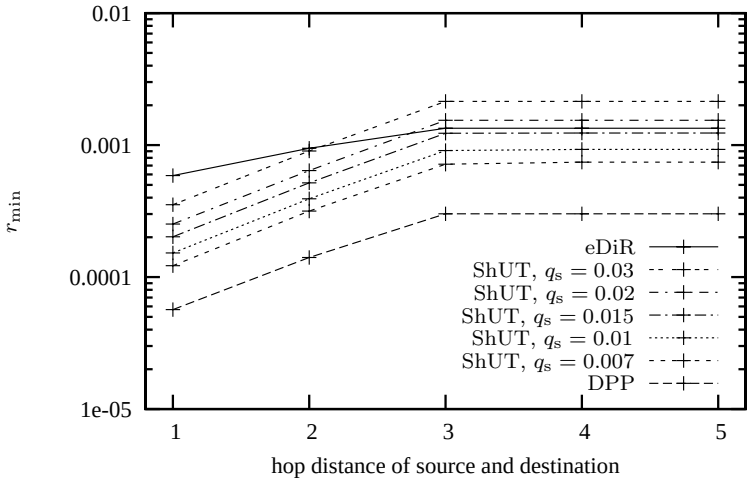


Figure 5.8: Best guarantees for connections of different length in the EU network

method may be higher (see Figures 5.3 and 5.7).

If the r availability requirement is decreased beyond the $r_{\min}$ of the extended DiR method, then the ShUT method may still be able to decrease the blocking compared to that of DPP by a factor of two, as shown on Figure 5.4. However, too low values of r lead to very strict backup resource sharing constraints, which practically means that backup resources need to be dedicated in order to fulfil availability requirements.

[26] defines standard connection availability requirements as a function of physical distance between connection endpoints. Due to the different nature of the estimation methods, the ShUT method potentially performs better in this respect. Figure 5.8 compares the values of $r_{\min}$ obtained for connections of different length according to the procedure described in section 5.4.1. Note that the lines are plotted only to show trends, because results are interpreted only at discrete values along the x-axis.

Even in case of $q_s = 0.02$, when the overall blocking probability of the ShUT method matches that of the extended DiR method for $r = 0.005$ (see Figure 5.2), the ShUT method may offer significantly better guarantees for one-hop and two-hop connections than the extended DiR method. It is another advantage of the ShUT method that these enhanced guarantees are provided without any additional measures to be taken.

The reason for the plateau of the curves after 3 hops on Figure 5.8 is the following. Intuitively, the most distant nodes in the network in terms of physical distance will necessarily have the path pairs with the least availability connecting them. Best guarantees that uniformly apply to a particular subset of node pairs are thus determined by the most distant node pair within the subset. In the EU network nodes number 3 and 13 are the two most distant nodes, and the shortest path between them consists of 3 links. Nevertheless, the largest distance between two nodes in terms of hops in this network is

r	$q_{\mathrm{s,maxS}}$	$q_{\mathrm{s,max}}$
0.0007	0.0815029	0.0063394
0.001	0.0815029	0.0111975
0.0015	0.0815029	0.0193748
0.005	0.0815029	0.0840775

Table 5.2: Estimations of ranges of interest for q_{s} in the EU network

4 (c.f. section A).

5.5.4 Estimations of ranges of interest for q_{s}

Table 5.2 shows values obtained with methods presented in section 5.4. The input parameters of the estimation methods are the same as that of the simulator.

Comparison of the table and the simulation results shows that the estimation of $q_{\mathrm{s,max}}$ is always a good upper limit for the range of interest for q_{s}. By definition, higher values do not make sense. However, there is no point in selecting a value that is higher than $q_{\mathrm{s,maxS}}$, because that would not increase sharing any further.

On the other hand, it seems that there is no need to enable maximal sharing, which also agrees with the observations made in [71–73]. The explanation of this is the same as that of the climbing stage of the blocking probability curve towards higher threshold values on Figures 5.6 and 5.7. Note that the phenomenon would be perceived even without the single-slot central buffer, but the buffer strengthens the effect.

5.6 Implementation of provisoning methods

The implementation of the extended DiR method, as presented in this thesis, requires path precomputation and source routing functions in the network. Path selection requires accurate information on the current network state, i.e., it is necessary to maintain a link state database. The estimation method ensures that the availability requirements of already established connections are always fulfilled.

The extended DiR method does not scale very well, since the pessimistic assumption used for the connection availability estimation becomes more and more pessimistic as the number of components in the network increases. Another theoretical problem is that if new components appear in the network (e.g. an extension takes place), then the guarantees are not necessarily maintained for already established connections.

The implementation of the sharing unavailability threshold based method, as presented in this thesis, also requires path precomputation and source routing functions in the network. However, additional state information (sets $L_{\mathrm{w}}^{(d)}$ for each $d \in D^{(e,w)}$ and $q_{\mathrm{s}}^{(e,w)}(d)$) has to be stored at nodes per each network resource and each connection.

Path selection requires accurate information on the current network sate including the additional information mentioned above.

Note that the evaluation of the third criterion in section 5.3 may be speeded up significantly, if a single $q_{\mathrm{s}}^{(e,w)}$ value is used for an individual backup resource. Thus the speedup applies to the case when a single value of q_{s} is used throughout the whole network, as well.

The idea of the speedup is the following. Instead of checking the condition for each demand and each shared backup resource, it suffices to check $\max_{d_i} \hat{u}^{(e,w)}(t_{\mathrm{a}}^{(d)}, d_i) \leq q_{\mathrm{s}}^{(e,w)}$ for each shared backup resource (e, w). The value of $\max_{d_i} \hat{u}^{(e,w)}(t_{\mathrm{a}}^{(d)}, d_i)$ may be obtained in constant time if $L_{\mathrm{w}}^{(d)}$ contributions of demands $d \in D^{(e,w)}$ to $S_{\mathrm{w}}^{(e,w)}$ are stored in an ordered list. Therefore, checking the third criterion requires the evaluation of $|L_{\mathrm{b}}^{(d)}|$ inequalities at most. The call admission control method (see Algorithm 3) ensures that the availability requirements of already established connections are always fulfilled.

The ShUT method scales better than the extended DiR method, since the feasible guarantees may always approach that of the DPP method. Moreover, these guarantees are automatically maintained even if the network topology changes, e.g. new components are added.

The resource reservation step, which is needed to establish the connection after the decision is made about the admission, may be carried out with help of a protocol such as an appropriate extension of RSVP, like [8]. Both provisioning methods are robust in a sense that each node along the working and protection paths has all the necessary information to decide whether the connection establishment may proceed. If the admission decision is made upon inaccurate network state information, any node may send an upstream PATH TEAR message thus notifying the nodes through which the PATH message already passed.

The on-line implementation of the conservative connection estimation method presented in [5, 102] would require the maintenance of essentially the same information as needed by the ShUT method. However, in case of [5, 102] it is not clear how one can easily check if a new connection d violates the availability requirements of already established ones without actually evaluating $\sum_{e \in L_{\mathrm{b}}^{(d)}} |D^{(e,w_{\mathrm{b}})}|$ inequalities. Moreover, the evaluation of each of these inequalities requires $O(\sum_{e \in L_{\mathrm{b}}^{(d)}} |D^{(e,w_{\mathrm{b}})}|)$ operations, whereas the inequalities to be checked with the ShUT method need only $O(E)$ operations each.

The on-line implementation of the conservative estimation method in [97] is questionable, because the estimation procedure involves state space sampling. Moreover, the approach discussed in [97] does not seem to be able to handle the change in sharing relationships among connections over time, which is an inherent characteristic of the dynamic traffic scenario.

5.7 Summary

This chapter presented an on-line RWA algorithm for shared (backup) path protection that is capable of connection provisioning with guaranteed availability in the presence of multi-component failures. The probabilistic guarantees are based on the efficient computation of bounds using a threshold that is assigned to each shared backup resource to determine the extent of sharing. The appropriate choice of a uniform threshold value may lead to a gain in blocking probability as high as an order of magnitude compared to the performance of dedicated path protection while fulfilling the availability requirements and requiring fairly simple computations. The proposed method thus offers an attractive and efficient adaptation of shared backup path protection without any significant overhead. The discussion in this chapter assumed that nodes are perfect in order to avoid complex notations. This assumption, however, may be relaxed.

5.8 Proof of NP-Completeness of maximum backup sharing set (BSS)

The problem of determining the largest set of demands whose backup lightpaths may share a wavelength on a particular link is NP-Complete even if the route candidates to be used by demands are known in advance.

The basic idea of the proof presented here is the following. First, the BSS problem is stated as a decision problem, K-BSS, which considers the existence of a subset of K demands that may share the backup resource in question. K-BSS is shown to be in NP afterwards. The K-independent set (K-IS) problem is then reduced to the K-BSS problem, that is, a polynomial time transformation is shown that assigns a problem instance of K-BSS to any arbitrary instance of K-IS. Next, the K-IS problem instance is shown to have a solution if and only if the K-BSS problem instance has one. The NP-Completeness of K-BSS and BSS immediately follows.

In what follows, the maximum backup sharing set problem is formally stated first.

Instance: let D be a set of demands that need to be served in the network represented by graph $G(V, E)$. Let each demand $d \in D$ have a single working lightpath $L_{\mathrm{w}}^{(d)} \subset E$ and a single protection lightpath $L_{\mathrm{b}}^{(d)} \subset E$ and let $z \in E$ be such that $z \in \bigcap_{d \in D} L_{\mathrm{b}}^{(d)}$.

Question: What is the size of the largest subset $D' \subseteq D$ of demands whose backup lightpaths may share a wavelength on link z subject to the mutual disjointness of their working lightpaths?

Theorem 2. *BSS is NP-Complete.*

Proof. First, the K-BSS problem is defined as the problem of deciding whether there exists a subset $D' \subseteq D$ of K demands whose backup lightpaths are allowed to share a

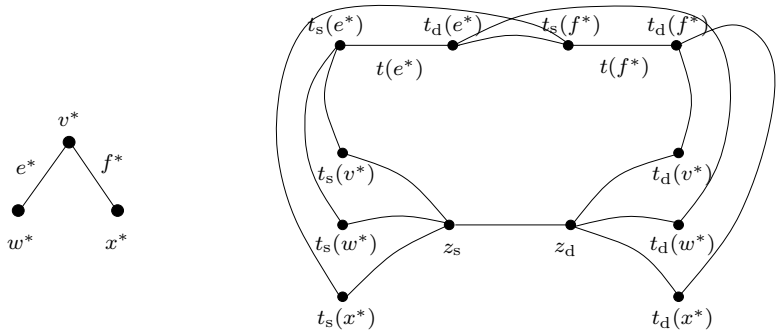


Figure 5.9: Reduction of an instance of K-IS to K-BSS

wavelength on link z. If this decision problem is NP-Complete, then BSS must also be NP-Complete.

K-BSS $\in$ NP, since for any solution $D' \subseteq D$ the number of demands in set D' and the mutual disjointness of the working lighpaths of demands in set D' may be verified in polynomial time.

The K-independent set (K-IS) problem is polynomial-time reducible to K-BSS. Let $G^*(V^*, E^*)$ be a graph representing an instance of the K-IS problem, for which another graph $G(V, E)$ is constructed.

Since the existence of parallel edges does not change the solution of the K-independent set problem, it is assumed that $G^*(V^*, E^*)$ does not contain such edges for the ease of discussion. $G(V, E)$ is not assumed to contain neither isolated points nor isolated components to make the discussion simpler. It is easy to see that these special cases may be dealt with in polynomial time as well by means of partitioning the graph and solving the K^*-independent set problems for the subgraphs for $K^* \leq K$.

The reduction may be accomplished as illustrated on Figure 5.9. Initially, let $V = \{z_s, z_d\}$ and $E = \{z = (z_s, z_d)\}$.

For each edge $e^* \in E^*$ two corresponding vertices, $t_s(e^*)$ and $t_d(e^*)$, and one corresponding edge, $t(e^*) = (t_s(e^*), t_d(e^*))$ are created in V and E, respectively.

For each $v^* \in V^*$ two vertices, $t_s(v^*)$ and $t_d(v^*)$ are added to V. Additionally, for each node $v^* \in V^*$ other $\deg(v^*) + 3$ edges are created in E, where $\deg(v^*)$ is the degree of node v^*, as follows. Let $E^{**} = \{e_1^*, e_2^*, ..., e_N^*\} \subseteq E^*$ be the set of edges in G^* incident to v^*. For each $e_i^* \in E^{**}, i < N$ an edge is added to E that connects $t_d(e_i^*)$ and $t_s(e_{i+1}^*)$. The four additional edges to be added to E are $(t_s(v^*), z_s)$, $(t_s(v^*), t_s(e_1^*))$, $(t_d(e_N^*), t_d(v^*))$ and $(z_d, t_d(v^*))$.

Moreover, for each node $v^* \in V^*$ a demand $d(v^*) \in D$ is generated, so that $L_w^{(d(v^*))} = \{(t_s(v^*), t_s(e_1^*)), t(e_1^*), t(e_2^*), ..., t(e_N^*), (t_d(e_N^*), t_d(v^*))\}$ and $L_b^{(d(v^*))} = \{(t_s(v^*), z_s), z, (z_d, t_d(v^*))\}$.

This reduction may be carried out in polynomial time.

Consider that by construction $v_1^* \in V^*$ and $v_2^* \in V^*$ are connected by an edge (v_1^*, v_2^*) if and only if $L_{\mathrm{w}}^{(d(v_1^*))}$ and $L_{\mathrm{w}}^{(d(v_2^*))}$ are non-disjoint. If $e^* = (v_1^*, v_2^*) \in E^*$ exists then $L_{\mathrm{w}}^{(d(v_1^*))} \bigcap L_{\mathrm{w}}^{(d(v_2^*))} = \{t(e^*)\}$. Conversely, if $L_{\mathrm{w}}^{(d(v_1^*))} \bigcap L_{\mathrm{w}}^{(d(v_2^*))} = \{(e_1, e_2)\} \in E$, then there is an edge $e^* = (v_1^*, v_2^*) \in E^*$ for which $t_{\mathrm{s}}(e^*) = e_1$ and $t_{\mathrm{d}}(e^*) = e_2$.

As a consequence, an independent set of size K exists in $G^*(E^*, V^*)$ if and only if there is a subset of demands $D' \subseteq D$ of size K, whose working lightpaths are disjoint, and, therefore, who are allowed to share backup resources on link z.

Since the K-IS problem is NP-Complete, so is K-BSS. Thus BSS is NP-Complete, as well. □

Chapter 6

Applicability of Results

The goal of this Chapter is to discuss the conditions and limits to the applicability of the algorithms and methods presented in this thesis. First, the algorithm presented in Chapter 3 is shown to be a general tool in reliability modeling. Then the effects of network state information inaccuracy are quantified, which appear when the provisioning schemes proposed in Chapters 4 and 5 are operated in a real network. Finally, a note is made with respect to the conversion of connection availability guarantees to SLA terms.

6.1 Applicability of the failure stratum probability computation algorithm

The algorithm presented in Chapter 3 to determine failure stratum probabilities only assumes the following properties of the reliability model of the system:

- Components have two states, and
- components may change states independently of each other, and
- the asymptotic probability that the component is in the failed state is known for each component.

The algorithm may be applied to any system, whose model conforms the requirements listed above. Thus the applicability of the proposed method may be extended from models of telecommunications networks to a general reliability modeling context.

6.2 Distributed operation of provisioning schemes

Both of the provisioning methods presented in Chapters 4 and 5 assume complete knowledge of the current network state, i.e., they rely on the existence of a link state database. Link state databases are maintained in parallel at each node in networks that require link state information for proper operation. These databases try to reflect real-time changes of the network state. Information dissemination, however, by its very nature introduces latency in the update of link state information at different nodes, which may in turn lead to faulty decisions that entail false demand rejections and additional delay perceived by user traffic.

This section presents a probabilistic model based on simple assumptions in order to derive an upper bound on the probability that such a link state database inconsistency occurs [74].

6.2.1 Related work

Maintaining up-to-date information about the current network state is fundamental for most routing mechanisms that rely on full information. Link state information complemented with resource allocation data is especially important for Traffic Engineering (TE), for service provisioning with Quality of Service (QoS) or Quality of Protection (QoP) guarantees, as well as for the operation of certain technologies that impose granularity constraints on resource management such as WDM and DWDM.

Management and dissemination of link state information is achieved by means of using intradomain routing protocols such as IS-IS [68] and OSPF [66], both of which have extensions for TE. In addition to this, the Generalized Multi-Protocol Label Switching (GMPLS) architecture [58], which was proposed as a unified control plane for future data and transmission networks, is defined so that it can rely on these protocols to manage the link state database.

Call admission control and network resource allocation functions may be implemented in general at each node in a network. Even though call admission is most likely to be operated in a distributed manner, the decisions may require knowledge of the global network state. While distributed call admission is highly desirable for scalability reasons it can be the primary source of the occurrence of inconsistencies in the link state database.

If a network node considers a particular resource not available the call admission control algorithm running at the node may reject calls because of the apparent lack of adequate resources. However, it may also occur that the information at the node is out-of-date and the real state of a particular resource may be the opposite to the one perceived by the node. It would result in rejecting calls that could be, in fact, admitted at the moment. On the other hand, if a particular resource is marked as available in the link state database of a network node while it is already not available in reality, then calls may be admitted without enough resources in the network to serve them. Any of these events is highly undesirable unless the service being offered by the network is best-effort. Obviously, an appropriate resource reservation mechanism and feedback in the call admission process may ensure that only available resources are reserved. However, it undoubtedly introduces additional delay, which is perceived by the users, and it does not solve the problem of false demand rejections.

This brief study intends to propose a simple probabilistic model for estimating how often the discussed inconsistencies may occur. Among the related works in the liter-

ature [4] was amongst the first ones to examine routing protocol overheads and information update policies, thus entering an area investigated by many other authors later on. Obviously, one approach to the problem is to accept the existence of information uncertainty and to propose methods that operate relying on uncertain information. An excellent survey of such methods is presented in [59]. However, these works do not quantify the significance of information uncertainty when routing protocols are used for link state database maintenance and update.

[55] proposed a theoretical model for OSPF-TE and GMPLS bandwidth usage and memory consumption in routers and reported results for various dissemination policies. Various information dissemination policies are analyzed also in [95] with respect to their influence on the call blocking ratio without assuming any specific routing protocol. The behavior of a simplified OSPF implementation is examined by means of simulation in [17] and probabilities of blocking due to effects of information inaccuracy at different stages of the connection establishment are presented. Finally, an experimental approach is adapted by [91] to obtain some performance data of OSPF implementations.

Even though there are several works on related subjects, to the best of the author's knowledge the specific study presented in this section, i.e. a probabilistic analysis of link state database inconsistency with emphasis on real-world routing protocol performance has not been addressed yet.

6.2.2 Network model

Consider a network represented by a directed graph $G(V, E)$, where each vertex $v \in V$ represents a node with all its equipment and each edge $e \in E$ represents a certain wavelength channel on a certain unidirectional fiber. Let $G_f(V_f, E_f)$ be a subgraph of G so that $V_f \subset V$ and $E_f \subset E$ and $E_f = \{(v_1, v_2) | v_1, v_2 \in V_f\}$. Let $G_f(V_f, E_f)$ contain only the resources that may be assigned to the future call requests.

Each call request d comprises a couple of vertices $n_s^{(d)} \in V$ (source) and $n_d^{(d)} \in V$ (destination) representing the end nodes of a connection to be provided by the network. The call duration defines the time period when resources have to be assigned to the call. In order to achieve this, node $n_s^{(d)}$ makes the decision about resource allocation and call admission based on its current information about the state of the network, $G_f^{(n_s^{(d)})}(V_f^{(n_s^{(d)})}, E_f^{(n_s^{(d)})})$. In other words, the decisions of node $n_s^{(d)}$ rely on the content of its link state database, $G_f^{(n_s^{(d)})}$. Node $n_s^{(d)}$ may then either reject the call because of the lack of resources, or assign some of the available resources to the call thus admitting it to the network. Note that in order to remain general no further properties are assumed with respect to the call admission control and resource allocation algorithm used in the network. Obviously, the model may be extended to include further state information that is fundamental e.g. for constrained routing algorithms. However, this does not

yield significant changes with respect to the expected results.

Nodes in the network use some protocol for the dissemination of link state information in order to update their link state databases. Thus, call admission and resource allocation can be done in a distributed manner. Each state change in the network is not assumed to trigger the information update mechanism of the protocol. Instead a periodical pacing timer [91] is assumed that initiates the distribution of link state information accumulated since the last tick in order to decrease signaling processor load. In other words, link state information aggregation is considered in the current analysis. Nevertheless, no additional properties of the protocol are assumed, in order to remain general.

State changes in the network

Inconsistency of the link state database of node s is a difference between $G_f^{(n_s^{(d)})}$ and G_f. There may be several reasons for temporary inconsistencies in the link state database. The discussion of protocol specific issues is out of the scope of the study; therefore, all of the issues to be addressed here are due to delayed notification about changes in the information to be represented, i.e., changes in G_f. The following is a list of possible changes:

- Links and/or nodes are added to or removed from the network. This happens whenever an investment is made in order to expand or restructure the network. This type of change may be pre-planned, therefore, it is also omitted in the analysis.
- Links and/or nodes go down or are brought up. The possible reasons include cable cuts, software failures, maintenance operations or other network management related decisions. The first two are related to failure characteristics of equipment and cable, which are of paramount importance, as they cannot be completely avoided. On the other hand, the operator has control over the latter two; therefore these two are omitted from the current analysis. Note that node failures may be considered as simultaneous failures of all the links attached to the node.
- Resources are assigned to admitted connections or are released by terminating connections. This happens each time a call is admitted or torn down, consequently, it is a function of the traffic to be served.

A probabilistic approach is adopted in order to characterize the effect of these changes, or more precisely, the effect of delays of notifications due to the distributed nature of the environment. Let $A^{(d)}$ denote the event that $G_f^{(n_s^{(d)})} \neq G_f$, and let $B^{(d)}$ denote the event that a call admission decision has to be made at node $n_s^{(d)}$, and let $C^{(d)}$ denote the event that a wrong decision is made at node $n_s^{(d)}$. Basically, the following probabilities need to be considered:

$P(A^{(d)})$, the probability that the link state database of node s is inconsistent with the real network state,

$P(B^{(d)}|A^{(d)})$, the probability that a decision has to be made at node s based on inaccurate state information, and

$P(C^{(d)}|A^{(d)}, B^{(d)})$, the probability that a wrong decision is made due to the inaccuracy.

Thus the probability that node $n_s^{(d)}$ is adversely affected by link state database inconsistency is

$$P(C^{(d)}|A^{(d)}, B^{(d)})P(B^{(d)}|A^{(d)})P(A^{(d)}). \tag{6.1}$$

Nevertheless, in order to simplify the analysis in the present study, $P(C^{(d)}|A^{(d)}, B^{(d)}) = 1$ is assumed. It means that each decision made using inaccurate state information is considered to be incorrect.

State change model

A two-state stochastic process, denoted by $N^{(v)}$, is assumed to model the changes in the network as perceived by a particular node v. The random variable $I^{(v)}(t)$ takes the value 1 if $G_f^{(v)} = G_f$ at time t, and 0 otherwise. Let the call request arrival instants at node v be modeled by a point process $R^{(v)}$. At each arrival instant t_i of this point process the value of $I^{(v)}(t_i)$ is observed, which corresponds to accessing the link state database in order to make a decision about call admission and resource assignment. Figure 6.1 illustrates the processes $N^{(v)}$ and $R^{(v)}$.

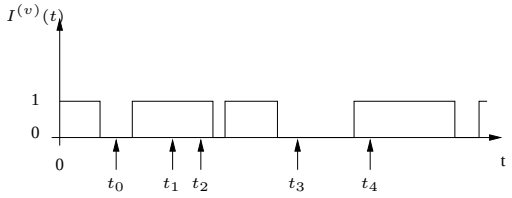


Figure 6.1: Stochastic process modeling the consistency of the link state database of node v.

Assuming stationary behavior of the system it follows directly that

$$P(A^{(d)}) = \lim_{t\to\infty} P(I^{(n_s^{(d)})}(t) = 0), \tag{6.2}$$

and

$$P(B^{(d)}|A^{(d)})P(A^{(d)}) = \lim_{i\to\infty} P(I^{(n_s^{(d)})}(t_i) = 0). \tag{6.3}$$

In some notable special cases, however, the computation may be simplified. If only equipment and cable failures are modelled by $N^{(v)}$, then processes $N^{(v)}$ and $R^{(v)}$ may

be considered to be independent. As a consequence,

$$P(B^{(d)}|A^{(d)})P(A^{(d)}) = P(A^{(d)}). \tag{6.4}$$

On the other hand, if resource assignments are also modelled by $N^{(v)}$ then the two processes become dependent as call admission affects the network state. In this case only a weaker claim may be made by applying the PASTA theorem [105]: if $R^{(v)}$ is a Poisson process then (6.4) still holds. In short, determining $P(A^{(d)})$ is sufficient provided that the conditions above are met.

Steady state analysis

The analytic derivation of $P(A^{(d)})$ is still complex, therefore an upper bound on this probability is computed as follows. Let T_0 be a stochastic variable that shows the amount of time that process $N^{(v)}$ spends in the 0 state and let $E(T_0)$ be the expected value of T_0. If $\bar{\mu}$ is the frequency of state changes of process $N^{(v)}$ from state 1 to state 0, then the following holds true:

$$P(A^{(d)}) = E(T_0)\bar{\mu} < \hat{P}(A^{(d)}) = \hat{E}(T_0)\hat{\bar{\mu}}. \tag{6.5}$$

A good value for $\hat{\bar{\mu}}$ is the intensity of network state change events. Two quantities contribute to this: component failure intensities, which may be computed using data published in [49] and [100], and the λ call request arrival intensity.

When analyzing the different time components of $\hat{E}(T_0)$ the notation introduced in [6] is followed. Consider the time elapsed from the occurrence of the state change in the network until the notification containing the respective information is processed by a certain node. This time consists of the following components:

T_1, the detection time that is necessary for the adjacent network nodes to perceive the event. It obviously depends on the nature of the event, and the detection mechanisms used. Based on [6] 20ms is used in this study.

T_2, the hold-off time, which is an additional delay introduced in order to let lower layers deal with failures before the more time consuming reactions of higher layers begin. The hold-off time is assumed to be zero in the present study.

T_3, the notification time, which may be calculated as detailed below.

In order to estimate T_3, the following additional definitions are used:

$T_{\mathbf{m}}$, the link state advertisement (LSA) message processing time in the nodes. This depends on the size of the LSA packet, but based on measurements 0.8ms seems to be a reasonable upper bound [91].

ρ**,** the load of the signaling processor. It is a function of the intensity of LSA message arrivals from all over the network, which depends heavily on the applied protocol. However, with some simple assumptions it can be upper bounded, as discussed later.

$\bar{s}(\rho)$**,** the sum of the waiting and processing time in the nodes, which is a function of the load of the signaling processor. If the arrival process of LSA's at a single node from all over the network is approximated with a Poisson process (which is a reasonable assumption since there are many independent event sources) and each LSA is assumed to require the same amount of processing time then the respective formula of the M/D/1 queue may be applied that yields $\bar{s}(\rho) = T_{\mathrm{m}} \frac{\rho}{2(1-\rho)}$. Note that the burstiness of the LSA arrival process, which would lead to worse waiting times, is implicitly assumed to be negligible.

$T_{\mathbf{f}}$**,** the pacing timer interval, which determines the frequency at which packets that contain aggregated LSA information are sent. In the worst case the LSA has to wait for the entire interval before it is forwarded. The value assumed for T_{f} is 33ms [91].

$T_{\mathbf{d}}(l)$**,** the propagation delay on link l, which is a function of the length of the link. Generally it is assumed to be 5μs on a fiber of 1km. $T_{\mathrm{d,max}}$ is used to denote the maximal $T_{\mathrm{d}}(l)$ over all links.

Information about each state change in the network sooner or later reaches any node (provided that the network remains connected). Due to the flooding strategy that is generally applied in link state dissemination a particular piece of information may be forwarded to a node from all of its neighbors in the worst case. If duplicate or outdated LSA's are not forwarded then the degree of node v is the upper bound on the number of LSA instances that carry the same information to node v.

From time to time the forwarding information base (FIB) also has to be updated at each node. This may involve route computations, whose time consumption depends on the network size, and also an update operation that may last up to 300ms based on measurements [91]. Measured route computation times in milliseconds may be approximated with $T_{\mathrm{FIB}} = 0.00252z^2 - 0.0108z + 1.2$, where $z = |V|$, in case of OSPF [91]. Note that the link state database update is included in the processing time of LSA messages, and the FIB update only appears here as a possible additional load factor for the signaling processor. The frequency of FIB updates is denoted by $\bar{\phi}$.

The signaling processor load may then be estimated as

$$\rho = \bar{\lambda} d_{\max} T_{\mathrm{m}} + \bar{\phi} T_{\mathrm{FIB}}, \tag{6.6}$$

where $d_{\max}$ is the maximal node degree in the network. Thus, (6.6) gives the total load for OSPF. Omitting the last term yields the load due to link state database updates only.

Note that in general there is a certain amount of additional work associated with aggregating LSA information, which is considered to be negligible in the present analysis based on the following considerations. [91] measured the time spent by a router in between the receipt of an LSA packet and the forwarding of LSA information and found that it is primarily determined by the pacing timer and not the amount of information to be aggregated. That is, the time necessary for the signaling processor to aggregate LSA information was found to be significantly smaller than the pacing timer interval. The load estimated by (6.6) remains reasonably low even if instead of T_{m} a factor of $(T_{\mathrm{m}}+T_{\mathrm{f}})$ is used in (6.6) because of the predominance of the work associated with FIB updates. Therefore, it is reasonable to omit the workload represented by LSA aggregation.

GMPLS permits the use of a separate network to transport control plane messages. However, it is assumed now that the topology of the control plane network is the same as that of the data plane network.

The length of the route that any LSA message needs to travel before reaching the most distant node is assumed to be one hop less than the length of the longest one of the shortest paths in between all node pairs in the network ($\hat{H}_{\mathrm{sp}}$). The rationale is that if LSA information is flooded throughout the network, then LSA messages reach any node first along the shortest path in between the failed component and the node. Since it is the neighboring nodes that start disseminating LSA's when a state change occurs, the distance that the LSA's need to travel to any node is one hop less than the distance of the place of the state change.

As a consequence,

$$\hat{E}(T_0) = T_1 + T_2 + \big(\bar{s}(\rho) + T_{\mathrm{f}} + T_{\mathrm{d,max}}\big)(\hat{H}_{\mathrm{sp}} - 1) + \bar{s}(\rho). \tag{6.7}$$

6.2.3 Results

The assumptions used so far are general with respect to routing protocols that maintain a link state database. Nevertheless, the results presented here only demonstrate the performance of OSPF with suitable extensions for traffic engineering (e.g. [23–25]) because the default values of parameters of the estimation match the properties of OSPF.

Throughout the experiments all of the network topologies in appendix A have been used for generating results.

The effect of FIB updates

Router implementations are often engineered so that they limit processing load. An example of this is the `spf_holdtime` parameter in Cisco routers that specifies the minimal

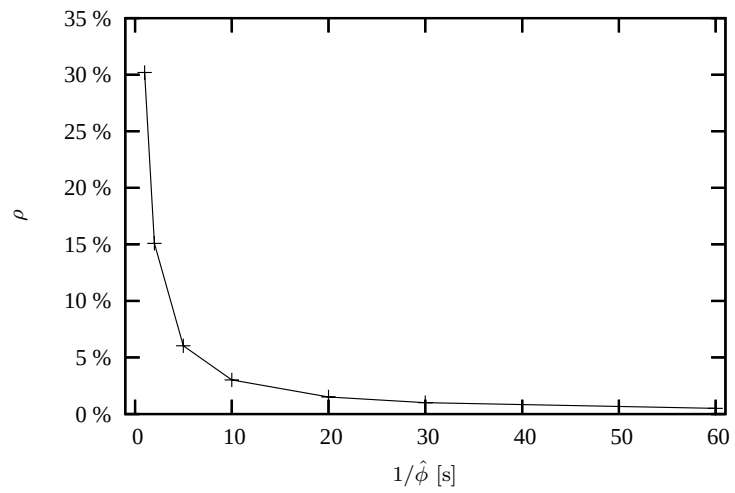


Figure 6.2: Signaling processor load in the EU network as a function of $1/\bar{\phi}$

time in between subsequent shortest path computations [13]. This corresponds to $1/\bar{\phi}$ in our model.

Figure 6.2 shows routing processor load as a function of `spf_holdtime` with $\hat{\lambda} = 0$. The default setting for this parameter is 10 s [13], which corresponds to a processing load of less than 5%. The figure suggests that by means of reasonably setting `spf_holdtime` or any corresponding parameter the additional load corresponding to the last term in (6.6) my be rendered negligible.

Figure 6.3 depicts results obtained with the default setting of `spf_holdtime`. The figure confirms that with this assumption it is still the FIB updates that dominate signaling processor load over a wide range of call request arrival intensities.

Finally, Figure 6.4 shows how the presence of the second term in (6.6) affects the estimated upper bound on $P(A^{(d)})$. The curves obtained with and without considering FIB updates almost perfectly overlap with a relative difference on the order of 0.01%. Hence the rest of the figures only depict the results obtained with FIB updates with the default `spf_holdtime` setting.

The expected difference between the behavior of "traditional" OSPF and OSPF with traffic engineering extensions is that the latter probably requires more processing at nodes. However, the model seems to be fairly insensitive to signaling processor load. Therefore, the results presented above justify that even though the aim is to examine the performance of OSPF with traffic engineering extensions, it is sufficient to set the parameters of the model to values that derive from experiments with "traditional" OSPF.

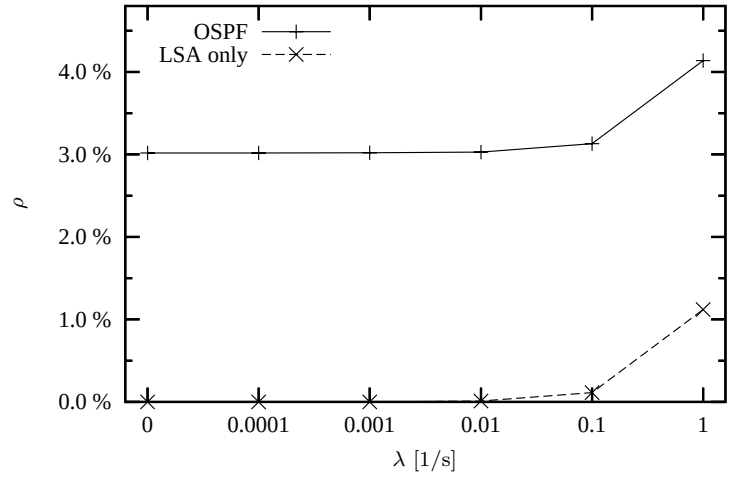


Figure 6.3: Signaling processor load in the EU network as a function of λ

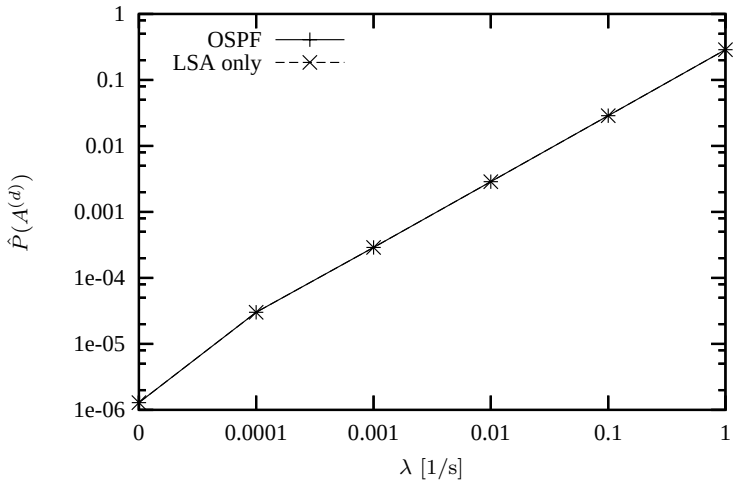


Figure 6.4: Upper bound on $P(A^{(d)})$ in the EU network as a function of λ

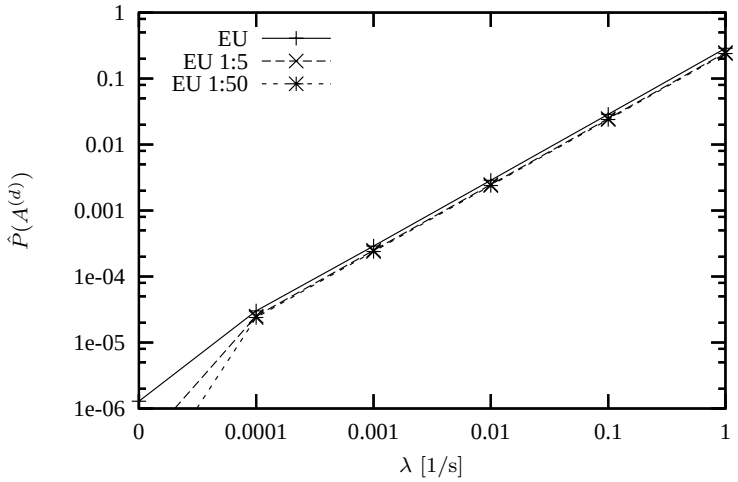


Figure 6.5: Upper bound on $P(A^{(d)})$ in the scaled versions of the EU network as a function of λ

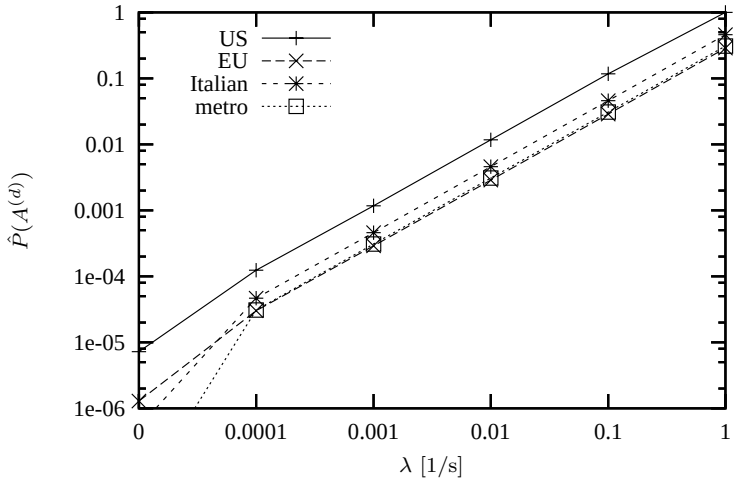


Figure 6.6: Upper bound on $P(A^{(d)})$ in different networks as a function of λ

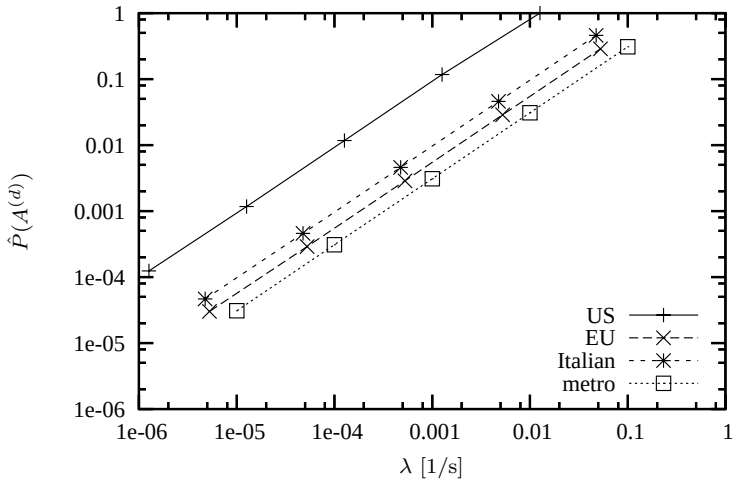


Figure 6.7: Upper bound on $P(A^{(d)})$ in different networks as a function of λ per node

Estimations of the upper bound on $P(A^{(d)})$

Clearly, offered load has a predominant influence on the estimated upper bound of the probability of link state database inconsistency. Compared with the case when only equipment and link failure and repair events are disseminated (c.f. zero arrival intensity on Figures 6.5 and 6.6) the difference may be as large as some orders of magnitude even at modest call request arrival intensities.

Figures 6.5 and 6.6 also shed light on how the estimation depends on the network topology. Scaling the same topology shows that it is not propagation delay that makes a significant difference. However, network scale does matter when no traffic is present, because link failure probabilities are determined by link length. This difference is then diminished as network load grows.

When different topologies are compared, two major factors may be identified that account for high values of $\hat{P}(A^{(d)})$. If the network is sparse, such as the US topology, or its $\hat{H}_{\text{sp}}$ diameter is large, which is the case with the Italian topology, then the estimated probability of inconsistency is higher.

When call arrival intensity is divided by the number of nodes, which reflects the offered "distributed" load, the handicap of the highly sparse US topology becomes more apparent (see figure 6.7).

Nevertheless, one has to note that the presented values are merely estimations on the upper bound of the probability that the link state database is inconsistent with the real betwork state. Consequently, the experienced link state database inconsistency is expected to be less frequent. Another argument in favour of this is that $P(C^{(d)}|A^{(d)}, B^{(d)})$ is undoubtedly less than 1, in contrast to the assumptions of the analysis presented in this brief study. Thus, the probability that the source node is adversely affected by

the link state database inconsistency is clearly expected to be lower than the indicated values.

Note that the presented call arrival intensities are absolute values, that is, they are measured in real-time, whereas call arrival intensities presented in Chapters 4 and 5 are normalized to the mean connection holding time.

6.2.4 Conclusions

To sum up, the offered load has paramount significance with respect to the frequency of link state database inconsistency. If the network traffic is not highly dynamic (overall call arrival intensities of less than 0.001 [1/s]), the results suggest that the effect of link state database inconsistencies is not significant. However, in case of more dynamic traffic a more detailed analysis seems to be necessary.

6.3 Converting availability guarantees to SLA terms and conditions

The provisioning methods that guarantee connection availability presented both in the literature and in this thesis attempt to ensure that asymptotic unavailability of any connection remains above the requirement associated with that connection. However, asymptotic unavailability assumes an infinite observation period, and SLA contracts have finite durations.

A single (asymptotic) unavailability value may be misleading, as both the frequency and the duration of the outages caused by failures is also important. An excellent study pointed out this obvious mismatch between availability values and SLA terms and conditions recently [114].

Consider the following. If failure occurence is modeled as a stochastic point process then the probability that exactly n failures occur ($P(n)$) during the whole SLA contract pedriod (T) that disrupt the connection may be determined for each non-negative integer n. The time of the outage during the SLA contract period is $n \times \mathrm{MTTR}$, if n is known. Based on this argument a statement may be made in the SLA that "the probability that the outage is above $\frac{n \times \mathrm{MTTR}}{T}$ is at most $\sum_{i=0}^{n} P(n)$ over the contract period T" [114].

An example of this is a contract that guarantees "five nines" of availability for one year. It means that the overall downtime during the contract period has to remain under approximately 5.2 minutes. Even if a connection is established using any of the methods proposed in this thesis with a guaranteed unavailability that conforms this requirement, a failure that disrupts the connection causes an outage. Since it is most likely that the repair takes more than the guaranteed maximum downtime, the SLA is violated and thus the service provider is bound to lose some revenues. The method proposed

in [114] offers a solution to avoid such situations by means of properly formulating SLA guarantees.

For further considerations and analysis the reader is referred to [114].

6.4 Summary

This chapter discussed the applicability of the results presented in this thesis. The algorithm presented in Chapter 3 was shown to be general enough to be applied in a wide reliability modeling context.

The chapter also presented a probabilistic analysis of link state database inconsistency, which threatens RWA algorithms that rely on complete knowledge of the current network state. Critical regions of network traffic dynamics were identified assuming that OSPF with suitable extensions for traffic engineering is used for link state information dissemination.

Finally, the pitfalls of converting asymptotic unavailability values to SLA guarantees were discussed.

Chapter 7

Conclusions and Future Work

This chapter gives an overview of the main concepts and contributions presented in the thesis, as well as the results obtained. One application of the results is presented and areas left for future research are identified and briefly stated.

7.1 Summary of contributions

The problem of computing the probability of multiple failure scenarios arises in modern telecommunications networks when service availability have to be evaluated or guaranteed in order to meet terms and conditions of the SLA. The problem is related to the K-terminal reliability problem in reliability theory, which is proven to be hard to solve. Albeit the two problems do not correspond exactly to each other, the latter result suggests that the real-world problem is neither an easy one.

This thesis deals with the efficient computation of multiple failure probabilities. First an algorithm is proposed for computing the failure stratum probabilities in order to support analytical network evaluation methods, such as stratified sampling and adaptive approximation. The proposed algorithm is proven to be more efficient than the one available in the literature for the purpose.

The most significant contributions of this thesis are two on-line connection provisioning methods that provide connection-level availability guarantees with shared (backup) path protection. Previous works on the topic use either non-conservative availability estimation methods or do not consider the dynamic traffic scenario. The proposed methods are based on different connection availability estimation techniques, which are suitable for application in a dynamic traffic scenario and are proven to be conservative.

The first provisioning method is an extension of the SPP-DiR concept to absolute probabilistic guarantees. The extended SPP-DiR method is capable of providing fine-grained levels of connection availability in between unprotected and shared path protected levels with potentially maximal backup sharing. The second method is based on the concept of sharing unavailability, which is defined as an additional unavailability of shared backup resources due to sharing. The sharing unavailability threshold based method coupled with S(B)PP is capable of providing higher availability guarantees than the extended SPP-DiR method, while its blocking probability may remain as low as half of that of dedicated path protection. The choice of the value of the sharing unavailability

threshold is fundamental for the performance of this method, therefore some algorithms are proposed to identify the relevant parameter range.

With help of the extended SPP-DiR method the significance of node failures in an all-optical network is assessed, which is often considered negligible by the literature. Moreover, some computation problems to be solved when applying the proposed methods are defined and proved to be NP-Complete, which justifies the proposal of heuristic solutions.

Additionally, the implementation and applicability of the presented methods are extensively discussed. The importance of the problem of routing based on inconsistent network state information is emphasized, and a theoretical model is presented to estimate the probability of decisions made using inaccurate network state information. With help of the model the performance of OSPF with traffic engineering extensions, a real-world protocol used for maintaining link state databases is evaluated, and a critical real-time call arrival intensity range is identified.

7.2 Application of results

The results presented in Chapter 5 have been applied to obtain results as part of the contribution of Magyar Telekom (formerly Matáv), the Hungarian incumbent telecommunications service provider, to the IST project MUPBED of the 6$^{\text{th}}$ Framework Programme of the European Union [44].

The main goal of the IST project MUPBED is to integrate and validate, in the context of user-driven large-scale testbeds, ASON/GMPLS technology and network solutions as enablers for future upgrades to European research infrastructures [44].

7.3 Future research directions

The work and contributions presented in this thesis may be extended, as some questions are left out of the scope of the study. These potential extensions may serve as grounds and motivations for future work on the subject.

As stated earlier, the most significant contribution of this thesis are two on-line connection provisioning methods. Both of them are presented using shared (backup) path protection, however, they may very well be combined with other resilience schemes. In case of the extended DiR method the application requires practically no modification to the availability bounding method; however, the estimated value of the lower bound on connection availability will not reflect the potentially higher availability provided by other resilience schemes. On the other hand, the sharing unavailability threshold based method does need appropriate modifications, but its bounding technique has the

potential to capture higher availability. The only limitation of the applicability seems to be that at most one backup path may be assigned to protect any working resource.

Another challenge is to discuss if and how the presented methods may be used to provide availability guarantees in networks of different technology. If some artificial bandwidth granularity is introduced, then the extension seems to be straightforward, but it may not be resonable or possible in all cases.

The presented work relies heavily on the accuracy of the asymptotical unavailability values. However, the "perceived" values may change over the lifetime of components, thus statistical averages may not always reflect reality. The sensitivity analysis of the methods to changes in asymptotical unavailability is a relevant and interesting area left for future investigation (and is addressed partly in [73]).

Requirements of applications may include QoS parameters other than blocking probability. However, the current study omitted these parameters, such as recovery time or various physical characteristics of the lightpath. To guarantee QoS with multiple constraints is clearly a challanging open problem for future research.

Bibliography

[1] K. Al-Begain, J. Barner, G. Bolch, and A. I. Zreikat. The performance and reliability modelling language mosel and its application. *International J. on Simulation*, 3(3–4):66–80, 2003.

[2] ALCOA FUJIKURA Ltd. Reliability of fiber optic cable systems: Buried fiber optic cable, optical groundwire cable, all dielectric self supporting cable. Technical report, 2001.

[3] V. Anand, S. Chauhan, and C. Qiao. Sub-path protection: A new framework for optical layer survivability and its quantitative evaluation. Technical report, State University of New York at Buffalo, 2002.

[4] G. Apostolopoulos, R. Guérin, S. Kamat, and S. K. Tripathi. Quality of service based routing: A performance perspective. In *ACM SIGCOMM*, 1998.

[5] D. Arci, G. Maier, A. Pattavina, D. Petecchi, and M. Tornatore. Availability models for protection techniques in wdm networks. In *International Workshop on the Design of Reliable Communication Networks (DRCN)*, 2003.

[6] A. Autenrieth. Recovery time analysis of differentiated resilience in mpls. In *International Workshop on the Design of Reliable Communications Networks (DRCN)*, 2003.

[7] M. O. Ball, C. J. Colbourn, and J. S. Provan. Network reliability. Technical report, University of Maryland at College Park, 1992. TR 92–74.

[8] L. Berger. Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions. RFC3473, 2003.

[9] CA*Net4 news mailing list archive. URL `http://www.canarie.ca/press/lists.html`.

[10] J. Carlier, Y. Li, and J. Lutton. Reliability evaluation of large telecommunication networks. *Discrete Applied Mathematics*, 76:61–80, 1997.

[11] S. Chaudhuri, G. Hjalmtysson, and J. Yates. Control of lightpaths in an optical network. draft-chaudhuri-ip-olxc-control-00.txt, 2000.

[12] K. C.-H. Chu, M. Mezhoudi, and Y. Hu. Comprehensive end-to-end reliability assessment of optical network transports. In *Optical Fiber Communication Conference and Exhibit (OFC)*, 2002.

[13] Cisco Systems, Inc. URL `http://www.cisco.com`.

[14] M. Clouqueur and W. Grover. Availability analysis of span-restorable mesh networks. *IEEE J. on Selected Areas in Communications*, 20(4):810–821, 2002.

[15] M. Clouqueur and W. Grover. Mesh-restorable networks with complete dual failure restorability and with selectively enhanced dual-failure restorability properties. In *Proceedings of OptiComm*, 2002.

[16] D. Commenges and M. Monsion. Fast inversion of triangular toeplitz matrices. *IEEE Transactions on Automatic Control*, AC–29(3):250–251, 1984.

[17] S. Darisala, A. Fumagalli, P. Kothandaraman, M. Tacca, L. Valcarenghi, M. Ali, and D. Elie-Dit-Cosaque. On the convergence of the link-state advertisement protocol in survivable wdm mesh networks. In *Conference on Optical Network Design and Modeling (ONDM)*, 2003.

[18] P. Datta and A. K. Somani. Diverse routing for shared risk resource groups (srrg) failures in wdm optical networks. In *IEEE BroadNets Conference*, 2004.

[19] P. Datta, M. T. Frederick, and A. Somani. Sub-graph routing: A novel fault-tolerant architecture for shared risk link group failures in wdm optical networks. In *International Workshop on the Design of Reliable Communication Networks (DRCN)*, 2003.

[20] W. E. Deming. *Some Theory of Sampling*. Dover Publishers, Inc., 1966.

[21] K. Dohmen. Improved inclusion-exclusion identities and bonferroni inequalities with applications to reliability analysis of coherent systems. Humboldt University, Berlin, Germany, 2000. Habilitation thesis.

[22] J. Doucette, M. Coloqueur, and W. D. Grover. On the availability and capacity requirements of shared backup path-protected mesh networks. *SPIE Optical Networks Magazine*, 4(6):29–44, 2003.

[23] D. K. et al. Traffic Engineering (TE) Extensions to OSPF Version 2. RFC3630, 2003.

[24] K. K. et al. Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC4202, 2005.

[25] K. K. et al. OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC4203, 2005.

[26] European Telecommunications Standards Institute. Network aspects (na); availability performance of path elements of international digital paths. ETSI EN 300 416, 1998.

[27] P. A. Fishwick. *Simulation Model Design and Execution*. Prentice Hall, 1995.

[28] A. Fumagalli and M. Tacca. Optimal design of differentiated reliability (dir) optical ring networks. In *International Workshop on QoS in Multiservice IP Networks (QoS-IP)*, 2001.

[29] A. Fumagalli, M. Tacca, F. Unghváry, and A. Faragó. Shared path protection with differentiated reliability. In *IEEE International Conference on Communications (ICC)*, 2002.

[30] O. Gerstel and G. Sasaki. Quality of protection (qop): A quantitative unifying paradigm to protection service grades. In *SPIE OptiComm*, 2001.

[31] W. Grover. The protected working capacity envelope concept: An alternative paradigm for automated service provisioning. *IEEE Communications Magazine*, 42(1):62–69, 2004.

[32] W. Grover and D. Stamatelakis. Cycle-oriented distributed pre-configuration: ring-like speed with mesh-like capacity for self-planning network restoration. In *IEEE International Conference Communications (ICC)*, 1998.

[33] W. D. Grover. *Mesh-based Survivable Transport Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking*. Prentice Hall, 2003.

[34] M. Halldórsson and J. Radhakrishnan. Greed is good: approximating independent sets in sparse and bounded-degree graphs. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, 1994.

[35] M. Hayashi, K. Ohara, H. Tanaka, M. Daikoku, T. Otani, and M. Suzuki. Highly reliable optical bidirectional path switched ring networks applicable to photonic ip networks. *IEEE J. of Lightwave Techn.*, 21(2):356–364, 2003.

[36] W. He and A. K. Somani. Path-based protection for surviving double-link failures in mesh-restorable optical networks. In *IEEE GLOBECOM*, 2003.

[37] P.-H. Ho and H. T. Mouftah. Slsp: A new path protection shceme for the optical internet. In *Optical Fiber Communication Conference and Exhibit (OFC)*, 2001.

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

[39] L. J. Hornbeck. Digital light processing and mems: An overview. In *Digest of IEEE/LEOS 1996 Summer Topical Meetings, Optical MEMS and Their Applications, WA3*, 1996.

[40] Y. Huang, J. P. Heritage, B. Mukherjee, and W. Wen. Availability-guaranteed service provisioning with shared-path protection in optical wdm networks. In *Optical Fiber Communications Conference and Exhibit (OFC)*, 2004.

[41] R. Inkret, A. Kuchar, and B. Mikac. Advanced infrastructure for photonic networks. Extended final report of COST action 266, 2003.

[42] International Electrotechnical Commission. Reliability data handbook — universal model for reliability prediction of electronics components, pcbs and equipment. IEC/TR 62380, 2004.

[43] International Electrotechnical Commission. Mathematical expressions for reliability, maintainability and maintenance support terms. IEC/TR 61703, 2000.

[44] IST MUPBED "Multi-Partner European Testbeds for Research Networking". URL `http://www.ist-mupbed.org`.

[45] IST NOBEL "Next generation Optical network for Broadband European Leadership". URL `http://www.ist-nobel.org`.

[46] M. Jaeger and R. Huelsermann. Service availability of shared path protection in optical mesh networks. In *European Conference on Optical Communication (ECOC)*, 2004.

[47] L. Jereb. Efficient reliability modeling and analysis of telecommunication networks. In *Proceedings of the 6th International Conference on Telecommunication Systems (ICTS)*, 1998.

[48] V. Jimenez and A. Marzal. Computing the k shortest paths: a new algorithm and an experimental comparison. In *Proceedings of the 3rd Workshop on Algorithm Engineering*, 1999.

[49] C. R. Johnson, Y. Kogan, Y. Levy, F. Saheban, and P. Tarapore. Voip reliability: A service provider's perspective. *IEEE Communications Magazine*, 42(7):48–54, 2004.

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

[51] A. Kodian and W. Grover. Failure independent path-protecting p-cycles: efficient and simple fully pre-connected optical-path protection. *IEEE J. of Lightwave Technology*, 23(10):3241–3259, 2005.

[52] A. Kodian and W. Grover. Multiple quality of protection classes including dual-failure survivable services in p-cycle networks. In *IEEE BroadNets Conference*, 2005.

[53] J. Levendovszky, L. Jereb, Z. Elek, and G. Vesztergombi. Adaptive statistical algorithms in network reliability analysis. *Elsevier Performance Evaluation*, 48 (1–4):225–236, 2002.

[54] L. Y. Lin, E. L. Goldstein, and R. W. Tkach. Free-space micromachined optical switches with submillisecond switching times for large-scale optical crossconnects. *IEEE Photonics Technology Letters*, 10(4):525–527, 1998.

[55] H. Liu, E. Bouillet, D. Pendarakis, N. Komaee, J.-F. Labourdette, and S. Chaudhuri. Distributed route computation algorithms and dynamic provisioning in intelligent optical mesh networks. *IEEE J. on Selected Areas in Communications*, 22(9):1626–1639, 2004.

[56] S. D. Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz. Pan-european optical transport networks: an availability based comparison. *Photonic Network Communications*, 5(3):203–225, 2003.

[57] M. Malhotra and K. S. Trivedi. Power hierarchy of dependability model types. *IEEE Transactions on Reliability*, 43(3):493–502, 1994.

[58] E. Mannie. Generalized Multi-Protocol Label Switching (GMPLS) Architecture. RFC3945, 2004.

[59] X. Masip. *Mechanisms to Reduce Routing Information Inaccuracy Effects: Application to MPLS and WDM Networks.* PhD thesis, Universitat Politécnica de Catalunya, 2003.

[60] D. A. A. Mello, J. U. Pelegrini, R. P. Ribeiro, D. A. Schupke, and H. Waldman. Dynamic provisioning of shared-backup path protected connections with guaranteed availability requirements. In *IEEE BroadNets Conference*, 2005.

[61] D. A. A. Mello, D. A. Schupke, M. Scheffel, and H. Waldman. Availability maps for connections in wdm optical networks. In *International Workshop on the Design of Reliable Communications Networks (DRCN)*, 2005.

[62] D. A. A. Mello, D. A. Schupke, and H. Waldman. A matrix-based analytical approach to connection unavailability estimation in shared backup path protection. *IEEE Communications Letters*, 9(9):844–846, 2005.

[63] G. Mohan and A. Somani. Routing dependable connections with specified failure restoration guarantees in wdm networks. In *IEEE Infocom*, 2000.

[64] P. Monti, M. Tacca, and A. Fumagalli. Resource-efficient path-protection schemes and online selection of routes in reliable wdm networks. *OSA J. of Optical Networking*, 3(4):188–203, 2004.

[65] E. Motamedi, M. C. Wu, and K. S. J. Pister. Micro-opto-electro-machanical devices and on-chip optical processing. *Optical Engineering*, 36(5):1282–1297, 1997.

[66] J. Moy. OSPF Version 2. RFC2328, 1998.

[67] B. Mukherjee. *Optical Communication Networks.* McGraw-Hill, 1997.

[68] D. Oran. OSI IS-IS Intra-domain Routing Protocol. RFC1142, 1990.

[69] C. Ou, J. Zhang, H. Zang, L. H. Sahasrabuddhe, and B. Mukherjee. New and improved approaches for shared-path protection in wdm mesh networks. *IEEE J. of Lightwave Technology*, 22(5):1223–1232, 2004.

[70] Z. Pándi. Analysis of public trouble ticket data. Technical report, Budapest University of Technology and Economics, Department of Telecommunications, 2005. URL `http://cntic90.hit.bme.hu/ zspandi/publ/2005/tech_report.pdf`.

[71] Z. Pándi and Á. Gricser. Availability analysis of shared protection schemes for on-line connection provisioning. In *Proceedings of the IV Workshop in G/MPLS Networks*, 2005.

[72] Z. Pándi and Á. Gricser. Analysis of the trade-off between availability and backup resource sharing. In *IEEE International Conference on Transparent Optical Networks (ICTON)*, 2005.

[73] Z. Pándi and Á. Gricser. Improving connection availability by means of backup sharing restrictions. to appear in OSA J. of Optical Networking, 2006.

[74] Z. Pándi and L. Wosinska. On temporary inconsistency of the link state database with prompt update policies. In *IEEE International Conference on Transparent Optical Networks (ICTON)*, 2005.

[75] Z. Pándi, M. Tacca, and A. Fumagalli. Impact of oxc failures on network reliability. In *Proceedings of the SPIE Photonics Europe Conference, Strasbourg, France*, 2004.

[76] Z. Pándi, M. Tacca, and A. Fumagalli. Efficient computation of multi-component failure stratum probabilities. *IEEE Communications Letters*, 9(10):939–941, 2005.

[77] Z. Pándi, M. Tacca, and A. Fumagalli. A threshold based on-line rwa algorithm with reliability guarantees. In *Conference on Optical Network Design and Modeling (ONDM)*, 2005.

[78] Z. Pándi, M. Tacca, A. Fumagalli, and L. Wosinska. Dynamic provisioning of availability-constrained optical circuits in the presence of optical node failures. submitted to IEEE J. of Lightwave Technology, 2005.

[79] M. Pióro and D. Mehdi. *Routing, Flow and Capacity Design in Communication and Computer Networks*. Elsevier, 2004.

[80] J. S. Provan and M. O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. of Computing*, 12:777–788, 1983.

[81] S. Ramamurthy, L. Sahasrabuddhe, and B. Mukherjee. Survivable wdm mesh networks. *IEEE J. of Lightwave Technology*, 21(4):870–883, 2003.

[82] R. Ramaswami and K. Sivarajan. *Optical Networks: A Practical Perspective (Second Edition)*. Morgan Kaufmann Publishers, 2001.

[83] RealityGrid Consortium. URL `http://www.realitygrid.org/SPICE/`.

[84] R. A. Sahner, K. S. Trivedi, and A. Puliafito. *Performance and Reliability Analysis of Computer Systems: An Example-Based Approach Using the SHARPE Software Package*. Kluwer Academic Publishers, 1996.

[85] E. Salvadori, R. L. Cigno, and Z. Zsóka. Dynamic grooming in ip over optical networks based on the overlay architecture. submitted to Optical Switching and Networking, 2005.

[86] D. A. Schupke. The tradeoff between the number of deployed p-cycles and the survivability to dual fiber duct failures. In *IEEE International Conference on Communications (ICC), Anchorage, AK, USA*, 2003.

[87] D. A. Schupke. Multiple failure survivability in wdm networks with p-cycles. In *International Symposium on Circuits and Systems (ISCAS)*, 2003.

[88] D. A. Schupke. The tradeoff between the number of deployed p-cycles and the survivability to dual fiber duct failures. In *IEEE International Conference on Communications (ICC)*, 2003.

[89] D. A. Schupke, C. G. Gruber, and A. Autenrieth. Optimal configuration of p-cycles in wdm networks. In *IEEE International Conference on Communications (ICC)*, 2002.

[90] D. A. Schupke, W. D. Grover, and M. Clouqueur. Strategies for enhanced dual failure restorability with static or reconfigurable p-cycle networks. In *IEEE International Conference on Communications (ICC)*, 2004.

[91] A. Shaikh and A. Greenberg. Experience in black-box ospf measurement. In *1st ACM SIGCOMM Workshop on Internet Measurement*, 2001.

[92] G. Shen and W. D. Grover. Extending the p-cycle concept to path segment protection for span and node failure recovery. *IEEE J. on Special Areas in Communication, Optical Communications and Networking Series*, 21(8):1306–1319, 2003.

[93] D. Shier. *Network Reliability and Algebraic Structures*. Clarendon Press, New York, NY, USA, 1991.

[94] L. Song, J. Zhang, and B. Mukherjee. Dynamic provisioning with reliability guarantee and resource optimization for differentiated services in wdm mesh networks. In *Optical Fiber Communications Conference and Exhibit (OFC)*, 2005.

[95] J. Szigeti, I. Ballók, and T. Cinkler. Efficiency of information update strategies for automatically switched multi-domain optical networks. In *IEEE International Conference on Transparent Optical Networks (ICTON)*, 2005.

[96] M. Tacca. *Differentiated Reliability in Wavelength Division Multiplexing Networks*. PhD thesis, Department of Electrical Engineering, University of Texas at Dallas, 2002.

[97] M. Tacca, A. Fumagalli, and F. Unghváry. Double-fault shared path protection scheme with constrained connection downtime. In *International Workshop on the Design of Reliable Communication Networks Conference (DRCN)*, 2003.

[98] M. Tacca, P. Monti, and A. Fumagalli. The disjoint path-pair matrix approach for online routing in reliable wdm networks. In *IEEE International Conference on Communications (ICC)*, 2004.

[99] Telcordia Technologies. Reliability prediction procedure for electronic equipment. Special report SR–332, 2001.

[100] M. To and P. Neusy. Unavailability analysis of long-haul networks. *IEEE J. on Selected Areas in Communication*, 12(1):100–109, 1994.

[101] I. Tomkos, D. Vogiatzis, C. Mas, I. Zacharopoulos, A. Tzanakaki, and E. Varvarigos. Performance engineering of metropolitan area optical networks through impairment constraint routing. *IEEE J. of Optical Communications*, 42(8):S40–S47, 2004.

[102] M. Tornatore, G. Maier, and A. Pattavina. Availability design of optical transport networks. *IEEE J. on Selected Areas in Communications*, 23(8):1520–1532, 2005.

[103] I. T. Union. Architecture for the automatically switched optical network (ason). ITU G.8080/Y.1304, 2001.

[104] L. Valcarenghi and A. Fumagalli. The preplanned weighted restoration scheme. In *IEEE Workshop on High Performance Switching and Routing*, 2001.

[105] R. W. Wolff. Poisson arrivals see time averages. *Operations Research*, 30(2): 223–231, 1982.

[106] L. Wosinska. *A Study of the Reliability of Optical Switching Nodes for High Capacity Telecommunications Networks (TRITA-MVT report 1999:4)*. PhD thesis, Royal Institute of Technology, Stockholm, 1999.

[107] L. Wosinska. Reliability study of fault-tolerant multiwavelength nonblocking optical cross connect based on ingaasp/inp laser-amplifier gate-switch arrays. *IEEE Photonics Technology Letters*, 5(10):1206–1209, 1993.

[108] L. Wosinska, L. Thylen, and R. P. Holmstrom. Large-capacity strictly nonblocking optical cross-connects based on microelectrooptomechanical systems (meoms) switch matrices: Reliability performance analysis. *IEEE J. of Lightwave Techn.*, 19(8):1065–1075, 2001.

[109] T.-H. Wu. *Fiber Network Service Survivability*. Artech House, 1992.

[110] D. Xu, C. Qiao, and Y. Xiong. An ultra-fast shared path protection scheme — distributed partial information management — part ii. In *IEEE International Conference on Networking Protocols*, 2002.

[111] D. Xu, Y. Xiong, and C. Qiao. New algorithms for shared segment protection. *IEEE J. on Selected Areas in Communication*, 21(8):1320–1331, 2003.

[112] J. Zhang, K. Zhu, B. Mukherjee, and H. Zang. Service provisioning to provide per-connection-based availability guarantee in wdm mesh networks. In *Optical Fiber Communications Conference and Exhibit (OFC)*, 2003.

[113] J. Zhang, K. Zhu, H. Zang, and B. Mukherjee. A new provisioning framework to provide availability-guaranteed service in wdm mesh networks. In *IEEE International Conference on Communications (ICC)*, 2003.

[114] L. Zhou and W. D. Grover. A theory for setting the "safety margin" on availability guarantees in an sla. In *International Workshop on the Design of Reliable Communication Networks (DRCN)*, 2005.

Appendix A

Network Topologies

The purpose of this appendix is to collect network topology data used throughout the thesis along with the respective references. A summary of the characteristic parameters of the topologies is also included.

A.1 Graphic data

In the thesis topologies of continental, national and metropolitan scale are also used to test the performance of the proposed methods.

Two topologies of continental scale are used: a European WDM network topology [96] and a US WDM network topology [96] depicted on Figures A.2 and A.1, respectively. The former one also appears in experiments with its link lengths scaled to 1:5 and 1:50 in order to represent a network of national and metropolitan scale, respectively. The rationale is to make comparisons between networks of different scale easier.

The topology of national scale considered here is that of the Italian WDM network [17], which is shown on Figure A.3.

Due to the difficulties of obtaining realistic topology data of metropolitan WDM networks the topology of metropolitan scale used in the study is taken from [101]. It is a metropolitan topology created for demonstrative purposes, which has realistic parameters. The topology is shown on Figure A.4.

A.2 Overview of characteristic parameters

The parameters of interest to the discussion of the thesis are number of links, nodes and node degrees, as well as shortest path hop lengths, and link length data. They are listed broken up into Table A.1 and Table A.2 due to typesetting reasons.

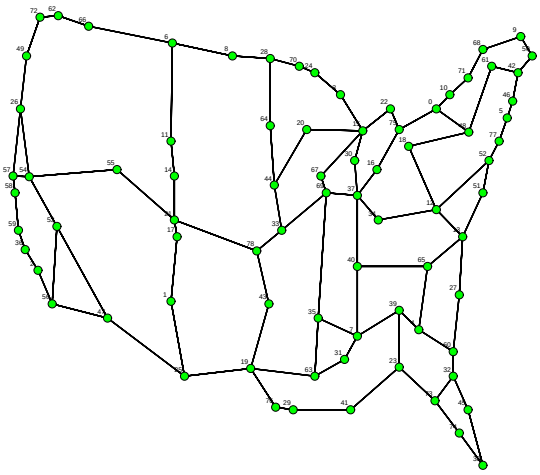


Figure A.1: US WDM network topology

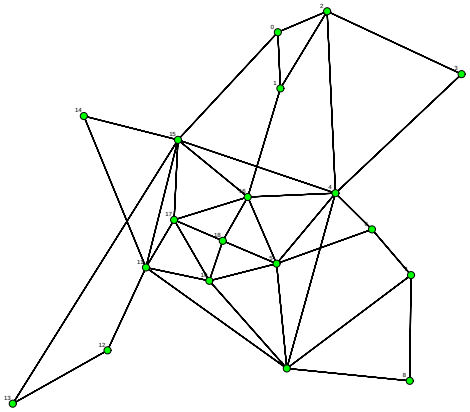


Figure A.2: European WDM network topology

Topology	Nodes	Bidirectional links	Avg. node degree	Max. node degree
US	79	102	2.6	5
EU 1:1	19	39	4.1	7
Italian	21	36	3.4	6
EU 1:5	19	39	4.1	7
metro	10	16	3.2	4
EU 1:50	19	39	4.1	7

Table A.1: Summary of parameters of network topologies — part I

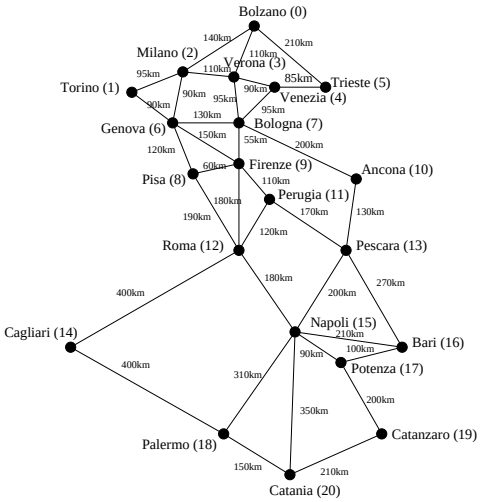


Figure A.3: Italian WDM network topology

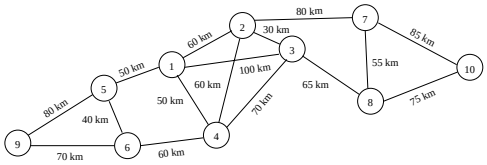


Figure A.4: Metropolitan WDM network topology

Topology	Longest shortest path [hop]	Average fiber length [km]	Longest fiber [km]	Total fiber sheath length [km]
US	16	324.565	914	33106
EU 1:1	4	644.615	1650	25140
Italian	7	163.750	400	5895
EU 1:5	4	128.923	330	5028
metro	5	64.375	100	1030
EU 1:50	4	12.892	33	503

Table A.2: Summary of parameters of network topologies — part II

Appendix B

Details of the Simulator

This appendix discusses the details of the event-driven simulator implemented for experimental testing of the methods proposed in Chapter 4 and Chapter 5. The purpose is to detach these details from the body of the thesis in order to collect them as a reference and to complete the documentation of the experiments.

B.1 General description

The simulator takes the network topology, the λ call generation rate and the $r^{(d)} = r$ availability requirement of calls as input, and it generates random calls in the network. Call generation is a Poisson process with parameter λ. Source and destination nodes of calls are selected uniformly and the same availability requirement is considered for all the generated calls. The duration of established connections is exponentially distributed with a mean of one time unit.

A single-slot central buffer is used. Call requests wait in this buffer until enough resources are freed in the network to serve them. If the single slot buffer is already occupied, newly generated call requests are rejected.

Without the buffer connection requests between closer nodes would be favored, that is, they would have a higher chance to be served successfully as opposed to requests between distant nodes. This would yield an 'average' blocking probability that would not apply uniformly to all possible end node pairs.

In this study, the system is fair if all source and destination pairs experience the same blocking. If the call generation process is Poisson, then the single-slot central buffer ensures fairness as a consequence of the Poisson Arrivals See Time Averages (PASTA) theorem [105].

The number of rejected and total call requests is recorded, and the ratio of the two gives the call blocking probability, which is used to characterize network performance.

The rationale of using the central buffer is to facilitate comparison of different methods. Obviously, it is not absolutely necessary, because benchmarking conditions may be defined in any arbitrary way. However, the present study uses this assumption because results obtained this way show a very picturesque property of the call admission control process, which is straighforward to interpret and compare.

In what follows, methods used for determining routing and wavelength assignment throughout the simulations are reviewed.

B.2 Storage of candidate paths

Path computations are carried out off-line, before the simulation starts. Paths are stored in a disjoint path-pair matrix (DPM) structure proposed in [64, 98]. The DPM stores for each source and destination node pair the shortest k_1 paths as working path candidates and at most k_2 protection path candidates for each of the working path candidates.

The DPM for a given network topology is obtained as follows. For each source and destination node pair the following steps are repeated. A k-shortest path algorithm [48] is run in order to find the shortest k_1 loopless paths for the current source and destination node pair. These are the working path candidates. For each working path candidate at most k_2 backup path candidates are found by running the k-shortest path algorithm on the graph which is obtained by removing the edges used by the current working path candidate from the original topology.

The size of the DPM in the simulations is set to $k_1 = 20$ and $k_2 = 5$. Note that k_1 and k_2 only limit the size of the path matrix and it is not guaranteed that $k_1 \times k_2$ path pairs actually exist for each source and destination node pair.

B.2.1 Interesting properties of the DPM

In spite of being simple, the DPM structure is remarkable for some of its properties, which justifies its application in the simulations.

Firstly, based on the link cost function used to determine path lengths different path ordering schemes are available. The cost function may assign unit cost to each link, which then results in a smallest-hop-first order of the paths in the DPM. However, it is also possible to assign availability related costs to links, such as $-\log(p_e)$, which in turn yields a highest-availability-first order.

In this thesis the cost function assigned to link lengths when building the DPM is $-\log(p_e)$.

Secondly, due to the fact that paths are ordered in the DPM, one may easily guarantee certain properties of the selected path pairs. For example a hop limit may be used to limit restoration time.

Thirdly, the DPM structure inherently supports the use of segment-based protection schemes, as well. To see this the closure property of a DPM is defined as follows [71, 73].

Definition 7. *The DPM is* closed *if for any p working path candidate each $p' \subset p$ segment of p can be found in the DPM as a working path candidate between the end nodes of p'.*

Computing the paths as described in [64, 98] does not yield a closed DPM in general. In order to close the DPM, once the DPM is constructed each missing segment of all working path candidates is also added to the matrix with disjoint backup path candidates computed for them. Then the missing segments of newly added working path candidates are also inserted with their backup path candidates and so on. Algorithm 6 gives a more formal description.

Algorithm 6 DPM closure

```
1: continue = true
2: while continue do
3:     continue = false
4:     for all segments of all paths in the DPM do
5:         if current path segment is not an entry in the DPM then
6:             add current path segment to the DPM
7:             add potential disjoint paths to the new entry
8:             continue = true
9:         end if
10:    end for
11: end while
```

The described process yields a closed DPM after only a few rounds, as in each round the longest path segment to be added to the matrix will be at least one hop shorter than the longest path segment added to the matrix in the previous round.

An alternative approach is to complement the DPM with the missing segments in an on-demand manner as the simulation evolves. However, if the closure is achieved by means of the discussed algorithm in advance, there is no need for path computations during connection admission.

In case of the European WDM network used in this study (c.f. Figure A.2) only about 7% of the segments of working path candidates are missing from the initial DPM. As only path-based protection schemes are considered in the study, the closure is not implemented in the simulator and is only considered here for the sake of completeness.

B.3 Routing and wavelength assignment

Routing and wavelength selection are carried out jointly by means of running a meta-heuristic to find an optimal solution in a solution space that consists of all potential candidates. The meta-heuristic selected for the purpose is the well-known simulated annealing [79]. The parameters of the annealing are selected based on an extensive set of experiments [64], which aimed at finding a trade-off between simulation time and the optimality of the solution.

The quantity to be minimized is the blocking probability for a long sequence of successive demands. However, the optimization is run demand-by-demand minimizing a given cost function due to the dynamic traffic scenario considered (c.f. section 2.4.1). One may think of the process as doing a greedy or local optimization in search for a global optimum. It is, therefore, a challenge to define a cost function so that the overall blocking will be as low as possible.

When defining an appropriate cost function two approaches may be followed. The first one is to try to evaluate the resource usage and the extent of overprovisioning in terms of availability as presented in [64]. The second one is to try to influence routing so that the routes less frequently used by others are preferred using the idea presented in [104].

The cost functions used during the optimization of the RWA are presented along with the simulation results in the thesis. Note that the range of the cost function should be more or less the same as that of the one used for finding the best parameters of the annealing.