



**EUROPEAN COOPERATION  
IN THE FIELD OF SCIENTIFIC  
AND TECHNICAL RESEARCH**

# Advanced Infrastructure for Photonic Networks

---

EXTENDED FINAL REPORT OF COST ACTION 266

Robert Inkret, Anton Kuchar, Branko Mikac



# Advanced Infrastructure for Photonic Networks

<http://www.ure.cas.cz/dpt240/cost266/index.html>



# Advanced Infrastructure for Photonic Networks

---

Extended Final Report of COST 266 Action

Edited by

**Robert Inkret, Anton Kuchar, Branko Mikac**

**Copyright © 2003 by Management Committee of  
the project COST 266**

**All rights reserved.**

No part of this publication may not be reproduced by  
any means, or transmitted or translated into machine  
language without permission of the Management  
Committee of the project COST 266.

Published by:

Faculty of Electrical  
Engineering and Computing,  
University of Zagreb

Unska 3, HR-10000 Zagreb, CROATIA

Prepared by Robert Inkret  
Cover design by Robert Inkret  
Printed by ITG d.o.o., Zagreb

**ISBN 953-184-064-4**

CIP - Katalogizacija u publikaciji  
Nacionalna i sveučilišna knjižnica - Zagreb

UDK 621.38/.39(082)

ADVANCED infrastructure for photonic  
networks : extended final report of COST  
266 action / edited by Robert Inkert...  
<et al.>. - Zagreb : Faculty of  
Electrical, Engineering and Computing,  
University of Zagreb, 2003.

Bibliografija.

ISBN 953-184-064-4

1. Inkret, Robert  
I. Optički prijenosni sustavi -- Zbornik

430917031

**Printed in Croatia**

# Contents

<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 COST 266 Action Objectives	1
1.2 Overview of Results Presented in this Extended Final Report	3
1.3 COST 266 Reference Scenario	19
1.3.1. Introduction	19
1.3.2. Network Scenarios	19
1.3.3. Traffic Model	21
1.4 References	23
<b>Chapter 2 Transmission Systems</b>	<b>25</b>
2.1 Introduction	25
2.2 N×40 Gbit/s Transmission Performance	27
2.2.1. Theoretical Background on N×40 Gbit/s Transmission Systems and Simulation Model	27
2.2.2. Single Channel Transmission at 40 Gbit/s	31
2.2.3. N×40 Gbit/s Transmission	33
2.2.4. System Performance in the Presence of Raman Amplification	35
2.3 DWDM in Metropolitan Area Networks	37
2.4 Optical Add/Drop Multiplexers and Optical Cross-Connects for Wavelength Routed Networks	40
2.4.1. Optical Add/Drop Multiplexers	41
2.4.2. Optical Cross-Connects	43
2.4.3. System Performance Discussion	46
2.5 All-Optical 3R Signal Regeneration	48
2.5.1. Clock Recovery and 3R Regeneration Schemes	49

2.5.2. Basic Devices for 3R-Regenerator	49
2.5.3. Clock Recovery for Flexible and Asynchronous Networks	51
2.5.4. Test of All-Optical 3R Regenerators in Loop Experiments	51
2.5.5. Novel 3R-Regenerator Scheme	52
2.5.6. Summary	53
2.6 Algorithm to Assign the Lightpaths and Example for an Optical Transport European Network	53
2.7 Cross Gain Saturation in Erbium Amplifiers	58
2.8 Conclusions	63
2.9 References	64
<b>Chapter 3 Circuit/Wavelength Switching and Routing</b>	<b>71</b>
3.1 Introduction	71
3.2 Control Plane Architectures for IP-over-WDM Networks	72
3.2.1. Control Plane Architectures for Optical Networks	72
3.2.2. Dynamic Provisioning over Multi-Provider Interconnected Networks	74
3.2.3. Comparing OPNNI versus GMPLS	76
3.3 Designing Optimised Reliable WDM Networks	78
3.3.1. Static Wavelength Assignment in WDM Networks with OXCs	78
3.3.2. Routing in Optical Networks under Inaccurate Network State Information	79
3.3.3. Protection in the Optical Layer for Mesh Networks	81
3.3.4. Multi-Ring Networks	85
3.3.5. Uncertainty in the Planning Process	85
3.4 Designing IP over WDM Networks	87
3.4.1. An Approach to an IP-over-WDM Transmission Network Modelling	87
3.4.2. Static Lightpath Routing in IP over WDM Networks	88
3.4.3. Design of Optimal Data Link Layer Topology for IP/WDM Networks	91



3.4.4. Dynamic Routing and Grooming in SDH/WDM Multi-Layer Networks	93
3.4.5. Multi-Layer Traffic Engineering in Data-Centric Optical Networks	95
3.4.6. Influence of Traffic Asymmetry on the Network Design	98
3.5 Providing Optical VPN Services	100
3.5.1. Overview	100
3.5.2. The Three Types of OVPNs Considered	102
3.5.3. Three Methods for Setting up OVPNs Optimally	103
3.5.4. Protection Alternatives for OVPNs	104
3.6 Conclusions/Summary	104
3.7 Issues Open for Further Research	107
3.8 References	108
<b>Chapter 4 Optical Burst Switching and Optical Packet Switching</b>	<b>115</b>
4.1 Optical Burst Switching and Optical Packet Switching Concepts	115
4.1.1. Packet/Burst Handling Schemes	117
4.1.2. OPS/OBS Properties and Comparison	118
4.1.3. Summary	120
4.2 Design and Analysis of OPS Nodes	120
4.2.1. OPS and OBS Nodes	121
4.2.2. General Design Considerations	122
4.2.3. Optical Technology Status	123
4.2.4. Packet/Burst Switch Architectures	125
4.2.5. Summary	127
4.3 Contention Resolution	127
4.3.1. Motivation and Overview	127
4.3.2. Wavelength Conversion	128
4.3.3. FDL Buffer Architectures	129
4.3.4. Electronic Buffering	130
4.3.5. Comparative Performance Evaluation	131

4.3.6. Summary	133
4.4 Quality of Service in OBS/OPS	133
4.4.1. Introduction	133
4.4.2. Quality of Service in Optical Burst Switching	134
4.4.3. Quality of Service in Optical Packet Switching	136
4.4.4. QoS in MPLS Optical Networks	139
4.5 Optical Packet Switching in Metro Networks	141
4.5.1. O-MAN Properties	142
4.6 Conclusion and Outlook	147
4.7 References	149
<b>Chapter 5 Availability Analysis</b>	<b>155</b>
5.1 Introduction	155
5.2 Availability Performance Analysis	156
5.2.1. Description of Analysed Network Architectures	156
5.2.2. Availability Model	160
5.2.3. Network Modelling Procedure	162
5.2.4. Results	163
5.3 Comparison of Pan-European Optical Backbone Networks in Terms of Expected Loss of Traffic	169
5.3.1. Introduction	169
5.3.2. Case Study	170
5.3.3. Conclusions	173
5.4 Instead of Conclusion – Further Work	173
5.5 References	174
<b>Chapter 6 Tools</b>	<b>175</b>
6.1 Introduction	175
6.2 The Cosmos Tool	175
6.2.1. Motivation and Idea	175

6.2.2. Main Capabilities	178
6.2.3. Cosmos Tool Structure	178
6.2.4. Cosmos Kernel Structure	179
6.2.5. Nyx Structure	180
6.2.6. Topology Description	183
6.2.7. Behaviour Description	186
6.2.8. Communication Mechanisms	187
6.2.9. Simulation Mechanisms	188
6.2.10. Examples	190
6.3 Computer Aided Network Planning Cockpit CANPC	192
6.3.1. Introduction	192
6.3.2. The CANPC Framework	192
6.3.3. Applications and Algorithms	194
6.3.4. The Extensions Manager	197
6.3.5. Conclusion	198
6.4 Other Tools	198
6.4.1. WDM Guru	199
6.5 Conclusion	201
6.6 References	202
<b>Chapter 7 Evolution of Next-Generation Optical Networks</b>	<b>205</b>
7.1 Introduction	205
7.2 A Technology Perspective	206
7.2.1. Introduction	206
7.2.2. Evolution of Circuit-Switched Networks	206
7.2.3. Gradual Introduction of Optical Burst/Packet Switching	207
7.3 An Economical Perspective	210
7.4 Optional Step: ORION	211
7.4.1. Principle	212
7.4.2. Node Architecture	213

7.4.3. Case Study	215
7.4.4. Migration Options	217
7.5 Conclusions – Opportunities for Future Research	218
7.6 References	219
<b>Chapter 8 Summary of Results and Future Work</b>	<b>221</b>
8.1 Summary of Results Achieved in COST 266 Action	221
8.2 Future Work - Proposal for New COST Action: Towards Digital Optical Networks	233
8.2.1. Introduction	233
8.2.2. Objectives	234
8.2.3. Technical Programme	235

## Contributors

Elise Baert {elise.baert@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Stefan Bauer {bauer@hhi.fraunhofer.de}  
Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, Berlin, Germany

Steinar Bjørnstad {steinar.bjornstad@fysel.ntnu.no}  
Internet Network Architecture, Telenor R&D, Oslo, Norway  
Department of Physical Electronics, Norwegian University of Science and Technology,  
Trondheim, Norway

Erik Van Breusegem {erik.van.breusegem@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Cristian Bungarzeanu {cristian.bungarzeanu@epfl.ch}  
Laboratoire des Télécommunications, EPF – Lausanne, Lausanne, Switzerland

Franco Callegati {fcallegati@deis.unibo.it}  
DEIS - University of Bologna, Bologna, Italy

Davide Careglio {careglio@ac.upc.es}  
Advanced Broadband Communications Centre (CCABA), Dept. Computer Architecture,  
Universitat Politècnica de Catalunya (UPC), Barcelona, Catalunya, Spain

Walter Cerroni {wcerroni@deis.unibo.it}  
DEIS - University of Bologna, Bologna, Italy

Jan Cheyns {jan.cheyns@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Tibor Cinkler {cinkler@tnt-atm.ttt.bme.hu}  
High-Speed Networks Lab, Department of Telecommunications & Telematics, Budapest  
University of Technology & Economics, Budapest, Hungary

Didier Colle {didier.colle@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Piet Demeester {piet.demeester@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Chris Develder {chris.develder@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Vincenzo Eramo {eramo@infocom.uniroma1.it}  
Infocom Department, University of Rome “La Sapienza”, Rome, Italy

Christoph M. Gauger {gauger@ikr.uni-stuttgart.de}  
Institute of Communication Networks and Computer Engineering (IKR), University of Stuttgart,  
Stuttgart, Germany

Dag R. Hjelme {dag.hjelme@fysel.ntnu.no}  
Department of Physical Electronics, Norwegian University of Science and Technology,  
Trondheim, Norway

Robert Inkret {inkret@tel.fer.hr}  
Department of Telecommunications, Faculty of Electrical Engineering and Computing,  
University of Zagreb, Zagreb, Croatia

Gabriel Junyent {junyent@tsc.upc.es}  
Advanced Broadband Communications Centre (CCABA), Dept. Signal Theory and  
Communications, Universitat Politècnica de Catalunya (UPC), Barcelona, Catalunya, Spain

Miroslav Karasek {karasek@ure.cas.cz}  
Institute of Radio Engineering and Electronics, Czech Academy of Sciences, Prague, Czech  
Republic

Mirosław Klinkowski {mklinkow@itl.waw.pl}  
National Institute of Telecommunications, Department of Transmission and Fiber Technology,  
Warsaw, Poland

Martin Köhn {koehn@ikr.uni-stuttgart.de}

Institute of Communication Networks and Computer Engineering (IKR), University of Stuttgart,  
Stuttgart, Germany

Marian Kowalewski {m.kowalewski@itl.waw.pl}

National Institute of Telecommunications, Department of Transmission and Fiber Technology,  
Warsaw, Poland

Anton Kuchar {kuchar@ure.cas.cz}

Institute of Radio Engineering and Electronics, Czech Academy of Sciences, Prague, Czech  
Republic

Marko Lacković {marko.lackovic@tel.fer.hr}

Department of Telecommunications, Faculty of Electrical Engineering and Computing,  
University of Zagreb, Zagreb, Croatia

Marije Ljolje {marije@tel.fer.hr}

Department of Telecommunications, Faculty of Electrical Engineering and Computing,  
University of Zagreb, Zagreb, Croatia

Sophie De Maesschalck {sophie.demaesschalck@intec.UGent.be}

Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Marian Marciniak {m.marciniak@itl.waw.pl}

National Institute of Telecommunications, Department of Transmission and Fiber Technology,  
Warsaw, Poland

Carmen Mas {cmas@ait.gr}

Athens Information Technology Center (AIT), Athens, Greece

Xavier Masip-Bruin {xmasip@ac.upc.es}

Advanced Broadband Communications Centre (CCABA), Dept. Computer Architecture,  
Universitat Politècnica de Catalunya (UPC), Barcelona, Catalunya, Spain

Francesco Matera {mat@fub.it}

Fondazione Ugo Bordononi (FUB), Rome Italy

Mario Mattiello {mario.mattiello@epfl.ch}

Laboratoire des Télécommunications, EPF – Lausanne, Lausanne, Switzerland

Christian Mauz {mauz@nari.ee.ethz.ch}  
Communication Technology Laboratory, ETH Zurich, Zurich, Switzerland

Branko Mikac {branko.mikac@fer.hr}  
Department of Telecommunications, Faculty of Electrical Engineering and Computing,  
University of Zagreb, Zagreb, Croatia

Hans-Peter Nolting {nolting@hhi.fraunhofer.de}  
Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, Berlin, Germany

Martin Nord {mn@com.dtu.dk}  
Internet Network Architecture, Telenor R&D, Oslo, Norway  
Research Center COM, DTU, Lyngby, Denmark

Mario Pickavet {mario.pickavet@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Bart Puype {bart.puype@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Carla Raffaelli {craffaelli@deis.unibo.it}  
DEIS - University of Bologna, Bologna, Italy

Sergio Sánchez-López (sergio@ac.upc.es)  
Advanced Broadband Communications Centre (CCABA), Dept. Computer Architecture,  
Universitat Politècnica de Catalunya (UPC), Barcelona, Catalunya, Spain

Bernd Sartorius {sartorius@hhi.de}  
Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, Berlin, Germany

Dominic A. Schupke {schupke@lkn.ei.tum.de}  
Institute of Communication Networks, Munich University of Technology, Munich, Germany

Marina Settembre {marina.settembre@eri.ericsson.se}  
Ericsson Lab, Rome, Italy



Josep Solé-Pareta {pareta@ac.upc.es}  
Advanced Broadband Communications Centre (CCABA), Dept. Computer Architecture,  
Universitat Politècnica de Catalunya (UPC), Barcelona, Catalunya, Spain

Norvald Stol {norvald.stol@item.ntnu.no}  
Department of telematics, Norwegian University of Science and Technology, Trondheim,  
Norway

Slobodanka Tomić {slobodanka.tomic@tuwien.ac.at}  
Institute of Communication Networks (IKN), TU Wien, Vienna, Austria

Ioannis Tomkos {itom@ait.gr}  
Athens Information Technology Center (AIT), Athens, Greece

Anna Tzanakaki {atza@ait.gr}  
Athens Information Technology Center (AIT), Athens, Greece

Sofie Verbrugge {sofie.verbrugge@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Qiang Yan {qiang.yan@intec.UGent.be}  
Department of Information Technology (INTEC), Ghent University – IMEC, Ghent, Belgium

Ioannis Zacharopoulos {izac@ait.gr}  
Athens Information Technology Center (AIT), Athens, Greece

Paolo Zaffoni {pzaffoni@deis.unibo.it}  
DEIS - University of Bologna, Bologna, Italy

THIS PAGE HAS BEEN LEFT BLANK INTENTIONALLY

## Chapter 1

# Introduction

*Anton Kuchar, Christian Mauz, Didier Colle, and Sophie De Maesschalck*

This report is the last one summarizing the most important results achieved within the framework of COST 266 action. It is organized in seven chapters.

**Chapter 1** – Introduction, presents

- The objectives and organizational structure of COST 266 Action
- Overview of work done by COST 266 partners during the lifetime of the Action described in this report
- COST 266 Reference Scenario of a model pan-European network referred to in the following chapters.

### 1.1 COST 266 Action Objectives

COST 266 Action *Advanced infrastructure for photonic networks* started in February 1999. Its main objective was:

*“to propose and evaluate suitable architectures and identify the key photonic components for the infrastructure of next generation photonic networks and work out alternative scenarios for the evolution towards next generation networks“* [1].

To achieve this objective, three working groups have been formed:

WG1 - *Physical limitations of photonic transport*, dealing with:

- Study on dimensioning tools for photonic networks by updating COST 239 results.
- Theoretical analysis of optical transmission limitations including a comparison of simulation tools for a given transmission path consisting of fibre links, optical amplifiers and optical network nodes. The impact of the limitations on implementation of a large optical transport layer.
- Optical signal processing and regeneration in the network nodes - modelling of all-optical 3R (Re-amplification, Re-timing, and Re-shaping) regenerator subsystems.
- Monitoring in the optical layer.

WG2 - *Network elements of advanced photonic infrastructure*, covers the following issues:

- Definition of the functionality requirements for the optical network nodes and the elements required to provide these functions.
- Development of strategies for the transport of variable length packets within an optical telecommunication layer.
- Development of a novel Internet Protocol over Wavelength Division Multiplexed (IP/WDM) network architecture enabling implementation of differentiated quality of service (QoS) by utilising the wavelength.
- Study of alternative MAN network architectures employing dense WDM (DWDM).
- Computation of availability of DWDM networks and study of protection mechanisms in optical networks.

WG3 - *Network evolution and the role of photonic infrastructure in the next generation networks*, is devoted to:

- Studies on the impact of photonic technologies on the architecture of next generation networks.
- Network dimensioning and development of planning routines.
- General architectural issues of photonic networks such as division of functionality between the electrical and optical domains and between the network edge and the network core, etc.
- Convergence of standard telecommunication networks and computer networks.

- Coexistence of the current networks with the next generation networks.

There was a substantial amount of overlap between the WGs.

The specific *secondary objectives* referred to the target photonic network architecture that should:

- a) provide virtually unlimited end-to-end bandwidth regardless of distance enabling instant movement of huge amounts of data for applications such as information retrieval, downloading (often multimedia) Web software, exchange of CAD-size software (hundreds of Mbytes) and data models (Gbytes, as part of remote network-enabled co-operative design, etc.). enabled efficient and simple multicasting and broadcasting of broadband signals,
- b) be flexible and agile in order to accommodate abrupt and unpredictable traffic surges and to react instantaneously when failures occur, allowing seamless and errorless transmission of large data blocks,
- c) be transparent to all sorts of digital signals and protocols so that it would enable to accommodate a number of network types and services simultaneously or in succession deployed over time, and
- d) trade bandwidth for simplicity to alleviate the complexity problem in the communication networks resulting mainly from the ever increasing amount of software residing in the networks that raises costs and lowers reliability of the networks.

## 1.2 Overview of Results Presented in this Extended Final Report

Advanced photonic networks will support traffic between its network nodes of the order of many Tbit/s. To fully exploit the potential of optical fibres, Dense Wavelength Division Multiplexing (DWDM) will be utilized as the transmission technique with the individual wavelengths carrying tens to hundreds of Gbit/s. Such huge data streams will have to be processed at the network nodes semi- or all-optically by means of some sort of Optical Cross-Connects (OXC) and Optical Add/Drop Multiplexers (OADM). Ultra-long links consisting of many sections of fibre and the photonic network devices will introduce physical limitations due to signal attenuation, dispersion, interchannel interference and various non-linear phenomena. These will have to be countered by optical amplification, dispersion compensation and full (3R) digital signal regeneration.

At present, the DWDM 40 Gbit/s/wavelength ( $N \times 40$  Gbit/s) transmission systems are being field tested and are close to commercial deployment but

many issues remain to be resolved especially if the most common G.652 optical fibres are to be fully exploited. In wavelength routing the lightpath assignment will not only depend on the traffic requirements among the nodes but the transmission impairments will have to be taken into account, as well.

In **Chapter 2** the results on transmission performance achieved in the framework of COST 266 Action are reported. The maximum achievable capacity of the  $N \times 40$  Gbit/s G.652 fibre links are investigated with the aim to facilitate the design of long-haul transport networks. The return-to-zero (RZ) pulses are only considered since they are superior over the non-return-to-zero (NRZ) and other signal formats. Periodical compensation of chromatic dispersion at the output of each fibre span with and without a pre-chirp at the link input and Erbium Doped Fibre Amplifier (EDFA) or hybrid Raman/EDFA optical amplification is assumed. The results of this investigation are reported in terms of the maximum realizable number of 40 Gbit/s channels versus distance and applied to wavelength assignment to create lightpaths in a transport network according to traffic requirements. A novel algorithm for the design of lightpaths taking into account the physical impairments of the fibre links using the technique of removing the lightpaths is presented. As an example, this algorithm is applied to a model pan-European network consisting of 26 nodes studied in the framework of the European COST 266 project as a reference. The performance of DWDM systems in metro and short haul applications by exploiting polarization modulation has also been studied from the physical limitations point of view.

Another issue treated in Chapter 2 is the cross-gain modulation effect in EDFAs considering either the WDM circuit-switching networks in which the number of channels present in an EDFA may vary due to network reconfiguration or channel failure or the WDM packet-switching burst-mode networks where no power is transmitted during empty slots. The cross-gain saturation in the fibre amplifiers induces power transients in the surviving channels causing severe service impairments. Concrete results of experimental as well as theoretical investigation of the effect of the cross-gain saturation on transmission of packetized Ethernet data at 10 Mbit/s over three wavelength channels in a cascade of five EDFAs are presented.

Another part of Chapter 2 is devoted to the important issue of all-optical 3R regeneration describing two novel semiconductor devices - the self-pulsating laser for clock extraction and the semiconductor optical amplifier (SOA) for switching. Mach-Zehnder Interferometer (MZI) with SOAs has been used as a fast optical gate. A novel all-optical alternating data clock 3R regenerator architecture has been proposed and demonstrated combining regeneration with

a flexible wavelength conversion which is an important advantage for application in WDM systems. All-optical clock recovery is especially difficult problem in multi-rate digital systems and asynchronous networks, where, for example, ultra fast locking to asynchronous IP packets is needed. The self-pulsating laser is shown to be a promising candidate to solve this problem offering a continuous broad tuning range covering several SDH/Sonet hierarchies up to 40 Gbit/s.

**Chapter 3** is devoted to network level aspects of circuit switched optical networks. In the optical network nodes complete wavelength channels are switched. Four main topics are treated in this chapter.

One of the main trends in the evolution of optical networks is towards Intelligent Optical Networks (IONs). The intelligence in the (optical) transport networks is expected to be introduced in the form of a distributed control plane.

The first topic of Chapter 3 concerns control plane architectures for IP over optical networks considering multi-layer as well as multi-domain inter-working schemes. The possibility of integrating this control plane with the control plane of the client networks (i.e., typically IP/MPLS networks) is investigated. Advantages and disadvantages of integrating/separating the two control planes are pointed out. The peer (full integration), the overlay (no integration) and the augmented (a compromise between the two) control plane models are considered.

To ensure dynamic global connectivity, the control framework of the optical network will have to account not only for inter-layer (client IP-layer over server optical layer) but also for inter-domain and inter-provider inter-working problems. Therefore, due attention is paid in this chapter to the issue of dynamic provisioning over multi-provider interconnected GMPLS-capable networks. The automatic provisioning over multiple domains requires that all the involved domains have to have some notion of the available network resources. Advanced infrastructure for photonic networks will need novel operational scenarios, e.g. such that enable dynamic bandwidth trading. The dynamic on-demand provisioning may be done at the points of domains' inter-connections giving rise to a novel interconnection Multi-Provider Edge (MPE) architecture employing GMPLS-enabled cross-connects (MPE nodes) to which clients and domains will be connected taking part in both the policy-based path calculation and the set-up of requested end-to-end connections. Two key functionalities of an MPE node are (i) provisioning of on-demand interconnections and (ii) routing or control mediation. Thereby MPE, as a multi-layer node, may be able to provide interconnections on the different layers. With the MPE

overlay acting as a middle tear between providers and customers, and the providers themselves, MPE can act as a trusted mediator.

A modified version of the Private Network-to-Network Interface (PNNI) is studied as an alternative to the Generalized Multi-Protocol Label Switching (GMPLS) scheme. PNNI is expected to be suitable for Automatic Switched Optical Networks (ASON) after some appropriate modifications. As a mature technology, the PNNI can be very practical for a seamless migration from current transport networks to ASON. The Optical PNNI (O-PNNI), as an adaptation of the ATM PNNI, has been proposed as an alternative control plane to the one provided by GMPLS. O-PNNI fits very well with the overlay model separating both the IP and OTN control planes.

The second topic of Chapter 3 concerns the design of the optical layer of WDM networks focused on Routing and Wavelength Assignment (RWA) problems taking into account not just the static traffic but also dynamic and uncertain traffic demands. In addition, the possibility of incorporating network recovery schemes in the RWA is considered.

Wavelength assignment is a major dimensioning step in designing large and highly loaded WDM networks. Therefore, efficient algorithms and tools able to perform this operation are needed. Two wavelength assignment algorithms have been considered - TABUCOL that has been conceived at TCOM within COST 266 Action and the Minimum Hop RWA proposed in another project. Both of them perform routing and wavelength assignment separately. In this way, the exact minimum number of fibres, for a given routing scheme, can be calculated. It is pointed out that, because of the optimal assignment, the use of wavelength converters would not have reduced the total number of used fibres.

In agile ASONs with frequent state of network changes it is better to set up and tear down lightpaths in a distributed manner based on local information supplied to the control plane. Therefore, since maintaining precise global network state information on each node is almost impossible, a new adaptive source routing mechanism that computes dynamically explicit lightpaths in an ASON without conversion capabilities has been proposed. Its aim is to reduce the connection blocking probability due to performing routing and wavelength assignment decisions under inaccurate global network state information.

In mesh networks, the path or span protection can be realized in the optical layer. After giving classification of common protection schemes in mesh networks it is shown that for all these mesh protection schemes a unified formulation as an ILP-problem can be adopted with the objective to minimize the required capacity of the network. When the routing of the working paths is



included in the optimisation process of the protection resources, the capacity requirements for a survivable network may be significantly reduced. Results of a case study for the different protection schemes using CPLEX 7.0 to solve the corresponding ILP-problems are presented in Chapter 3.

In some situations it is advantageous to organize transport networks in rings employing the simpler and cheaper optical add-drop multiplexers (OADM) as network elements. However, planning of multi-ring-networks with several interweaved rings may be very complex. In rings, the protection resources may be dedicated or shared. Within COST 266 Action an optimised algorithm based on graph-theory enabling to deduce an optimal ring-coverage has been developed and is described in this report.

Since the communication networks are becoming increasingly agile with high dynamics, a systematic treatment of the vagueness of the traffic pattern is required. The uncertainty of the future traffic can be modelled by entering random variables in the traffic matrix with appropriate distribution. Then, the probabilities for the realization of a specific connection may be calculated. By combining several statistically independent traffic sources, the relative variance and, therefore, the required transmission capacity can be minimized. The results of a case study of the COST 266 pan-European model network for different distributions are shown.

In the third part of this chapter, the modelling, design and dimensioning of multi-layer IP over optical WDM networks that include intermediate layers like SDH/SONET is treated. Several techniques to build the logical network topology are investigated considering also the dynamic and asymmetric nature of the traffic in IP networks. Several data link layer topologies are compared, then routing and grooming in IP transport networks are studied. Traffic engineering concepts for dynamic WDM networks are presented.

An approach to an IP-over-WDM transmission network modelling using two main node types - IP Point of Presence (IPPoP) and WDM Point of Presence (WDMPoP) - is described. An IPPoP node aggregates IP traffic from many sources and lower levels of aggregation generating IP traffic for the WDM transmission network. A WDM network is composed of WDMPoP nodes interconnected with optical fibres serving as ingress and egress points, also providing switching functionality as intermediate nodes inside the WDM network.

In the following part of this chapter a solution to the problem of static light-path routing in IP over WDM networks is proposed. In IP-over-WDM networks the IP routers are interconnected by lightpaths with the aid of OXCs

bypassing the ATM or SDH layers. To find an optimal virtual topology seen by the IP layer and a routing and wavelength assignment (RWA) on the WDM layer a feasible and efficient (static) configuration of the lightpaths is needed. Due to the complexity of the problem it is split into three sub-problems solved by heuristic algorithms. The aim is to minimize the number of lightpaths. Wavelength conversion at the WDM nodes is not assumed, therefore, for each fibre the wavelength constraint has to be fulfilled. The described approach is applied to a case study.

The design of optimal data link layer topology for IP/WDM networks, i.e., the connectivity of IP backbone routers, is another key issue dealt with in this chapter. It is assumed that the physical topology, i.e., the length and placement of the optical fibres, is given. The traffic matrix at the data link layer is also known and the traffic demands in this layer are routed along the shortest paths in the optical network layer. Assuming that the number of fibres on the links is unlimited, three scenarios to obtain a data link layer topology are presented: 1-to-1 mapping between the physical and data link topologies, full mesh data link layer topology and an optimal topology. Two most popular scenarios for mapping IP on WDM are considered: IP/POS/WDM and IP/GbE/WDM. The described algorithm has been applied to a network sample taken from the case study developed within COST 239 and COST 266.

One possible evolution path for IP transport networks is dynamic SDH network on top of a dynamic optical network (IP/SDH/WDM). In this architecture low-bandwidth electrical connections are groomed and then assigned to high-speed optical lightpaths. Efficient transport of dynamic traffic demands requires optimised multi-layer routing and grooming algorithms. Non-integrated routing schemes treat both layers separately while integrated schemes try to improve the performance by combining both layers. Results of a simulation study using a fictitious 9-node network presented in this report show that the impact of routing and grooming algorithms on the performance of multi-layer networks is not negligible. They reduce the blocking probability in the network as well as increase the usage of the wavelength channels.

Intelligent optical networks enable to set up or tear down the lightpaths dynamically. They then support logical IP topology in multi-layer data-centric optical networks. Multi-layer traffic engineering allows reconfiguring the logical resources. The logical network layer will use signalling to request new lightpaths from the optical layer or to remove existing lightpaths. A good traffic engineering strategy needs to integrate both multi- as well as single-layer traffic engineering strategy aspects. Issues in designing multi-layer traffic engineering strategies and benefits of multi-layer traffic engineering are highlighted in this chapter by a case study.

It is well known that IP traffic is asymmetric the degree of which depends on the location in the network. When this IP traffic is imposed on the current networks large portions of the network bandwidth may in one direction be under-utilized while at the opposite direction it can be congested. Using unidirectional WDM line-systems instead of the currently in use bi-directional ones could thus lead to a significant cost reduction. In this chapter, this cost advantage is quantified for three of the COST 266 optical backbone topologies described later in the report.

In the last part of Chapter 3 a classification of Optical Virtual Private Networks (OVPNs) and some general methods for setting them up optimally with various additional aspects including protection and grooming are presented. OVPNs are a special kind of VPN built over optical channels, e.g., wavelength paths using optical cross-connects. The advantages of traffic grooming can be exploited. VPNs can be dynamically reconfigured or re-dimensioned. Three Types of OVPNs are considered in this chapter depending on how the lightpaths are used by concurrent VPNs. The goal is to configure the VPNs and the lightpath system optimally without separating the network layers. Three methods for setting up OVPNs optimally and two alternatives to protect them are described - at the higher (electrical) and at the lower (wavelength) layer.

**Chapter 4** deals with optical burst switching and optical packet switching. Truly agile networks require sub-second reaction time to sudden traffic surges. This is virtually impossible to achieve in circuit-switched networks. Moreover, the telecommunication networks are becoming increasingly data-centric. Therefore, the future telecommunication networks should be optimised for carrying bursty data traffic with utmost granularity. Optical packet switching (OPS) is the technology best fit to fulfil such requirements. At the same time, however, it is the most difficult one to realize. Optical burst switching (OBS) is regarded as a compromise and a bridge between the two extremes. As a result, due attention has been paid to it in COST 266 Action.

In optical packet switching as well as in optical burst switching forwarding is done separately from switching in the network nodes. Forwarding decisions are made by means of a packet header or a burst control packet, respectively. Since forwarding involves a lot of signal processing, in most proposed schemes - including those presented in this report - it is supposed to be realized electronically. On the other hand, the burst or packet ultra-high speed payload is expected to be switched all-optically to avoid the electronic bottleneck presented by the current electronic network nodes.

Although requirements placed on individual modules of OPS and OBS nodes may differ, the basic principles for switching architectures and functionalities are similar in both cases, therefore, a generic packet/burst handling schemes are treated simultaneously. Four potential packet/burst handling schemes classified according to synchronisation and size of the data units are introduced in the report. This enables the authors to suggest the application areas for OPS and OBS, respectively. All possible variations of these two technologies are considered. Further on, WDM mesh networks are assumed for application of the two schemes.

Functionality of ingress and egress edge routers and of core routers is explained in order to compare the various concepts proposed in the literature and by the authors. The differences discussed concern node realisations, contention resolution and QoS differentiation.

In this report, OPS node design assumes IP packets 40 to 1500 bytes long that are representative of today's Internet traffic. OPS then require a rather low degree of client layer packet assembly but the level of aggregation needed for the same packet duration increases with the bitrate. More extensive burst aggregation to realise bursts with payloads typically carrying tens of kbytes is performed in OBS schemes. While in OPS systems the control information is typically encoded *in-band* (on the same wavelength), in OBS the control header containing the timing information on the burst is sent *out-of-band* on a separate wavelength enabling advanced burst scheduling including QoS differentiation.

This chapter describes the main functionalities of OPS and OBS nodes and includes some design considerations. A brief overview of technology status is given, before presenting OPS node architectures proposed within COST 266.

A generic OPS/OBS node is introduced composed of four main building blocks - (1) the input interface that detects a preamble/synchronisation pattern and marks the packet arrival enabling clock recovery, (2) the electronic control unit that makes look-ups in the routing table and is responsible for implementing the scheduling policy, (3) the switch matrix characterized by its switching time, maximum throughput, internal blocking properties and signal degradation, and (4) the output interface that updates the OPS headers or the burst control packets and conditions the signal.

The OPS/OBS networks are expected to provide to their client layers throughput in the Tbit/s range. Among others it depends on the packet/burst switch architecture/designs and is a function of node degree, load and traffic characteristics. In realistic network scenarios external blocking or *contention* occurs,

leading to loss of packets/bursts. Apart from acceptable loss rates, the packet/burst switch should support QoS differentiation. The delay of an OPS/OBS node is negligible compared to the transmission delay in wide area networks whereas jitter caused e.g. by buffering may be significant since it may lead to reordering of packets/bursts increasing the egress node complexity. The inevitable overhead and optical guard bands decrease the bandwidth efficiency depending on payload packets duration and the switching time. To keep it at a reasonable level some aggregation at the OPS ingress nodes is required.

The OPS/OBS require rather advanced optical as well as electronic devices. A brief overview of the technology status of relevant devices is given including the electronic control and header processing units and types of optical switch fabrics suitable for OPS/OBS.

In the section on packet/burst switch architectures, results of performance simulation of an internally blocking OPS switch employing arrayed waveguides and wavelength converters at its input are given. Slotted as well as asynchronous mode of operation is considered. Then, as an example of non-blocking OPS architecture, an original (*feedback*) switch design employing tunable input wavelength converters, arrayed waveguides, shared electronic buffers and fixed output wavelength converters suitable for asynchronous packet switching is described in detail.

Contention resolution is an important issue in OPS/OBS networks since they rely on statistical multiplexing in order to achieve good utilisation of network resources in the presence of bursty traffic. Here, temporary overload situations called contention situations occur that have to be resolved otherwise burst or packet loss occurs. Such a contention situation can be resolved in the wavelength, time or space domain. In this report contention resolution in wavelength and time domain is discussed and results of a joint comparative performance evaluation are presented. Most node OPS/OBS designs assume full I/O wavelength conversion, however, to reduce the cost, using shared converter pools have been proposed. For contention resolution in the time domain, as a rule, the use of fibre delay lines (FDLs) is assumed. In a DWDM network, contention resolution may also exploit the wavelength domain by sharing the wavelength pool of a fibre and then by transmitting contending packets on different wavelengths. When a packet needs to be forwarded to an output fibre specified in the routing table the wavelength and delay selection problem (WDS) to implement a just enough time (JET) based QoS differentiation scheme arises. Several resource allocation policies for the case of asynchronous variable-length optical packets have been proposed and are considered in

the report. In OBS, the JET reservation scheme has been studied in detail in the framework of COST 266 Action. It offers the flexibility to reserve time intervals to newly arriving bursts in gaps left by already reserved bursts. Thus, JET provides another solution for the problem of gaps induced by FDL buffers. Also, regarding the WDS, the sequence in which wavelength conversion and buffering are applied can be exploited to trade off wavelength converter and FDL buffer usage.

Since optical memory with random access in the time-domain is still immature, FDLs are used as optical buffers in OPS/OBS schemes. As an alternative to FDLs, the use of simple electronic FIFO memory with few opto-electronic interfaces is suggested and treated in detail in the report. In a common evaluation scenario, the impact of different WDS algorithms and individual FDL delays in an FDL buffer is compared for OBS and OPS both assuming asynchronous operation and variable length bursts or packets. Then, both approaches are compared to OPS with electronic buffers based on the number of buffer interfaces. Burst/packet loss probability versus normalized buffer delay, number of buffer ports and number of electronic interfaces when using electronic buffering for various numbers of wavelengths has been studied by simulation. A conclusion is made that both FDL and electronic buffers can significantly reduce burst/packet loss probability when used together with wavelength conversion.

The IP protocol, regarded as the converging protocol layer, does not itself support the quality of service (QoS) differentiation. Therefore, when implemented, the OBS/OPS layers should be able to provide QoS differentiation to the IP-layer. This will depend on resources available in the network nodes like buffering and wavelength converters.

Service differentiation in OBS is achieved by allowing early reservation of high priority bursts by assigning them an extra offset time called QoS offset. Therefore, high priority bursts have a lower loss probability than the low priority ones. The impact of the QoS offset on the burst loss probability of the high priority class is described in the report and pros and cons of the offset-based schemes are discussed.

In OPS nodes, where it is not possible to change the order of packets coming out of the delay lines, some form of reservation of the resources managed by the WDS policies must be applied, i.e. the available wavelength and delay, in order to privilege one traffic class over the other. The time-threshold-based and the wavelength-based techniques have been investigated applied to the minimum gap queue WDS policy described in the report leading to two new QoS-oriented policies which use the time-threshold-based and the wavelength-

based technique respectively. Assuming two different classes of QoS in an OPS node employing shared electronic buffer the packet loss has been evaluated by simulations when the buffer resources were divided into two different blocks of inputs quantifying the share of traffic to the buffer inputs reserved for the two classes.

Before QoS can be implemented in the optical domain in the OPS/OBS networks, it can be based on Multi Protocol Label Switching (MPLS) that offers resource reservation with proper control algorithms. Within the optical node, the optical packet label is read and compared with a look up table. The payload is then transparently routed in the optical domain. In a QoS capable packet network the MPLS controller performs packets classification in edge nodes, differentiated packets servicing in core nodes and traffic engineering. The label switched paths are established across the optical network reserving resources such as buffers or wavelengths over predetermined paths for service differentiation. Currently, two main MPLS features for supporting QoS are used, i.e., differentiated servicing (packets are classified at the edge of the network into classes of service - CoS) and traffic engineering (data traffic is routed on the various links and nodes in the network in order to balance the traffic load).

One of the trends in optical networking is migration of the packet-based technologies from access networks to Metropolitan Area Networks (MANs). A packet-based transport technology is regarded as a natural fit with the now ubiquitous IP protocol offering better cost/efficiency trade-off in the metro environment with the rapidly increasing data traffic volume than the existing transport infrastructures such as SONET/SDH and ATM. Therefore, researchers world-wide are making great efforts to propose new MAN architecture solutions. Apart from solutions based on the more traditional approaches such as the Resilient Packet Ring, OPS at the optical level that can provide high throughput is being investigated as a good candidate for future optical MAN (O-MAN) architectures. In this report the main requirements for O-MANs are identified that are used to compare five different O-MANs architectures including one solution investigated within the COST 266 action and another one studied in co-operation with the IST project DAVID.

Due to wide-scale deployment of DWDM-based all-optical networks carrying huge amount of traffic, the aspect of resilience to the various failures have become of utmost importance. **Chapter 5** on network reliability/availability starts with an overview of considered network architectures and protection switching schemes. In this report a generic term called *network survivability* is introduced expressing the level of resilience to network failures. Three groups

of survivability measures for quantitative analysis are considered, namely, probability-based (*availability*), traffic-based (e.g. *expected loss of traffic*), and topology-based measures (e.g. *connectivity*, *diameter*, etc.). When designing all-optical network architecture, it is very important to pinpoint optical components, devices and even architectural choices that are critical to network survivability. Then, various network protection and restoration (P&R) architectures or concrete networks can be compared. Within the framework of COST 266 Action the probability-based and the traffic-based survivability measures have been studied. For this purpose the COST 266 case study topologies (*Reference scenario*) were used.

Two types of WDM optical networks are considered in the model used for availability analysis - passive (no automatic switching or rerouting of wavelength paths) and those enabling automatic circuit switching in the network nodes by establishing wavelength paths or virtual wavelength paths across the networks. All connections between any two nodes are assumed symmetrical, i.e., lightpaths in both directions use the same links and nodes.

1+1 protection of the passive optical networks composed of WDM multiplexers, de-multiplexers, optical amplifiers, cables and fibres is considered first. It is concluded that survivability of such networks after any single failure is ensured when elements of any two lightpaths do not share the same physical elements, i.e., the lightpaths are placed onto two link-independent paths.

Automatically switched WDM networks employing optical non-blocking switches are treated next. The 1+1 protection is in this case realised by the management software that manages allocation of network resources. For survivability after a single failure each lightpath is assigned an additional spare node-independent lightpath. In the 1:1 path protection scenario the spare path is used exclusively in the case when there is a failure on the primary path. Under normal conditions the spare path can be used for providing lower priority lightpaths. In the 1: $N$  case the network elements allocated to the spare lightpaths can be shared by a number of primary lightpaths.

Restoration, as the most efficient P&R mechanism in terms of utilisation of network resources, is also discussed. The restoration time can be kept acceptable by pre-computation of the spare lightpaths whenever a demand is routed through the network. A new spare lightpath is computed only if the primary or spare lightpaths are affected by a failure.

An availability model is then developed and availability results for transport entities calculated from the availabilities of network components are presented. Exponential failure and repair times and the knowledge of the mean time to failure (MTTF) and to repair (MTTR) for each component are



assumed. The transport entities including the wavelength channel, the lightpath and a logical connection are characterised in detail and the availability for each connection in the network is calculated.

The network modelling procedure is described step by step. These steps produce a WDM network capable of supporting the chosen P&R mechanism using selected network architecture. This is followed by a network analysis that can include equipment statistics, cost calculation and availability calculation by analytical calculation and Monte Carlo simulation. Extensive results for the COST 266 case study topologies and different P&R architectures are presented. In all calculations, 16 wavelengths per fibre, no wavelength conversion, amplifier spacing of 80 km and traffic matrices estimated for the year 2002 have been assumed.

Since requirements concerning availability of the connections for voice differ from those placed on providing the IP-based best effort services, there is an opportunity for network operators to save costs by applying different recovery schemes for voice and transaction data on one hand and IP traffic on the other hand. In the last part of this chapter a comparison is made between (a) the use of 1+1 dedicated optical protection for all traffic and (b) 1+1 protection for voice and transaction data traffic while IP data is routed unprotected. The COST 266 pan-European reference network is used for the case study. The basic reference topology, the ring topology and the triangular topology are compared in terms of the network design integrity. The average expected loss of traffic (AELT), i.e., the average amount of traffic that is expected to be lost each year due to failures, is used to express the availability of the network. It is concluded that the changing traffic pattern, characterized by the growing importance of Internet traffic, has a serious influence on the AELT. The recovery scheme used for the IP traffic has quite a big impact. Transporting IP traffic as a best-effort service is much cheaper than protecting this type of traffic, but it results in a much lower availability of this service than when this traffic would be 1+1 protected. The ring topology design is the most cost-efficient one while the triangular topology gives the best values for the AELT. Therefore, a strategic trade-off between network installation cost and availability of the connections has to be made.

Due to increasing complexity of problems faced by the telecom engineers, more sophisticated design, analysis and simulation aids are needed. Large number of simulators intended for various purposes have been developed. **Chapter 6** describes some network tools created in the framework of COST 266 Action primarily with the aim to help in the design and analysis of various options of advanced photonic telecommunication networks considered during

the Action. Two major network simulation tools are described in detail in this chapter - the COSMOS and the CANPC simulation packages.

The COSMOS tool allows a uniform network description and creation of network and optimisation algorithm libraries. It is an all-purpose tool enabling to solve a wide range of problems including modelling, simulation and optimisation processes of diverse systems ranging from physical effects in opto/electronic and optical devices to telecommunication networks with the focus on optical networks. The tool is highly efficient due to its uniform approach to network design using three structural elements - the modules, layers and the system. It is simple to use yet flexible enough allowing user modifications that can be done by redefining the functions of the simulation mechanism. The user code is portable between various computer platforms as well as different processor architectures. This feature has been accomplished by using freely obtainable C++ compilers for different platforms and keeping the code structure within ANSI C++ specifications. This portability includes just the console part of the tool restricting the GUI to Windows platform. The tool is fully programmed using C++ object-oriented programming language on the MS Windows based platform thus providing great flexibility. All parts of the tool, from the simulation mechanisms to the graphic interface, are coded in the same language.

As a result of the object-oriented paradigm (OOP) encapsulation feature implemented in the tool, the simulated systems can be described by their features and functionality. The OOP classes (general templates) and the objects (classes' instances) consist of the attributes and functionality (methods). A class hierarchy, with the more abstract classes at the top and the more specialized ones representing a particular system in more detail residing at the bottom, results in the so called *inheritance* where each derived (inherited) class contains all functionalities and features of the basic class. Inheritance is the basis of the OOP code reusability because the more abstract classes are expected to be used in different model development processes sharing the same basic model (class).

The tool enables structure (topology) description, behaviour description, simulation based on the network structure and its model and optimisation based on network algorithms and/or simulation results.

The COSMOS tool comprises the following parts: cosmos kernel, user libraries, analytical algorithms, Nyx tool and a graphical user interface (GUI). The kernel allows a user network description. It provides basic elements of network topology and the framework for behaviour description. The topology description defines a passive network structure, while the behaviour descrip-

tion defines a simulation model. Analytical and optimisation algorithm used in network structuring can be directly parameterised from the GUI. The Nyx tool consists of three modules: the problem module, optimisation module and the user interface module. The problem module is the user-defined module containing the knowledge about the problem being optimised. Its main function is to implement the fitness function for a specific problem. The optimisation module contains implementation of various general optimisation methods independent of the problem type. The user interface module enables setting and reading parameters of the other two Nyx modules.

The network topology or the system structure consists of a list of network elements and a list of connections between the network elements. The network elements include *attributes* - describing the features of the real-world network elements, *ports* - specifying connections and communication between the elements and *behaviour* - giving the model description. The modules are the bases for network topology description. Each module contains a number of attributes that can be included and accessed from the GUI. The modules can be nested to arbitrary depth. Logical connections between modules are created between ports. The connections together with the list of elements describe the topology and serve as input data for network and optimisation algorithms. Layers serve as logical aggregations for modules. In telecommunication networks design the layers are mostly used to represent the same topology corresponding to different OSI layers. The network and optimisation algorithms can be applied to the modules of one layer due to different topology of the same network (a mesh in TCP layer and rings in SDH layer).

The system is the highest part of the topological hierarchy. After defining topology and aggregating the modules into layers, the user adds the modules and layers to the system. Modules, layers and systems are the three main parts of topology description.

The simulation requires a simulation model representing the module behaviour. A system initialisation method is called before the simulation run of each module. The communication process can be accomplished in three ways: (1) *Direct method calls* where one module directly calls a method of some other module; (2) *Events* that are the characteristics of discrete event simulation; (3) *Message Exchange* which is the most used communication mechanism in telecommunication network simulation.

The simulation mechanisms are the core of simulation. They are implemented by a task-specific simulation domain, run a user provided behaviour description and give some meaning to the code. The discrete event domain is preferred in telecommunication networks simulation while electronic devices

are described best by differential equations. The discrete event simulation mechanism is implemented by an event handler and an event heap.

By way of examples, possible application areas of the COSMOS tool are indicated. These include availability and reliability simulation, modelling, optimisation and analysis of a WDM all-optical network based on configurable cross-connects and circuit switching, protection and restoration schemes in the automatic circuit switched optical network, Gigabit Ethernet network modelling, simulation and analysis, broadband IP-router network modelling and analysis, traffic modelling, logical topology optimisation and burst switching.

The Computer Aided Network Planning Cockpit (CANPC) is another major network tool developed in the framework of COST 266 Action. It is suitable for designing multi-service networks that have to provide guaranteed service quality and must be optimised according to various cost constraints.

The CANPC framework provides the main user interface and acts as a nexus for the network editor, the applications and the extension manager. Within the CANPC network editor, nodes, links and traffic demands can be placed graphically. The network may be structured at different levels, each node representing a sub-network. The network model contains all input parameters that are taken into account by applications and algorithms to perform evaluations. Editing the network model involves placing and interconnecting the network objects and specifying the object properties.

Applications and algorithms: To perform a network evaluation, the appropriate application is selected from the menu *applications*. Some basic applications have already been implemented into the framework, however, it enables developers to build and integrate new user applications. Thus, the application library can permanently grow by adding new applications in different fields.

The main applications available today in CANPC include network editing applications (Network generator, Scripting application, Link editor, Traffic editor), routing applications (Shortest path routing, Hierarchical routing, Routing with load balancing) and dimensioning of WDM photonic networks (wavelength assignment, dimensioning of optical packet switched networks, dimensioning optical transport network with optical add-drop multiplexers, planning a logical topology for IP over WDM networks).

The integration of new applications, protocols and algorithms into the CANPC framework is controlled by the extensions manager. It allows the user to extend the set of active applications without any programming. The main procedures required to do this are registration of a new application, new protocols and network objects.

The other network tools developed in the framework of COST 266 Action include a simulation program called “DEMOS” made especially for simulating OPS-switches with a strictly non-blocking switching matrix and a limited number of buffer inputs. The availability studies from the Ghent group were based on the WDMGuru tool of OPNET. This is also the new name of the WDMNetDesign tool.

### **1.3 COST 266 Reference Scenario**

#### ***1.3.1. Introduction***

For the comparison and evaluation of the architectures and algorithms developed within the Action COST 266, and in order to have a common platform, the need for a reference scenario for a pan-European transport network has soon been recognized. Since the definition of a similar network within the predecessor COST Action 239 *Ultra High Capacity Optical Transport Networks*, tremendous progress in the technology of photonic devices and systems has taken place. In addition, the requirements for transport networks have been shifted (due to the spectacular success of the Internet) from telephone and SDH dominated networks towards a data centric ones based on IP (Internet Protocol). Therefore, the intention was to update the reference network and to generate a more flexible model that is capable of being adapted to the new situation.

The work has been done in cooperation with the IST project LION. Here, only the key features of the reference scenario will be summarized. Details may be found in [2] from which this section has been extracted. All data is available in an electronic form under [3].

#### ***1.3.2. Network Scenarios***

##### *Reference network topology*

After some extensive study of (currently deployed) pan-European network topologies [4], [5], a reference topology, suited for a pan-European fibre-optic network, was designed. It consists of 28 nodes in major European cities connected by 41 links in a mesh topology (see Figure 1.1.). The nodes were chosen in such a way that they include some of the European Internet Exchange Points [6].

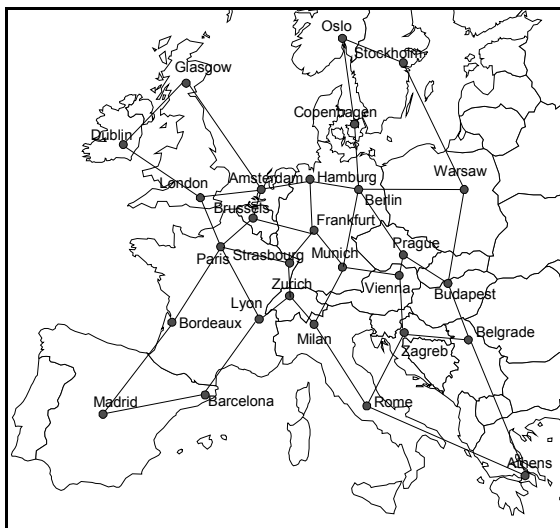


Figure 1.1. - Basic reference Topology (BT) for a pan-European fibre-optic network

### *Modified network topologies*

Starting from this Basic reference Topology (BT) two variations of dimensions can be addressed:

- The average node degree
- The number of nodes in the network.

In that way, the effect of these two parameters on the network characteristics may be studied. Four topologies have been derived from this BT. The first one is called *Core Topology* (CT) (see Figure 1.2.), as it only comprises the core part of the BT. The *Large Topology* (LT), illustrated in Figure 1.3., on the other hand, is built by extending the BT to more European countries. The CT and the LT contain thus, respectively, less and more nodes than the BT.

The other two derived topologies contain the same nodes as the BT. The difference is in the average node degree of the network, and thus the degree of meshedness. The topology called *Ring Topology*<sup>1</sup> (RT) illustrated in Figure 1.4., is a quite sparse one, whereas the *Triangular Topology* (TT) of Figure 1.5., is highly meshed. In fact, the last topology is built in such a way that it consists of triangles.

---

<sup>1</sup> Although this topology is called Ring Topology, this network is a meshed network. We only called this topology RT, because it very much resembles a ring-based network.

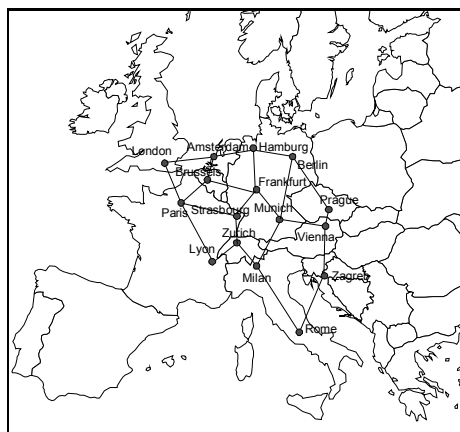


Figure 1.2. - Core Topology (CT)

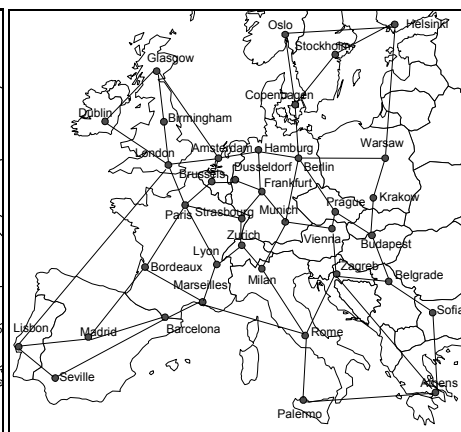


Figure 1.3. - Large Topology (LT)

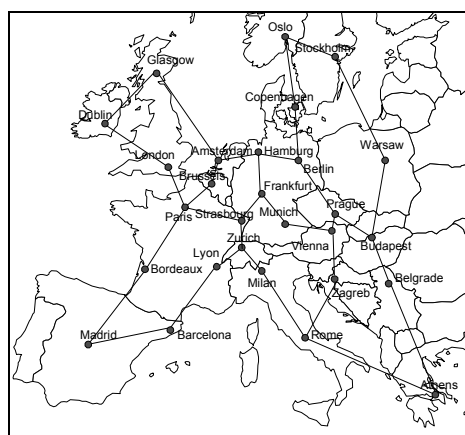


Figure 1.4. - Ring Topology (RT)

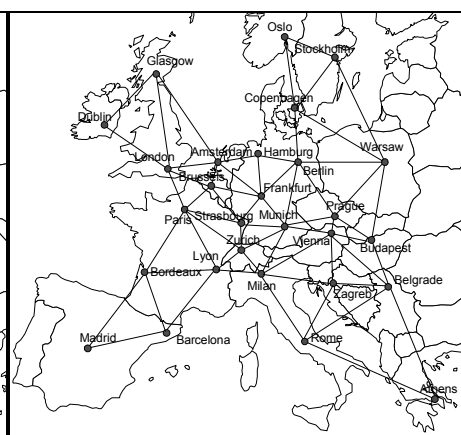


Figure 1.5. - Triangular Topology (TT)

### 1.3.3. Traffic Model

Several routes can be taken to define a suitable traffic matrix for such networks. The primary focus of the model in [7] has been the telephone traffic. According to [8], the total traffic between city  $i$  and city  $j$  can be determined as a function of a constant  $K$ , the population  $P$ , the number of non-production business employees  $E$ , the number of Internet hosts  $H$  of the city region and the distance  $D$  between the cities. The traffic is partitioned in three classes with the following characteristics:

$$\text{Voice traffic} \quad K_v \cdot P_i \cdot P_j / D_{ij} \quad (1-1)$$

$$\text{Transaction Data traffic} \quad K_t \cdot E_i \cdot E_j / D_{ij}^{1/2} \quad (1-2)$$

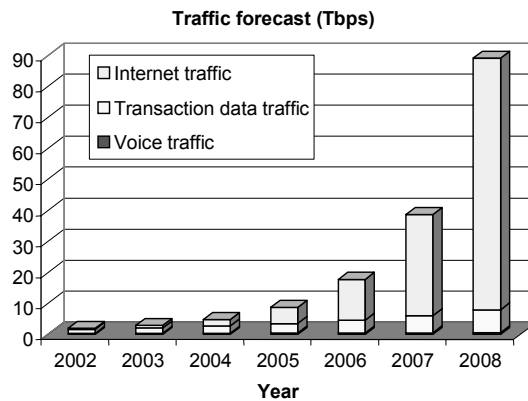
$$\text{IP traffic} \quad K_i \cdot H_i \cdot H_j \quad (1-3)$$

The constants  $K_v$ ,  $K_t$ , and  $K_i$  have to be determined based on the total volume of voice, transaction data and IP traffic, respectively, and the parameters  $P_i$ ,  $P_j$ ,  $E_i$ ,  $E_j$ ,  $H_i$ ,  $H_j$  and  $D_{ij}$ . Taking into account the consumed bandwidth per user, the average time of usage per day and the total number of users, estimates for the amount of voice, transaction data and Internet traffic in 2002 (reference year for our traffic studies) are obtained (see Table 1.1). A security factor of 5 for rush hours is used for voice and transaction data traffic.

**Table 1.1. - Figures for the different traffic classes**

	Voice traffic	Transaction data traffic	Internet traffic
<b>Traffic volume for 2002</b>	332 Gbit/s	1275 Gbit/s	341 Gbit/s
<b>Traffic annual growth rate</b>	10%	34% [8]	150% <sup>2</sup>
<b>User annual growth rate</b>	0.20% [9]	4% [10]	40% [11]

The traffic estimations for the reference year 2002, the values for the traffic constants over time and the estimation of the traffic growth allows us to calculate the current and future traffic matrices for voice, transaction data and Internet traffic. Adding these three traffic matrices yields the total traffic estimation.



**Figure 1.6. - Traffic forecast and distribution**

<sup>2</sup> Annual growth factor of 4 in the US according to [12], [13] only assumes a growth of 80% per year.



Note that the total volume of traffic is growing very fast (see Figure 1.6): from 1.9 Tbit/s in 2002 to 17.7 Tbit/s in 2006 to 89.1 Tbit/s in 2008. These numbers match very closely a forecast that appeared recently in the Financial Times. They estimated that the European traffic volume would grow from 1 Tbit/s in 2001 till 8.9 Tbit/s in 2005 (8.8 Tbit/s in our forecast).

In the first years, the transaction data traffic makes up the largest part of the total traffic volume. As time passes, the Internet traffic volume becomes more and more dominant, and from 2005 on, Internet traffic will be the most important traffic source.

As explained before, in the model used, the voice traffic between two cities is assumed to be proportional to  $1/D$ , the transaction data traffic to  $1/D^{1/2}$ , where  $D$  is the distance between those two cities, while the Internet traffic is regarded independent of the distance between the considered cities. This means that the traffic pattern will greatly change over time: in the first years most of the traffic will have to be conveyed between cities that are geographically close. With time, this distance factor will have less and less influence as Internet traffic gains importance and becomes the dominant traffic type.

#### 1.4 References

- [1] COST 266 Memorandum of Understanding.  
<http://www.ure.cas.cz/dpt240/cost266/index.html>
- [2] S. De Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. De-meester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz, "Pan-European Optical Transport Networks: An Availability-based Comparison," *Photonic Network Communication*, Vol. 5, No. 3, pp. 203-225, (2003).
- [3] [http://ibcn.atlantis.rug.ac.be/projects/COST266\\_IST\\_lion/NRS/index.html](http://ibcn.atlantis.rug.ac.be/projects/COST266_IST_lion/NRS/index.html)
- [4] <http://www.carrier1.net>, <http://www.ebone.com>,  
<http://www.globalcrossing.com>, <http://www.i-21future.com>,  
<http://www.kpnqwest.com>, <http://www.level3.com>,  
<http://www.viatel.com>, <http://www.colt-telecom.com>
- [5] M. J. O'Mahony, M. C. Sinclair, and B. Mikac, "Ultra-high Capacity Optical Transmission Networks - European Research Project COST 239", *Conf. on Telecommunications ConTEL'93*,

- Zagreb, Croatia, June 1993, *ITA*, vol. 12, no. 1-3, pp. 33-45, (1993).
- [6] [http://www.ep.net/naps\\_eu.html](http://www.ep.net/naps_eu.html)
- [7] M. C. Sinclair, "Improved model for European international telephone traffic". *IEE Electronics Letters*, vol. 30, no. 18, pp. 1468-1470, (1994).
- [8] A. Dwivedi and R. Wagner, "Traffic model for USA long-distance optical network", *Proc. of Optical Fiber Communication Conference, OFC 2000*, Baltimore, USA, March 2000, Vol 1, TuK1-1, pp. 156-158.
- [9] <http://www.odci.gov/cia/publications/factbook/indexgeo.html>
- [10] [http://www.census.gov/epcd/cbp/images/cbp\\_hist.gif](http://www.census.gov/epcd/cbp/images/cbp_hist.gif)
- [11] <http://www.infoworld.com/cgi-bin/displayStory.pl?features/990510mids.htm>
- [12] L. Roberts and C. Crump: "US Internet IP traffic growth", Caspian networks, <http://www.caspiannetworks.com/pressroom/press/08.15.01.shtml> Aug. 2001.
- [13] R. Ranganathan, L. Blair, and J. Berthold, "Benefits of grooming capable cross-connects in a Pan-European Optical Network," *Proc. of 27th European Conference on Optical Networking, ECOC'01*, Amsterdam, The Netherlands, Oct. 2001, pp. 38-39.

## Chapter 2

# Transmission Systems

*Francesco Matera (editor), Stefan Bauer, Vincenzo Eramo, Miroslav Karásek, Hans-Peter Nolting, Martin Nord, Bernd Sartorius, Marina Settembre, Ioannis Tomkos, Anna Tzanakaki, and Ioannis Zacharopoulos*

### 2.1 Introduction

Future telecommunication networks operating on wide areas will have to support a total traffic of the order of some Tbit/s. Such a huge capacity will be transmitted by means of the Wavelength Division Multiplexing (WDM) technique, or better with Dense WDM (DWDM) [1][2][3][4] and processed at the nodes by means of devices such as Optical Cross Connects (OXC) [1] and Optical Add Drop Multiplexers (OADM) [1]. Due to the traffic increasing and to the fact that transmission at 40 Gbit/s seems to be close to field testing, future telecommunication transport networks, especially if operating on wide geographical areas, can be based on the  $N \times 40$  Gbit/s transmission systems [1], even though the network design in this environment requires many ingenious contrivances. First of all the transmission at 40 Gbit/s requires sophisticated techniques in transmission (pulse generation, signal multiplexing), as well as in detection (high photodiode response, clock recovery, signal demultiplexing) and several signal control methods along the line (chromatic dispersion compensation, gain equalization of the optical amplifiers) especially if G.652 optical fibres are adopted that are the most installed ones in the world. Another fundamental aspect for wide geographical transport networks will be the intro-

duction of the optical 3R regeneration that will permit the implementation of networks operating at high bit rate (40 Gbit/s) with very long links (thousands of km).

Furthermore, another critical issue for future networks will be the lightpath assignment [5] since such a procedure does not only depend on the traffic requirements among the nodes but it critically depends on the transmission impairments [6][7][8]. Recently, a certain sensibility towards this issue has been shown also in IETF standardization body.

Conversely, for shorter distances other solutions, mainly based on high spectral density DWDM systems with lower bit rate, can be considered.

In this chapter we report the results on the ultra-high capacity transmission network physical limitations achieved in the framework of COST 266 action. In section 2.2 we report on investigation of the maximum capacity that can be achieved in links operating with  $N \times 40$  Gbit/s signals and we use the results to investigate the design of wide transport networks. The investigation of the  $N \times 40$  Gbit/s systems is limited to the G.652 links since they are the most installed ones in many countries of Europe and America. We only consider RZ pulses since previous works have shown their better performance with respect to NRZ and other signal formats [8]. Also for the chromatic dispersion compensation technique, often called *dispersion management*, we consider the two techniques that have shown the best performance for RZ pulses in G.652 links and they are periodically compensated at the output of each fibre span with and without a prechirp at the link input. For optical amplification we mainly assume the Erbium Doped Fiber Amplifiers (EDFA), but we also report results on investigation on the hybrid Raman/EDFA amplification. The results of this investigation are reported in terms of maximum number of 40 Gbit/s channels versus distance and these numbers will be used to assign the wavelengths, or lightpaths, in a transport network according to the traffic requirement, as reported in section 2.6. For such an aim we introduce a novel algorithm to assign the lightpaths taking into account the physical impairments of the fibre links, using the technique of removing lightpaths [5]. We also report an example of wavelength assignment by considering a European network consisting of 26 nodes that is a reference model studied within the European COST 266 Action.

In section 2.3 we report on the performance of DWDM systems operating in metro and short haul applications by exploiting polarisation-multiplexed schemes.

In section 2.4 we discuss the requirements and performance of different OADM and OXC architectures and technology options. The impact of the

impairments introduced by these types of nodes in a practical WDM network is also analyzed.

In section 2.5 we report the experimental activity performed on the Optical 3R regeneration considering different schemes and bit-rates. The network implementation with wavelength assignment is illustrated in section 2.6. In section 2.7 we report the activity on the Cross Gain Modulation in EDFAs considering the WDM packet-switching burst-mode networks, where no power is transmitted on empty slots.

The main conclusions on the implementation of future transport networks are reported in section 2.8.

## **2.2 N×40 Gbit/s Transmission Performance**

### ***2.2.1. Theoretical Background on N×40 Gbit/s Transmission Systems and Simulation Model***

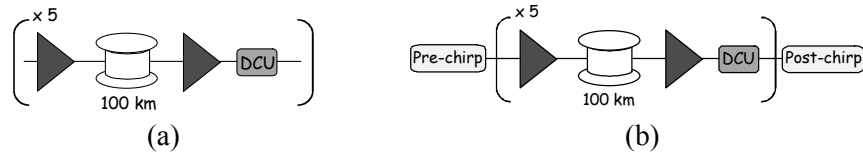
This contribution was obtained also in the framework of the IST ATLAS project. As is well known the 40 Gbit/s transmission systems in links encompassing G.652 fibre can be achieved only by compensating the chromatic dispersion by means of compensating devices such as fibre gratings or dispersion compensating fibres. Previous papers have shown that the best performance for 40 Gbit/s signals in G.652 fibres can be achieved by using Gaussian RZ pulses and several compensation schemes have been analysed to achieve the best system performance. The chromatic dispersion can be periodically compensated at each amplifier position and the compensating devices can be located either at the input or at the output of the optical amplifiers. The scheme based on the compensation at the amplifier input is generally called Post-Compensation (POC), referring to the fact that the compensation is achieved at the output of each fibre span, while the one based on the compensation at the amplifier output is generally called Pre-Compensation (PRC). From a linear analysis point of view the two schemes are equivalent, but due to the presence of the Kerr nonlinearity they show a different performance. In particular, in [4] a theoretical investigation was performed in the case of compensation obtained by using fibre gratings. Such devices have two fundamental advantages since they have very small loss and absence of nonlinearity. As a consequence, their location either before or after each fibre span has not any influence in terms of SNR. Conversely, their position can produce different nonlinear behaviour along the fibres. In fact, it was shown that for single-channel systems the Kerr effect mainly manifests itself in terms of nonlinear pulse interaction for PRC

case and spectral narrowing for POC respectively, with a worse impact in the PRC case. It was also shown that the impairments generated in the POC scheme can be limited by introducing a small chromatic dispersion pre-compensation at the link input and an opposite one at the link output [4]. Such a scheme, called POC with prechirp (POC-PRE) is the one that permits to reach the best performance in terms of maximum propagation distance for a single-channel. In [4] it was found that the optimum value of the chromatic dispersion precompensation (prechirp) is given by:

$$\frac{\exp\left(\alpha \left| \frac{T_0^2 C_{pr}}{\beta_2} \right| \right) - 1}{\alpha \left| \frac{T_0^2 C_{pr}}{\beta_2} \right|} \frac{C_{pr}^2}{\left(1 + \frac{C_{pr}^2}{4}\right)^{3/2}} = 0.77, \quad (2.1)$$

where  $\alpha$  is the loss,  $T_0$  is the time duration of the pulse and  $C_{pr}$  the value of the prechirp. Such a value is independent of the link length.

Another interesting scheme is the one based on the total chromatic dispersion compensation at the end of the link [9]. Such a scheme is very profitable from the practical point of view since the control of the GVD is not so critical in each fibre span. The system performance that can be achieved by this scheme is good, but is worse with respect to both POC and POC-PRE techniques [10]. Due to this fact we only consider in this work the POC and the POC-PRE techniques with their schemes reported in Figure 2.1.



**Figure 2.1. - Different dispersion map investigated; periodic post-compensation, straight (a) and with pre-chirp (b)**

It has to be pointed out that if we consider the use of DCF, the results can be much different since a suitable power distribution is necessary along the link in order to limit the nonlinear Kerr degradation in the DCF fibres. Comparison of the achievable propagation distances in the presence of fibre gratings and DCF have been reported in [10]. Here we can summarise that using DCF, in order to limit the strong DCF loss, a suitable double amplification before and after the DCF is necessary and even in this case the scheme with prechirp permits to achieve the best performance showing only some dB of penalty with respect to the grating case.

The model that we used for the WDM transmission is based on transmitters generating a different sequence of 128 bits with a Gaussian shape with optimised pulse duration. On the receiver side we assumed an ideal detection and clock recovery operation with electronic filtering obtained with a Bessel-Thomson filter according to the G.957 recommendation. The performance is evaluated by means of the  $Q$  factor taking into account the patterning effect of three adjacent bits [7]. We remember that in the hypothesis of Gaussian decision variable the Bit Error Rate (BER) is related to  $Q$  by means of the equation  $BER = 0.5 \cdot \text{erfc}(Q/\sqrt{2})$  [1]. It is well known from the statistical communication theory that to have a BER lower than  $10^{-9}$  a  $Q$  factor higher than 6 is required. However to achieve a transmission with a reasonable tolerance we introduce a novel requirement for the  $Q$  factor:

$$Q > 7.2 \text{ (BER} = 10^{-12}) \text{ for } P \in (P_{min}, P_{MAX}) \text{ with } P_{MAX} / P_{min} > 3 \quad (2.2)$$

that means that the BER has to be lower than  $10^{-12}$  in a large power interval in which the extremes are  $P_{min}$  and  $P_{MAX}$ . This constraint for system performance has been introduced to take into account further imperfections that can be present in the system and the ageing factors.

Our simulations were obtained by using the conventional split-step model [7] [11] that permits to evaluate the signal transmission in the presence of the dispersive effects and of the Kerr nonlinearity in the fibre. Furthermore, also the Raman nonlinearity was considered and two different routines were introduced to take into account both the Raman crosstalk and the Raman amplification. The Raman effect is responsible for a power exchange among the channels with two different consequences: different power distribution among the channels and cross talk [7]. The first effect can be compensated by means of equalisers, while the latter is much more severe in terms of performance degradation. Moreover, it is well known that the Raman effect can be used to amplify the signal [12]. In particular, by using a counter-propagating pump scheme, a distributed signal amplification along the line is realised obtaining two important goals: the signal can propagate with lower power and Amplified Spontaneous Emission (ASE) noise contribution can be lowered with respect to a conventional EDFA amplification. Another impairment due to the Raman effect is the self frequency shift [6], but it can be neglected for pulses longer than 3 ps, and this is the reason why we do not consider such a contribution in this paper. The routine for Raman cross-talk solves the coupled equations describing the Raman interaction among the channels and details and tests of this routine are reported in the deliverable D211 that can be found in the ATLAS web <http://www.fub.it/atlas/>. The Raman amplification has been

simulated with a simple routine assuming an un-depleted pump. Assuming a continuous-wave Raman pump, it is possible to calculate the Raman gain distribution along the fibre and the exact Raman pump power necessary to achieve the desired gain. Under the assumption of an un-depleted pump, the Raman gain has a  $z$ -dependence given by:

$$\begin{aligned} G_R(z) &= \frac{P_s(z)}{P_{s0} \exp(-\alpha_s z)} = \\ &= \exp\left(g_R \frac{P_p(L) \exp(-\alpha_p L)}{A_{eff} \alpha_p} (\exp(\alpha_p z) - 1)\right) \end{aligned} \quad (2.3)$$

where  $g_R$  is the Raman coefficient,  $P_p(L)$  the pump power at the fibre end,  $A_{eff}$  the fibre effective area and  $\alpha_s$  and  $\alpha_p$  are the fibre loss coefficients at the signal and pump wavelength, respectively.

As shown by several works, the optimal amplification consists of a hybrid Raman/EDFA amplification scheme in which a counter-propagating pump is launched at the output of a fibre span, before the EDFA amplifier (see Figure. 2.1). We chose the gain of the Raman amplifier equal to 8 dB with a noise figure of 0.5 dB, respectively.

EDFAs were considered employing a simple routine based on the flat gain with the presence of the ASE Gaussian noise with a noise figure of 6 dB ( $n_{sp}=2$ ).

We suppose to mainly operate in the C bandwidth of the EDFA (1528.7 nm – 1562.2 nm) that permits to locate 40 channels with a frequency spacing of 100 GHz. We have also considered the possibility of using the L bandwidth of the EDFA (1571.0 nm – 1603.5 nm) in such a way as to achieve a maximum number of channels equal to 40 with a frequency spacing of 200 GHz. As shown in [3] this is possible by using dual-band EDFAs, carefully adjusted to deliver a suitable gain in the C and L bandwidths. The flat gain can be assured by using a dynamic gain equaliser and to mitigate SRS a suitable slightly higher gain is required in the C bandwidth. A higher number of channels could be used also by exploiting the S bandwidth by employing the Thulium doped amplifiers [2], but this would require several technological considerations that are outside the scope of this report.

The gain of the EDFA (or of the hybrid Raman/EDFA amplifier) compensates for the fibre and the Dispersion Compensator Unit (DCU) losses (see Figure 2.1). Furthermore, an optical equaliser at the EDFA output was located to compensate for the channel tilt due to the Raman effect [13]. An optical filter



with a bandwidth equal to the WDM signal bandwidth was located at the amplifier output to limit the ASE amplification due to the Raman effect for frequency components that are located out of the WDM bandwidth [13].

The DCU were simulated by fibre grating devices that compensate both for the first- and third-order chromatic dispersion at each fibre span introducing a loss of 1.5 dB.

The introduction of optical filters as demultiplexers may be a source of impairments and therefore a numerical investigation was also performed on the best filter bandwidth to reach the minimum frequency spacing among the channels. We used third-order Butterworth filters since they have a shape that is quite similar to the currently commercially available demultiplexing filters based on multilayer dielectric thin-films. Laboratory tests performed on this component revealed good features in terms of in-band flatness, extinction ratio and low penalties when cascading multiplexing and demultiplexing filters. We have analysed the intrinsic penalty due to optical filtering in a back-to-back scheme considering WDM channels spaced by 100 and 200 GHz according to the ITU grid. The simulations were made by assuming RZ Gaussian pulses with time Full Width at Half Maximum ( $T_{FWHM}$ ) between 3 and 10 ps and we have found that only for pulses with time duration longer than 8 ps we could use minimum frequency spacing between the channels of 100 GHz, conversely, for shorter pulses the minimum frequency spacing was 200 GHz.

In our investigation we assume the parameters reported in Table 2.1. A long amplifier spacing of 100 km was considered.

**Table 2.1. – Link parameters**

Fibres	In-line Gratings	EDFA Amplifiers
$\alpha=0.25$ dB/km	$\alpha T=1.5$ dB	$G=26.5$ dB
$D=16$ ps/nm/km	$DT=-1600$ ps/nm	$F=6$ dB
$D_3=0.06$ ps/nm <sup>2</sup> /km	$DT_3=-6$ ps/nm <sup>2</sup>	
$\gamma=1.3$ (Wkm) <sup>-1</sup>	$\gamma=0$	
SRG= $4.48 \cdot 10^{-18}$ m/WGHz		

### 2.2.2. Single Channel Transmission at 40 Gbit/s

In Figure 2.2 we report the  $Q$  factor versus input average power for different link lengths for two different 40 Gbit/s single-channel systems: (2.2.a) transmission with POC-PRE compensation technique using pulses with  $T_{FWHM} = 5$  ps, and with an optimum prechirp equal to  $100$  ps<sup>2</sup> [4] and (2.2.b) transmission with POC scheme using pulses with  $T_{FWHM} = 8.5$  ps. We used third order Butterworth optical filters in front of the receiver that permitted to have both the minimum frequency spacing among the channels (according to

the ITU grid) and the narrowest bandwidth to limit the ASE contribution without adding a relevant penalty due to the filter distortion. For case a), according to the ITU grid, the wide bandwidth of the signal did not permit to have a frequency spacing among the channels narrower than 200 GHz and the optimum bandwidth of the filter was found to be at least equal to 120 GHz, conversely in case (b) the frequency spacing of 100 GHz can be used and we measured an optimum filter bandwidth of 80 GHz.

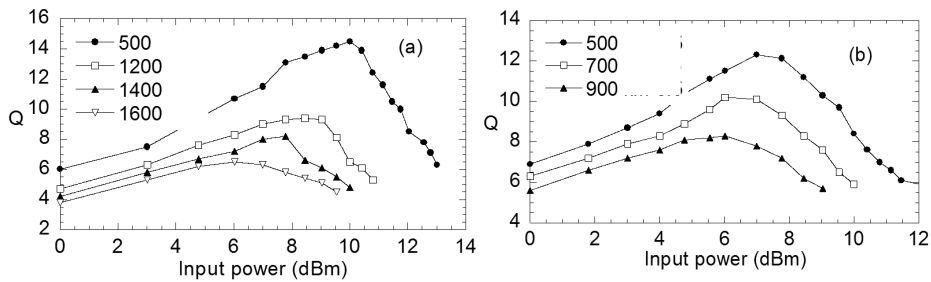


Figure 2.2. -  $Q$  factor vs. input average power for POC-PRE (a) and POC (b) technique.

It should be underlined that in case (b) for small input power the  $Q$  factor shows slightly higher values with respect to case (a) because of the narrower bandwidth of the filter. Conversely, as shown in [6], the Kerr nonlinearity in case (a) induces a reduced penalty with respect to case (b) as shown by the higher power threshold and overall by the longer distances that can be reached. Figure 2.2 shows that a transmission distance up to 1600 km is theoretically feasible with a  $Q > 6$  ( $\text{BER} = 10^{-9}$ ) in case (a), but, as reported in section 2.2.1., we require both a lower BER and a better tolerance that are formulated by eq. (2.2). According to this requirement, the maximum propagation distance for single-channel transmission reduces to 1300 km in case (a) and 900 km in case (b). The advantage of case (a) with respect to (b) is mainly due to the prechirp effect.

In Figure 2.3 we report the maximum propagation distance that can be achieved in 40 Gbit/s both with POC-PRE and POC technique assuming different  $T_{FWHM}$ . Both the optical filter in front of the receiver and prechirp for the POC-PRE technique were optimised for each  $T_{FWHM}$ . As shown by this figure  $T_{FWHM} = 5$  ps appears to be the best-performing time width and it is important to notice that for pulses longer than 8 ps the use of the prechirp does not produce any relevant improvement. Therefore we conclude that for pulses with time width of 8.5 ps the use of the sophisticated technique of the prechirp is useless.

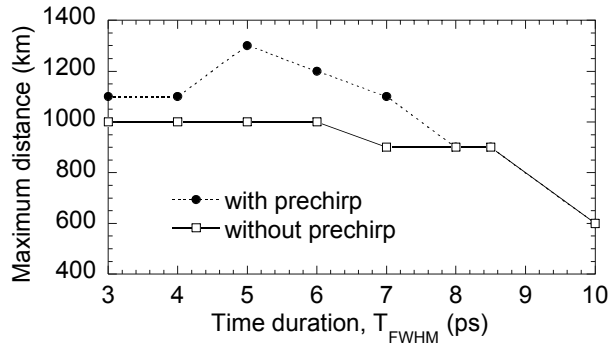


Figure 2.3. - Maximum achievable distance vs. pulse time width, with or without prechirp.

2.2.3. N×40 Gbit/s Transmission

In Figure 2.4 we report the  $Q$  factor versus channel position in the transmission of 25 channels at 40 Gbit/s in a link 500 km long by using the POC-pre techniques by assuming the frequency spacing between the channels equal to 200 GHz with time duration of the pulses equal to 5 ps.

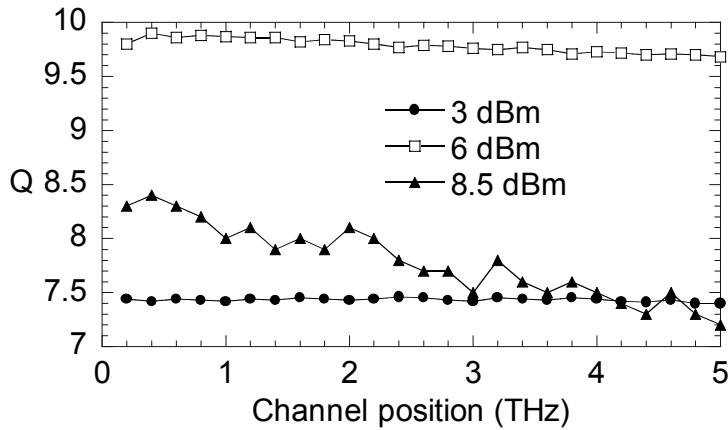
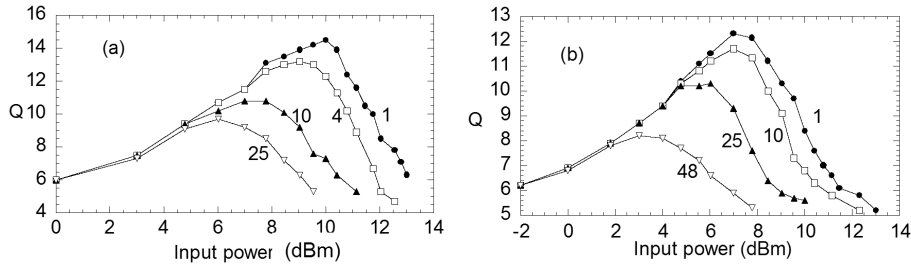


Figure 2.4. -  $Q$  factor for each of the 25 channels in the system with POC-PRE. Channel spacing = 200 GHz and  $T_{FWHM} = 5$  ps.

Three different input powers are considered. The bandwidth of the filter in front of the receiver is equal to 120 GHz. An ideal compensation of the first- and third-order chromatic dispersion at each amplifier span is assumed. As shown by the figure, for low input powers the  $Q$  factor assumes the same value for each channel since we suppose an ideal behaviour both of the amplifiers

and of the compensating devices (in each fibre span the average chromatic dispersion is zero for all the channels). Conversely, by increasing the input power, the Raman effect becomes relevant and the channels on the left transfer power to the channels on the right. This way the channels on the left manifest a loss that is compensated by introducing a routine that introduces an extra gain for the optical amplifiers. Due to this power, imperfect equalization the channels on the left only show a negligible reduction of the  $Q$  factor due to the introduction of the ASE contribution caused by the extra gain. Conversely, for the channels on the right the Raman transfer of energy produces a crosstalk effect that cannot be compensated. This is the reason for the degradation of the  $Q$  factor in the curve relative to the 8.5 dBm. Starting from this point, in the  $N \times 40$  Gbit/s system we will refer to the performance of the channels exhibiting the worst performance.

In Figure 2.5 we report the behaviour of the  $Q$  factor versus input power for different  $N \times 40$  Gbit/s systems operating over a distance of 500 km.



**Figure 2.5. -  $Q$  factor vs. average power per channel changing the number of transmitted channels Channel spacing = 200 GHz and  $T_{FWHM} = 5$  ps (a); Channel spacing = 100 GHz,  $T_{FWHM} = 8.5$  ps w/out pre-chirp (b);**

Figure 2.5a reports the results for the transmission with  $T_{FWHM} = 5$  ps,  $\Delta f = 200$  GHz and by using POC-PRE compensation technique, while Figure 2.5b reports the results for the transmission with  $T_{FWHM} = 8.5$  ps and  $\Delta f = 100$  GHz, and by using the POC compensation technique. The adopted filter bandwidth in case (a) is equal to 120 GHz and 80 GHz in case (b). In both cases the main limitation is given by the Raman crosstalk, as shown by the fact that if we neglect this effect the degradation due to the Kerr nonlinearity is negligible. In case (a) we observe that for  $25 \times 40$  Gbit/s the power margin is close to the limit expressed by eq. (2.2) for a propagation distance of just 500 km. According to this constraint we found that in case (a) the maximum number of channels achievable for  $L = 500$  km is 27. Conversely, by increasing the pulse width, a narrower frequency spacing can be adopted. In this condition the Raman cross-talk decreases due to reduced

optical bandwidth employed [22]. As a consequence, for  $L = 500$  km using 8.5 ps pulse, a much higher number of channels can be propagated. In particular, the difference between the performances when transmitting a single channel compared to transmission of 10 channels is small and, according to the constraint given by eq. (2.2), the maximum number of channels in case (b) is equal to 48.

Therefore, we conclude that, even though the POC-PRE technique with short pulses permits to reach longer distances in the single channel configuration, it cannot guarantee optimal working conditions for more channels. Conversely, the use of larger pulses adopting the simpler technique of the POC is more suitable for multi-channel operation. It has to be pointed out that the advantage of case (b) in terms of channel number is also due to the fact that the channel power can be lower with respect to case (a).

The relation of the achievable channel number versus the frequency spacing and channel power can be simply theoretically explained as reported in [14]. There, it was shown that the following equality must be satisfied in order to ensure a SNR degradation of less than 0.5 dB in the worst channel:

$$n(n-1)P_{th}\Delta fL_e = K_s \quad (2.4)$$

where:

$$L_e = m \frac{(1 - e^{-\alpha L_{amp}})}{\alpha} \quad (2.5)$$

is the effective length,  $m$  the number of fibre span and  $P_{th}$  the power threshold.  $K_s$  is a constant that depends on the system parameters, and in [14] it was found to be equal to  $8.7 \cdot 10^{12}$  HzWkm. It has to be pointed out that such an equation is valid for CW signals only and hence its application in the case of pulse signals and in the presence of strong chromatic dispersion is not valid. However, eq. (2.4) can provide some useful information about the way to vary the system parameters so as to increase the maximum number of channels. For instance, in Figure 2.5b the maximum number of channels is almost double with respect to the one shown in Figure 2.5a. This can be explained by the fact that in Figure 2.5b the product  $\Delta f \cdot P_{th}$  can be about 5 dB lower with respect to Figure 2.5a, as shown by the Figure 2.2.

#### **2.2.4. System Performance in the Presence of Raman Amplification**

The introduction of wavelength division multiplexing (WDM) has highly enhanced the transmission capacity of fibre optic communication systems, but

has also posed severe technical challenges to optical engineers. One of the main issues to be tackled is the development and deployment of large bandwidth amplifiers, because standard Erbium doped fibre amplifiers (EDFAs) cannot provide amplification over the entire band offered by optical fibres. Moreover, another critical issue is the choice of the optimal average input power, because, as explained in the previous sections, it should be a compromise between a good optical signal to noise ratio (OSNR) and negligible nonlinear effects.

The stimulated Raman scattering (SRS) has been envisaged for many years as a possible candidate for large bandwidth optical amplification, but only recently the semiconductor lasers technology has reached the maturity for providing suitably high power pump lasers [3]. The advantages of Distributed Raman Amplification (DRA) with respect to EDFAs are: 1) amplification can be provided at any frequency by simply changing the pump wavelength; 2) no special fibre is required and thus any deployed cable can be upgraded to DRA; 3) equivalent noise figure of DRA is lower than that of Erbium doped fibre amplifiers.

To introduce the Raman amplification we have first investigated the performance of a  $N \times 40$  Gbit/s system utilizing hybrid Raman/EDFA periodical amplification, by considering the two dispersion management techniques previously analysed. In Figure 2.6 we report the  $Q$  factor versus input average power for different distances, for single channel systems corresponding to cases (a) and (b) of Figure 2.5 but adopting Raman/EDFA amplifiers with a counter-propagating Raman amplification having a gain of 8 dB and a noise figure of 0.5 dB.

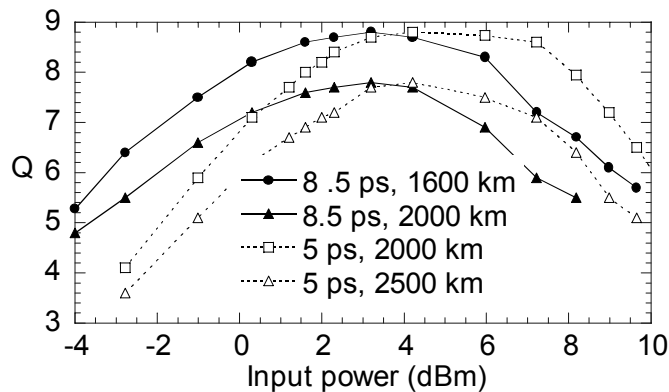


Figure 2.6. - Single channel transmission over G.652:  $Q$  factor vs. input power varying the overall link distance and the FWHM pulse width (5 and 8.5 ps).

We can see that in both cases the improvement in terms of maximum propagation distance is very relevant. It is due to two important facts: the reduction of the ASE noise level and of the nonlinear impairments (thanks to the lower input power).

Also in WDM systems the Raman amplification permits to deeply improve the system performance since the use of a lower power level along the line permits to limit the Raman crosstalk. In Figure 2.7 we report the maximum number of channels versus distance for the same cases (a), (b) of Figure 2.5 considering both EDFA and of Raman/EDFA amplification. The advantage of the Raman amplification and the fact that the narrower frequency spacing in case (b) permits a higher number of channels is evident. In these simulations for each number of transmitters considered, the average input power per channel was optimised according to the results previously discussed.

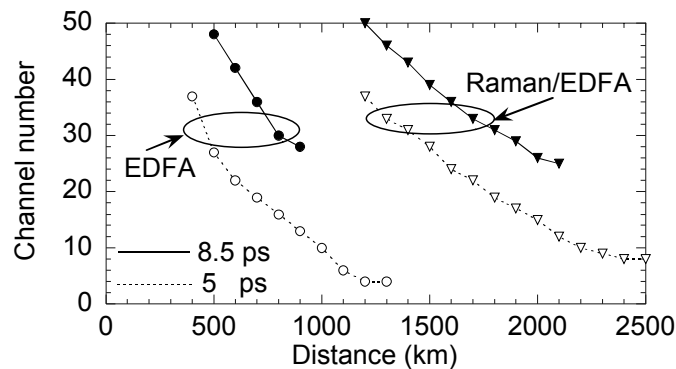


Figure 2.7. - Maximum number of transmitted channels vs. distance (over G.652 fibre) considering pure EDFA or hybrid Raman/EDFA amplification scheme.

### 2.3 DWDM in Metropolitan Area Networks

This section is based on work described in [18] and [19], which also contain references to further work on the topic.

The capacity of most fibre transmission systems is limited by the available amplifier bandwidth. The spectral efficiency (S.E.) in WDM systems is the ratio of the channel bitrate and the channel spacing within this bandwidth. The S.E. hence defines how well the bandwidth is exploited. WDM systems with both very high channel bitrate and very small channel spacing cannot be achieved, because the spectral width of the channel will be too large compared

to the channel spacing. This puts an upper limit on S.E., but advanced modulation formats and polarisation multiplexing (two channels transmitted at the same wavelength with orthogonal polarisation) techniques aim to push this limit towards an S.E. of around 1 bit/(s·Hz). Typical field systems today may have an S.E. of 0.1, e.g. using 10 Gbit/s channel bitrate and 100 GHz channel spacing. Quadrupling the bitrate and doubling the channel spacing doubles the S.E., but experiments with channel bit-rate at 40 Gbit/s and frequency spacing of 100 GHz have also been demonstrated.

High bitrate systems require complex transmitter and receiver modules. Furthermore, to obtain acceptable transmission performance, such systems often need dynamically adjustable dispersion, dispersion slope and polarization mode dispersion (PMD) compensation. DWDM systems have much-relaxed requirements to these factors, but have higher requirements to the stability of the carrier frequencies and the demultiplexing module. Higher channel count gives a higher number of transmitter and receiver modules. However, the lower bitrate opens up for alternative approaches. E.g. successful experiments with a super-continuum source, modulation and demultiplexing for 12.5 GHz channel spacing systems have been demonstrated. In the context of optical networks, other factors also influence the choice between DWDM and WDM. Higher channel count is beneficial in optical wavelength-routed systems, since a higher number of paths can co-exist on the same fibre. As discussed in chapter 4.3, optical packet or burst switched (OPS/OBS) networks also benefit from high channel count to resolve contention.

In systems with channel spacing equal to or below 50 GHz, four-wave mixing (FWM) becomes an important limitation on transmission distance and maximum transmission power, even when high dispersion fibre, like the standard SMF (SSMF) is used. We have addressed transmission issues and explored the benefit of using a polarisation-multiplexed scheme (PMS) to reduce FWM. The proposed PMS is very simple, and consists of launching every other channel with a state of polarisation (SOP) orthogonal to its adjacent channels. This corresponds to doubling the frequency spacing of the channels interacting to generate the FWM products. To prevent interaction between adjacent channels, the PMS must be maintained during transmission. However, PMD de-correlates the SOPs of different wavelengths, quantified by the differential group delay (DGD). Our measurements reveal that no significant de-correlation occurs over 150 km SSMF when channel spacing is 50 GHz or less.

We consider the PMS for application in metropolitan and medium-haul DWDM systems with 12.5 GHz channel spacing and 2.5 Gbit/s bitrate. The performance of the PMS was evaluated by considering a four-channel system



over 150 km SSMF. Demultiplexing becomes critical in such DWDM systems. We obtained 20 dB suppression of adjacent channels, mainly using two lock-in fibre Fabry-Perot filters with FWHM bandwidths of 8 and 15 GHz, in combination with broader Bragg grating filters.

Both experiments and simulations have been conducted. First, a 5-channel system with a void centre channel was studied in terms of FWM product build-up. The FWM product generated in the void channel by the four other channels was measured for different channel powers using the PMS, and compared to the case when parallel SOPs was launched. At 0 dBm/ch the suppression relative to the signal channel power was higher than 30 dB in both cases. At 8 dBm/ch the suppression was only 20 dB in the parallel case, but more than 26 dB using the PMS. Hence, significant FWM is only generated at sufficient channel powers, and the PMS reduces the amount of FWM in this case.

To estimate the transmission performance we measured BERs, both in the parallel and in the PMS case, at channel powers from 0 to 8 dBm/ch. The sensitivities at BER of  $10^{-9}$  were expressed as “normalized penalties”, with respect to the sensitivity at 0 dBm/ch. Experimental results reveal that the PMS has 5 dB less normalized penalty than the orthogonal case at 6 dBm/ch. Using the PMS, normalized penalties were below 1.1 dB up to 8 dBm/ch. Numerical simulations of a similar system have confirmed this trend.

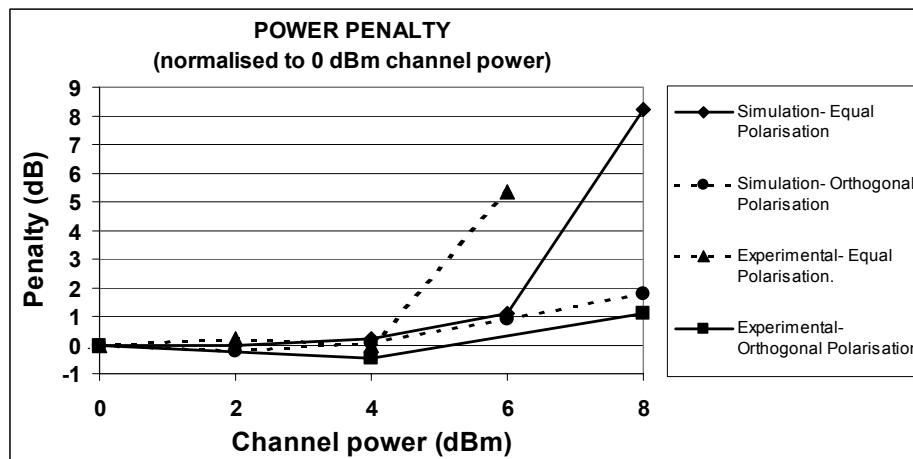


Figure 2.8. - The penalty induced by FWM depends on the relative polarisation state of the 12.5 GHz spaced DWDM channels.

In conclusion, the proposed PMS significantly reduces FWM, and thereby enables successful transmission of DWDM systems with relatively high channel launch power at 12.5 GHz channel spacing. The increase in channel

launch power can be used to increase tolerance to transmission distance, component loss or splitting of signals. DWDM can be an attractive solution in metropolitan and medium haul optical networks to increase the number of paths in an OCS network and to reduce the need for buffering in an OPS/OBS network employing wavelength conversion.

## **2.4 Optical Add/Drop Multiplexers and Optical Cross-Connects for Wavelength Routed Networks**

Currently, telecommunications networks widely employ WDM to interconnect discrete network locations and offer high capacity and long reach transmission capabilities. The information transmitted in the optical domain is transferred through simple point-to-point links terminated by SONET/SDH equipment forming ring and mesh network topologies. This solution requires a number of intermediate service layers introducing complex network architectures. Such a scenario provides unnecessarily high switching granularity, numerous opto-electronic conversions and complicated network management resulting in poor scalability for data services and slow service turn up with high installation, operation and maintenance cost.

With the recent technology evolution in the area of optical communications, the WDM transport layer is migrating from simple transmission links into elaborate networks providing similar functionality to that of the SONET/SDH layer, with improved features, higher manageability, lower complexity and cost. Integrated WDM networks performing switching and routing are deployed in order to overcome the need for multi-layer network architectures [20]. In such network scenarios, high capacity optical routes are set in the transport layer forming connections between discrete points of the network topology. The application of generalised multi-protocol label switching (GMPLS) as the standardised and common control plane enables interoperable optical networks bridging the gap between the traditional transport infrastructure and the IP layers [21]. These networks support an intelligent optical layer utilising optical add/drop and cross-connect nodes [22] and provide traffic allocation, routing and management of the optical bandwidth. They also accommodate network expansion, traffic growth, churn and network survivability.

This section focuses on OADM and OXC solutions based on different architecture and technology choices supporting a variety of features.

### ***2.4.1. Optical Add/Drop Multiplexers***

Optical add/drop multiplexers (OADMs) are elements that provide capability to add and drop traffic in the network (similar to SONET ADMs). They are located at sites supporting one or two (bi-directional) fibre pairs and enable a number of wavelength channels to be dropped and added reducing the number of unnecessary optoelectronic conversions, without affecting the traffic that is transmitted transparently through the node. The features that an OADM should ideally support are listed below:

- drop capability of any channel and insertion capability of a channel on any unused wavelength,
- dynamic reconfiguration supporting fast switching speed ( $\sim$  ms),
- variable add/drop percentage up to 100%,
- scaleable architecture in a modular fashion,
- drop and continue functionality enabling network architectures such as dual homing rings,
- span and ring protection capabilities, and
- minimum performance degradation in terms of noise, crosstalk, filtering etc., for add, drop & through paths.

An OADM can be used in both linear and ring network architectures and in practice operate in either fixed or reconfigurable mode. In fixed OADMs the add/drop and express (through) channels are predetermined and can only be manually rearranged after installation. In reconfigurable OADMs the channels that are added/dropped or pass transparently through the node can be dynamically reconfigured as required by the network. These are more complicated structures but more flexible as they provide provisioning on demand without manual intervention, therefore, they can be set up on the fly to allow adding or dropping of a percentage of the overall traffic.

Enabling technologies that are commonly used in fixed OADM solutions are based on thin film filters and fibre Bragg gratings. The technology of choice is determined by the functional and the performance requirements of the node. Reconfigurable OADMs can be divided into two categories: partly and fully reconfigurable. In partly reconfigurable architectures there is the capability to select the channels to be added/dropped, but there is a predetermined connectivity matrix between add/drop and through ports restricting the wavelength assignment function. Fully reconfigurable OADMs also provide the ability to select the channels to be added/dropped, but also offer flexible connectivity

between add/drop and through ports, which enables flexible wavelength assignment with the use of tuneable transmitters and receivers. Reconfigurable OADMs can be divided into two main generations [23]. The first is mainly applied in linear network configurations and does not support optical path protection while the second is applied in ring configurations and provide optical layer protection to support network survivability.

The two most common examples of fully reconfigurable OADMs i.e. “Wavelength Selective” (WS) and “Broadcast and Select” (B&S) architectures are illustrated in Figures 2.9 (a) and (b).

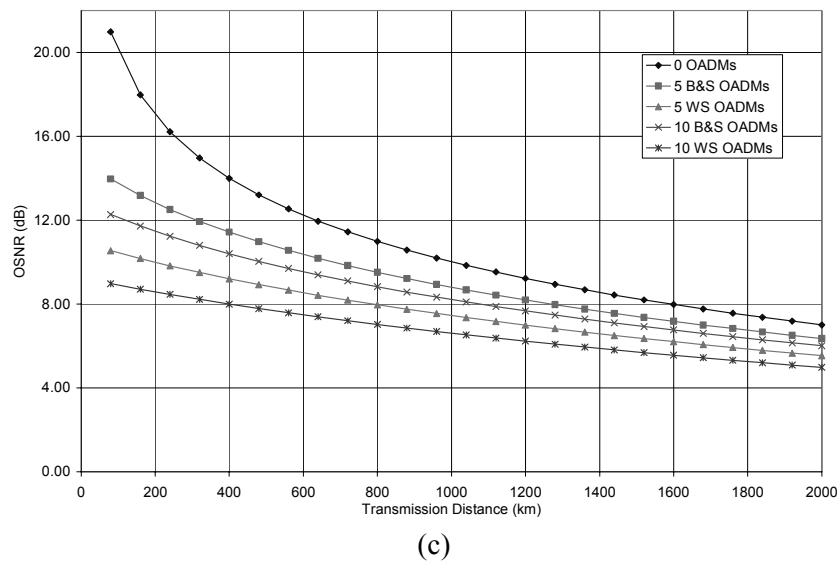
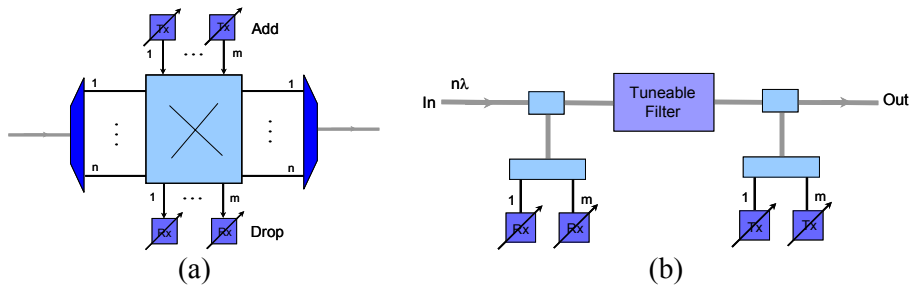


Figure 2.9. - (a) WS and (b) B&S OADMs architectures, (c) system OSNR performance

The WS architecture utilises wavelength de-multiplexing/multiplexing and a switch fabric interconnecting all express and add/drop ports, while the B&S is based on passive splitters/couplers and tuneable filters. The tuneable filtering,

present in the through path, should provide selective blocking of any dropped channels and can be achieved using technologies such as acousto-optic filters or dynamic channel equalisers (DCEs) [24] based on diffraction grating and liquid crystal or micro-electromechanical system (MEMS) technologies. Extensive investigations on different OADM architectures have been published in the literature [22]. In this report we compare the performance of the two architectural options using an analytical optical signal to noise ratio (OSNR) model. The study is based on detailed OADM designs and specifications of commercially available optical components such as amplifiers, filters, DCEs, muxes/demuxes, couplers etc. This analysis has shown that the overall loss introduced by the through path of the B&S solution is noticeably lower than the loss of the WS approach significantly improving the OSNR of the node and therefore its concatenation performance in a practical transmission link or ring. Figure 2.9 (c) illustrates the OSNR evolution across a link supporting different numbers of OADMs based on the B&S and WS architectures. The analysis is assuming 80 km amplifier spans (20 dB loss) and 6 dB noise figures for all the amplifiers used in the system. In addition to the OSNR performance benefit, the B&S design offers superior performance in terms of filter concatenation effects. The excellent concatenation performance of this architecture is confirmed by experimental results published in [25]. It also offers advanced features such as drop and continue functionality and provides good scalability in terms of add/drop percentage [26].

#### **2.4.2. Optical Cross-Connects**

Optical cross-connects (OXC) are located at nodes cross-connecting a number of fibre pairs and support add and drop of local traffic providing the interface with the service layer. To support flexible path provisioning and network resilience, OXCs normally utilise a switch fabric to enable routing of any incoming channels to the appropriate output port and access to the local client traffic. The features that an OXC should ideally support are similar to these of an OADM, but additionally OXCs need to provide:

- strictly non-blocking connectivity between input and output ports
- span and ring protection as well as mesh restoration capabilities

A number of OXC solutions based on different technologies have been proposed to date and depending on the switching technology and the architecture used they are commonly divided into two main categories: opaque and transparent [20].

Opaque OXCs are either based on electrical switching technology or on optical switch fabrics surrounded by o/e/o conversions, imposing the requirement of expensive optoelectronic interfaces (Figure 2.10). In OXCs using electrical switching, depending on the technology and architecture, sub-wavelength switching granularities can be supported providing edge and intermediate grooming capabilities for more efficient bandwidth utilisation. Opaque OXCs also offer inherent regeneration, wavelength conversion and bit-level monitoring. In transparent OXCs the incoming signals are routed transparently through an optical switch fabric without the requirement of optoelectronic conversions. The switching granularity may vary and support switching at the fibre, the wavelength band or the wavelength channel level.

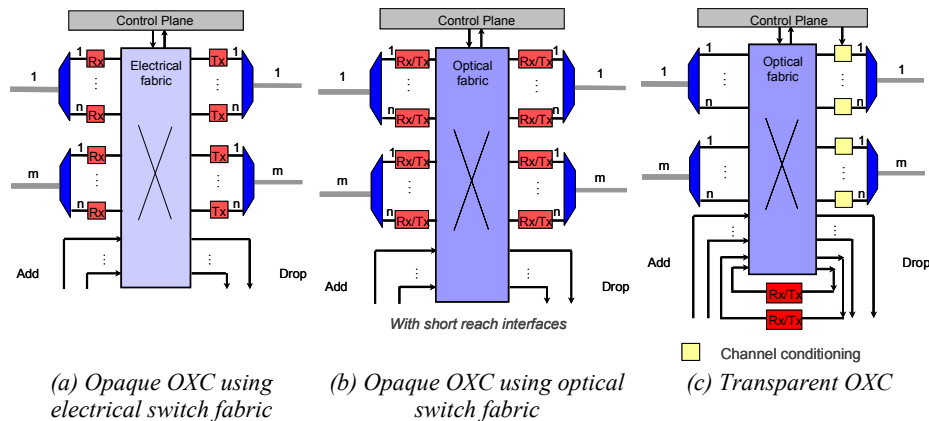
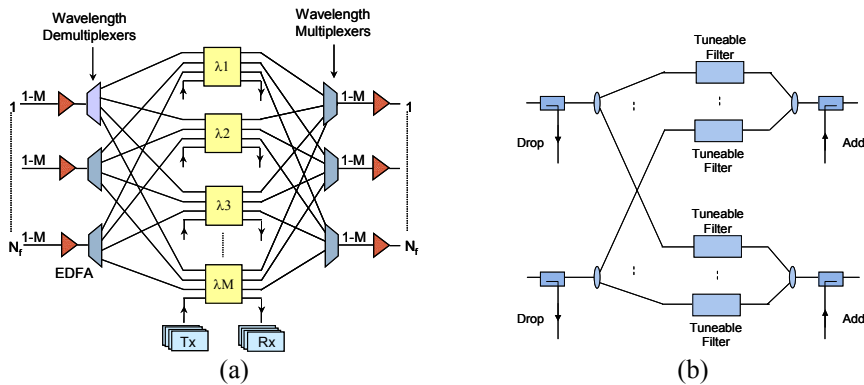


Figure 2.10. - OXC technologies

Such nodes support switching of optical signals without regeneration capabilities. This may significantly impact the scalability of the overall solution as present optical networks are analogue in nature introducing penalties on optical channels transmitted over fibre, due to effects such as amplifier noise and wavelength dependent gain spectrum, dispersion, nonlinear effects, polarisation mode dispersion (PMD) etc. It should be also noted that transparent OXCs are introducing transmission impairments themselves such as noise, crosstalk, polarisation dependent loss (PDL), PMD, dispersion, filtering effects etc. To overcome these physical limitations, partially regenerating architectures have been proposed in the literature [27]. These provide the ability to selectively regenerate individual wavelength channels of degraded signal quality when forming paths that exceed the transparency distance, i.e., the length an optical channel can traverse without the need for optoelectronic conversion. To achieve this, a set of regenerators that can be selectively accessed by any of the incoming wavelength channels is used. In addition to selective regeneration,

full and partial wavelength conversion [28][29] can be also applied to reduce wavelength blocking, offer improved bandwidth utilisation and support a control plane scheme compatible with GMPLS.

All-optical solutions offer transparency to a variety of bit-rates and modulation formats. However, it should be noted that different bit-rates and modulation formats may exhibit different transmission characteristics and impairments affecting the overall network design and implementation. A variety of optical switching technologies have been proposed and developed to date such as: 2-D and 3-D MEMS, bubble jet, semiconductor optical amplifier gates, holographic switches, liquid crystal, thermo-optic, electro-optic technologies [30]. The table illustrated in Figure 2.11c summarizes the basic characteristics of the main switch technology options.



Technology	Loss (dB)	Crosstalk (dB)	PDL (dB)	Switching time	Max Switch Dimension	Maturity
2D MEMS	5	55	0.2-0.5	<10 ms	32x32	high
3D MEMS	7	55	0.5	<10 ms	4000x4000	low
Bulk Mechanical	3	55	0.2	<10 ms	16x16	high
Bubble based Waveguide	7.5	50	0.3	<10 ms	32x32	high
Liquid Crystal	1	35	0.1	<10 ms	2x2	high
Electro-optical	8	35	0.5	<1 ns	16x16	low
Thermo-optical	8	35	0.5	<10 ms	16x16	high
Holographic	1	35	low	<10 ns	64x64	low
SOA	gain	35	low	<10 ns	4x4	low

(c)

Figure 2.11. - (a) WS and (b) B&S OXC architectures and (c) optical switching technologies

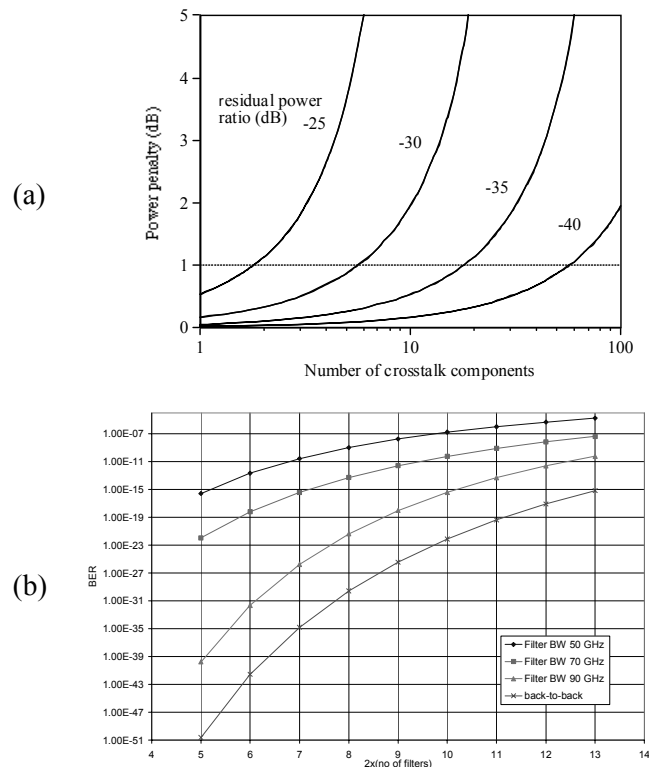
A number of different OXC architectures have been reported in the literature. The most straightforward solution is based on a central switch fabric which can potentially support a very high port count equal to  $[(\text{no. of fibres}) \times (\text{no. of } \lambda\text{s}) + (\text{no. of add/drop ports})]$ . Based on the above formula, an OXC, which supports 4 fibre pairs each carrying 80 wavelength channels and supporting 50% add/drop capability, would require a  $960 \times 960$  optical switch fabric. A scaleable optical switching technology that could support the requirement for very high port count would be based on beam stirring technology such as 3-D MEMS. However, most other (more mature) optical switching technologies are limited to significantly smaller fabrics typically  $32 \times 32$ . Therefore, various multistage optical switch structures have been proposed based on combinations of switching stages, such as the Clos and the wavelength selective switch architecture. The main idea behind the WS switch is that it exploits the wavelength dimension in order to conveniently segment the switching fabric, so, following the wavelength demultiplexing stage, the incoming wavelength channels are directed to discrete switches each supporting a single wavelength. The switching structure offers the ability to route any channel of a specific wavelength to any output port of the OXC offering strictly non-blocking connectivity. Another architecture that supports strictly non-blocking functionality is the broadcast and select [31]. In this case instead of optical switch fabrics the architecture is based on passive splitters and couplers combined with tuneable filters, which can be implemented using acousto-optic filters or DCE technologies as in the case of the OADMs mentioned above. The B&S architecture can offer low loss OXC solutions for nodes that support a limited number of input and output fibres, as in this case the loss of the node is significantly affected by the splitting ratio of passive splitters used to broadcast and recombine the transmitted WDM signals. This is a very attractive feature as in most practical optical network applications a small number of fibres (typically 4) are interconnected at the cross-connect nodes, each carrying a large number of wavelengths (typically up to 160). It should be noted that, when considering OXCs supporting capability to add and drop local traffic, the simple architectures described above need to be enhanced to provide the required add/drop connectivity.

#### ***2.4.3. System Performance Discussion***

The system concatenation performance of OADMs and OXCs is determined by a number of parameters such as the OSNR as discussed in paragraph 2, the crosstalk performance and the filtering characteristics of the node etc. A number of different studies have focused on the description of crosstalk implications in WDM systems [28][32] when leakage of optical signals transmitted



through muxing/demuxing, switching and other optical components interferes with their neighbouring channels. Figure 2.12a illustrates that severe power penalties may be introduced in the presence of interferometric crosstalk in the system [32].



**Figure 2.12. - Power penalty vs. number of crosstalk components for different levels of crosstalk (b) BER evolution with filter concatenation for different filter bandwidths**

These results are derived assuming optimised receiver threshold. Another effect that may cause significant penalties in an optical network is abrupt changes in the traffic loading conditions. In a long chain of amplifiers a change in the spectral loading introduced by adding or dropping in OADMs or switching in OXCs may cause gain tilt and severe transient effects as rapid variations of the amplifier input power level may cause unwanted sudden gain increase [33]. This problem is commonly overcome by using fast amplifier gain control, gain clamping, Raman amplification or dynamic spectral loading schemes. In this Section we will focus on the filtering characteristics of this type of nodes and their concatenation effects. As discussed above, OADMs

and OXCs require WDM multiplexing and demultiplexing functions in order to isolate and route individual wavelength channels through the nodes. If a number of these nodes are cascaded within the optical path of a WDM wavelength channel, the channel experiences the combined effect of all the cascaded filters, which mathematically is described by the multiplication of the transfer functions of each individual filter present in the path. This introduces a much narrower filter bandwidth compared to that of the individual filters depending on the shape of the filters used. This effect is further emphasised by any filter misalignments present in the system increasing further the transmission penalty due to filter concatenation [34]. The penalty introduced by filtering effects strongly depends on the spectral characteristics of the transmitted signal and therefore on the modulation format of the system. Therefore, very different penalties may be introduced in the system when using NRZ and RZ (33% and 50% duty cycle) modulation formats. Novel modulation schemes (e.g. carrier suppressed RZ) can be used to suppress filtering related penalties. Figure 2.12b demonstrates the BER performance of a point-to-point link incorporating a variable number of filters for different filter bandwidths. The results are taken for NRZ modulation format at 10 Gbit/s.

## 2.5 All-Optical 3R Signal Regeneration

In present telecommunication networks only the links are optical. Switching and processing in the nodes is performed in the electrical domain.

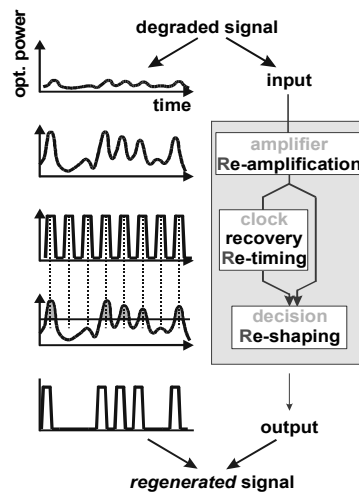


Figure 2.13. - The key functional blocks of 3R-regeneration: (1) *re-amplification*, (2) *re-timing* and (3) *re-shaping*

With the introduction of the optical cross-connect (OXC) in the near future, transmitted signals will be optically by-passed in the node. Noise and cross talk from transmission and switching nodes degrade the signal quality and limits the transmission distance. This leads to islands of transparency unless optical 3R signal regeneration is accomplished [35-50]. The key functional blocks and the architecture of 3R-regeneration (**3R** is the abbreviation for *Re*-amplification, *Re*-shaping, and *Re*-timing) is shown in Figure 2.13.

### 2.5.1. Clock Recovery and 3R Regeneration Schemes

The general scheme of a 3R regenerator is shown in Figure.2.14. One part of a degraded data signal is injected into the clock recovery. The function of the clock is to emit a stream of optical pulses, stable in amplitude and time, synchronized to the data stream. The other part of the data stream drives the decision element (e.g. a gate based on a MZ-interferometer with integrated SOAs as switching devices) that encodes the data signal onto the clock pulses. The signal quality has to match demands for transmission over the next fiber link. Two features are essential for regeneration: First, a non-linear transfer characteristic of the gate to suppress noise and amplitude fluctuations. And second, the switching window has to be larger than the clock pulse width. In this case jitter and pulse shape degradation of the data signal are not transferred to the clock pulses. The quality of the output signal thus is defined only by the quality of the clock. The clock is not only needed for retiming, it is also responsible for reshaping.

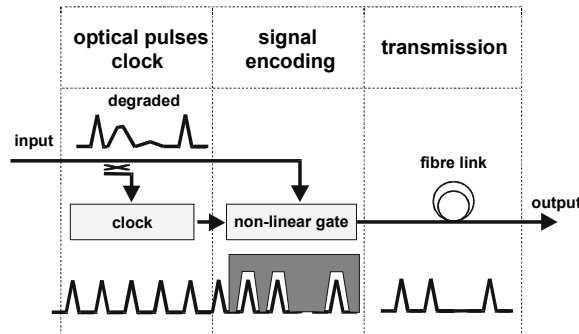


Figure 2.14. - Scheme of a 3R regenerator

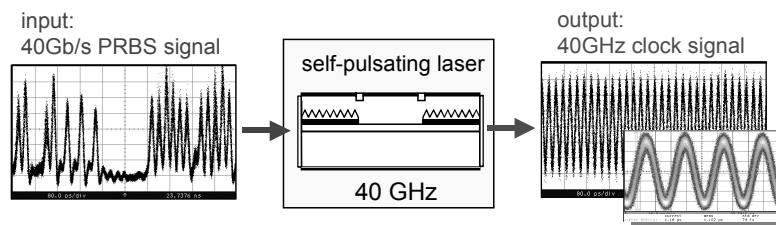
### 2.5.2. Basic Devices for 3R-Regenerator

We use two lines of basic optical signal processing devices, both are based on semiconductor laser (amplifier) material: the self-pulsating laser for clock extraction and the SOAs for switching. The advantages of this choice are (i)

short device length (compared to fibre), (ii) robust and stable operation for integrated devices, (iii) cheap to fabricate and they offer (iv) gain.

### *Clock Recovery*

*All-optical clock recovery* needs an optical oscillator, which can be synchronized to the injected optical data signals. Two options are available so far. The first is based on the well known mode-locked lasers [35] the second uses a novel device type, a self-pulsating multi-section DFB laser [36]. In both cases the optical data signal is injected into the device, the oscillation frequency is adjusted close to the data rate, and the synchronized clock pulse trace is obtained at the output port (Figure.2.15). For both device types clock recovery with a good system performance has been demonstrated at 10 Gbit/s [37] and 40 Gbit/s [38][39].



**Figure 2.15. - Self-pulsating DFB laser as all-optical clock.**

However, the different physical effects utilized in these devices lead to distinct features. The mode-locking frequency is defined by the cavity length that has to be adjusted very precisely in integrated devices. Mechanical frequency tuning is only possible using external cavity lasers. In self-pulsating lasers the frequency is depending on the laser dynamics and can be tuned electrically.

### *Fast optical gate*

Basic physical effect for *all-optical switching* is the change of the refractive index of the gain material due to the injection of optical control signals, which are contra- or co-propagating with the signals. Non-coherent effects are used. Thus for data input and control input we need always two different wavelengths.

Dependent on the utilization of the gain change (cross gain modulation: XGM) or the phase change (cross phase modulation: XPM) different architectures are possible. The phase change has to be translated by a Mach-Zehnder-Interferometer (MZI) to an amplitude change. In a SOA we have two time constants for switching: a slow switching time based on the intra-band effects (ns) and a fast switching time based on the intra-band effects (ps to 200 fs). Usually the

non-linear MZI is used as an optical gate to stop or transmit the clock pulses. Another application of the MZI is wavelength conversion. Interferometers with monolithically integrated SOAs (Figure. 2.16) as well as hybrid solutions are used and operation for DEMUX up to 160 Gbit/s has been shown.

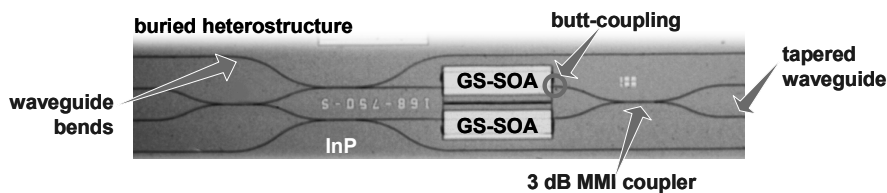


Figure 2.16. - MZI with SOAs as fast optical gate

### 2.5.3. Clock Recovery for Flexible and Asynchronous Networks

The electrical frequency tuning of a self-pulsating laser has been demonstrated to be a continuous tuning range e.g. extending from 6 to 46 GHz, covering two SDH/Sonet hierarchies (10, 40 Gbit/s). System tests at various bit rates have demonstrated the good performance of the clock recovery [42]. The bit rate flexible operation of the optical clock represents a functionality that is not available in present commercial available electronic solutions.

In a packet switched asynchronous network the bits are not synchronized from packet to packet. Therefore an ultra-fast locking clock is essential. The self-pulsating laser offers this feature.

A 10 Gbit/s packet containing “one” bits is injected into the non-synchronized self-pulsating laser. The synchronization can be observed to be within 10 bits [43]. This is an ultra-fast locking time of 1 ns. After switching off the data signal the clock holds the synchronization for more than 100 zero bits. These features have been used for error free clock recovery from asynchronous data packets [44].

### 2.5.4. Test of All-Optical 3R Regenerators in Loop Experiments

Loop experiments are the crucial test for 3R regenerators. Self-pulsating lasers have been applied and tested in 3R regenerators. Figure 2.17a shows one setup. A non-linear Mach-Zehnder interferometer gate is used as decision element. The regenerated signal is circulated in a loop containing 50 km of dispersion-shifted fibre and the BER performance [45] (sensitivity for BER  $10^{-9}$ ) is measured for several laps (Figure 2.17b). The signal quality is stable from lap 1 to lap 300. This demonstrates that the 3R regenerator suppresses the accumulation of noise and jitter. Cascading of 300 laps represents a transmission over 15 000 km.

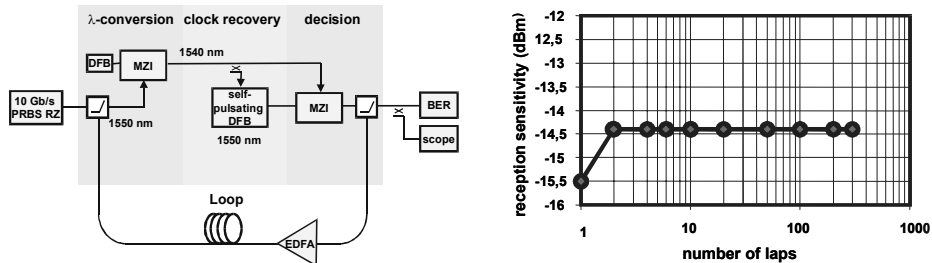


Figure 2.17. - Scheme of loop testbed (left), and BER performance in loop (right)

The 40 GHz optical clock has been tested in a modified regenerator (soliton compression and synchronous modulation). The eyes of the 40 Gbit/s and the demultiplexed 10 Gbit/s signals are clearly open after 10 000 km of transmission [46]. BER measurements verify that the transmission is error free and that the clock works at various lengths of the PRBS pattern.

### 2.5.5. Novel 3R-Regenerator Scheme

A novel all-optical ADC (Alternating Data Clock [47][48]) 3R regenerator architecture has been proposed (Figure 2.18) and demonstrated. It combines regeneration with a flexible wavelength conversion, which is an important advantage for application in WDM systems. Compact semiconductor devices are applied for both, all-optical clock recovery and the decision stage. EDFAs are not required within the ADC 3R regenerator. Thus cost effective 3R circuits based on SOA and laser devices can be integrated e.g. on a Si-motherboard.

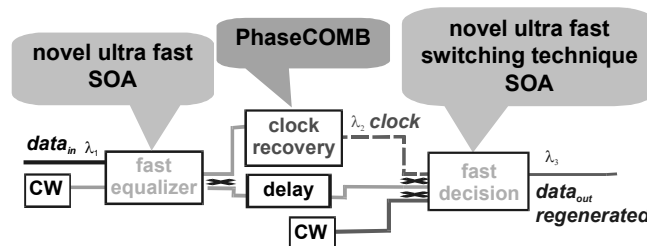


Figure 2.18. - Architecture of ADC-3R-Regenerator

Excellent re-timing and re-shaping of ADC-3R has been demonstrated already in system experiments at 40 Gbit/s. The speed potential of this architecture by modification of the devices (PhaseCOMB for clock recovery [49] and Fast Decision for clocked switching [50]) up to 160 Gbit/s has been shown by simulation studies.

### 2.5.6. Summary

Clock recovery is a key function in optical 3R regenerators. The self-pulsating lasers offer bit rate flexible operation and an ultra fast locking to asynchronous IP packets. The regenerating performance is demonstrated in loop experiments at 10 and 40 Gbit/s. A novel all-optical ADC (Alternating Data Clock) 3R regenerator architecture has been demonstrated experimentally for 10 and 40 Gbit/s. It combines regeneration with a flexible wavelength conversion and has the potential to avoid EDFAs for the switching function. Both together are important advantages for application in WDM systems.

## 2.6 Algorithm to Assign the Lightpaths and Example for an Optical Transport European Network

In this Section we show how a wide geographical transport network, based on the  $N \times 40$  Gbit/s transmission links encompassing G.652 optical fibres, can be designed according to the results found in section 2.1.

At the moment several works have reported algorithms for optimum design of logical topologies in WDM networks [5], but to our knowledge only in [51] the physical limits of a network were taken into account, even if it referred to a metropolitan network with a ring topology. We use an algorithm based on the *removing lightpath* [5] since it also permits to consider the transmission impairments in a simple way; in fact, as the first example, the lightpaths that have a length longer than the one permitted by the transmission impairments have to be removed.

We suppose that our network is composed of nodes having an OXC that permits to route the lightpaths without wavelength conversion. The nodes are connected, according to their geographical disposition, by fibre links and the connections are described by the  $[L]$  matrix, where  $l_{ij}$  is the length of the fibre link between  $i$  and  $j$  if a connection is present (0 otherwise). As input we have a static traffic matrix,  $[T]$ . In a Multi-protocol Label Switching (MPLS) in the WDM architecture [52] the element  $t_{ij}$  of the matrix  $[T]$  denotes the flow (in Gbit/s) of packets belonging to the Label Switched Paths (LSP) to be routed between the nodes  $i$  and  $j$  where edge Label Switch Routers (LSR) are attached.

We proposed an algorithm able to route, through the optical network, all of the offered LSPs with the object of minimising the used resources and with the constraint of guaranteeing to the routed lightpaths a given optical signal quality expressed in term of the  $Q$  factor as mentioned in section 3. In our

algorithm, according to [7], first we connect all the  $i$ - $j$  having  $t_{ij}$  different from 0, choosing the shortest lightpath between  $i$  and  $j$  found from the matrix [L]. Sequentially we remove the lightpaths carrying less traffic flow and we groom (at electrical level) such traffic with the traffic present in the remaining lightpaths, trying to reach the capacity of 40 Gbit/s for each lightpath. We stop the procedure when all the traffic is carried by the minimum number of lightpaths, taking into account two constraints: (i) for restoration reasons in each country we must have at least two lightpaths as input and as output, (ii) some lightpaths can be fixed for strategic consideration. During this procedure it has to be checked that the maximum number of lightpaths for a link has to satisfy the WDM transmission impairments according to the results shown in Figure 2.7. A diagram of the algorithm is reported in Figure 2.19.

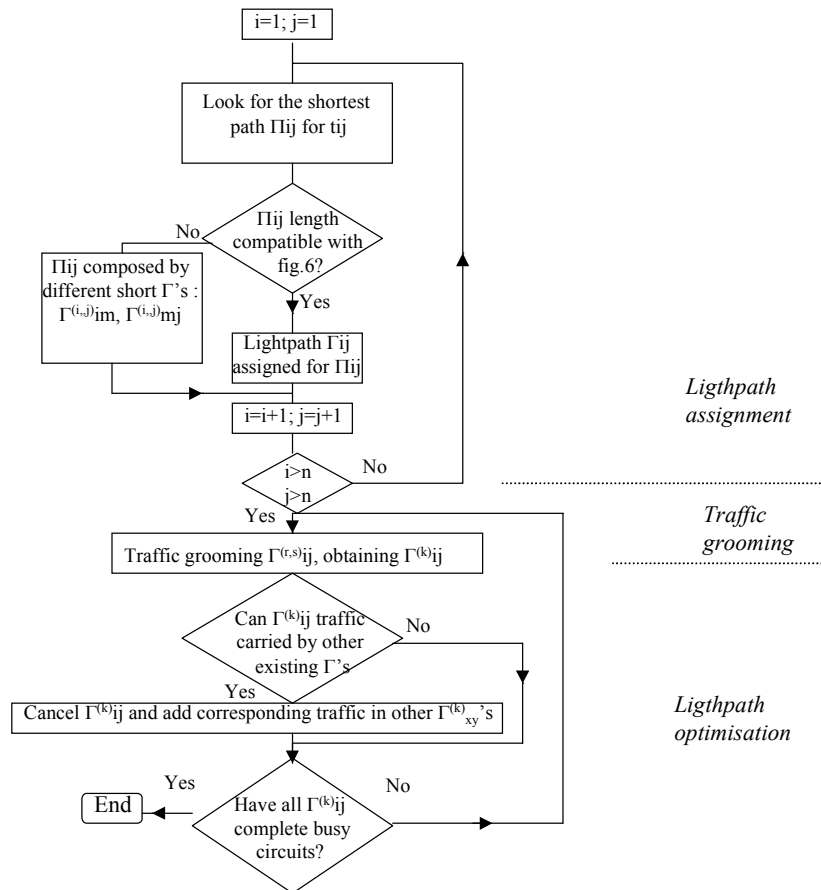


Figure 2.19. - Algorithm for wavelength assignment.



As an example, we have taken into consideration a model for a European network proposed within the European COST 266 project, assuming  $M=26$  nodes reported in Figure 2.20 with the respective optical links [53]. We have assumed a traffic matrix  $[T]=\psi[T_0]$ , where  $\psi$  is a scaling coefficient that permits to vary the total traffic,  $T_T=\sum_i^M \sum_j^M t_{ij}$ , between 100 and 3000 Gbit/s.

The assumed traffic matrix is for the case of a total traffic of 2400 Gbit/s. The elements  $t_{ij}$  have been chosen to have high average input/output traffic for big cities. Furthermore, high input traffic has been foreseen for Lisbon representing the America gate and also a relevant traffic for Rome (gate for Africa by means of the link via Palermo) and Warsaw (gate for eastern countries).

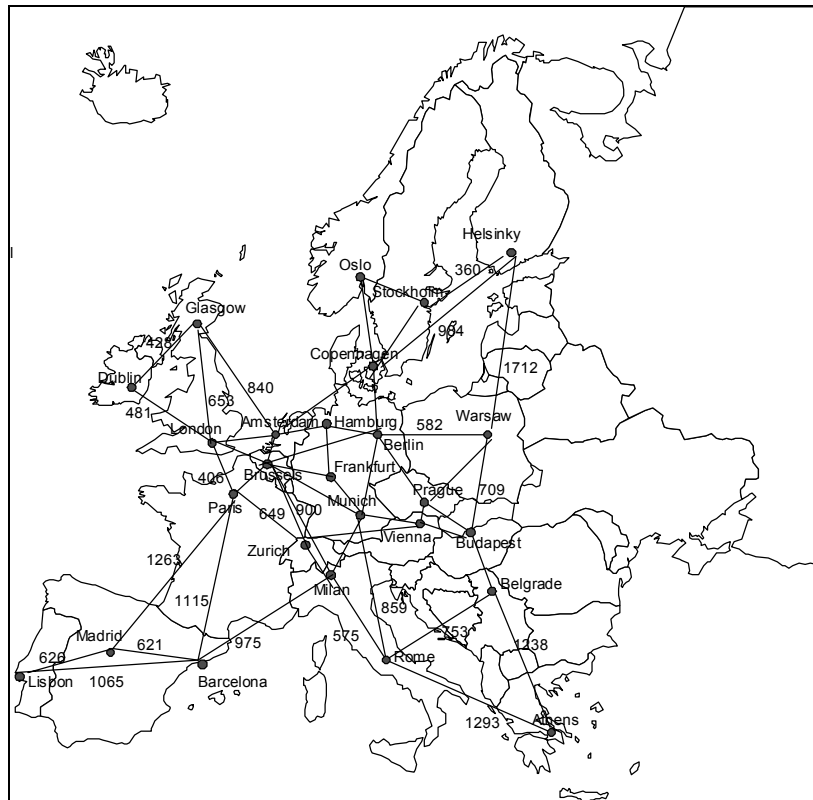
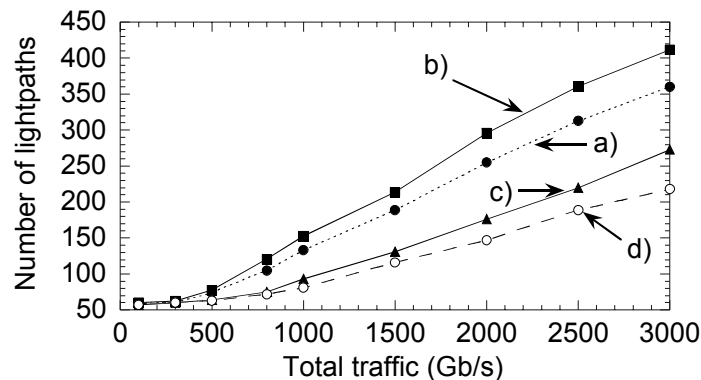


Figure 2.20. - European network

In Figure 2.21 we report the number of lightpaths versus the total traffic by using different  $N \times 40$  Gbit/s transmission techniques:

- (a) transmission with the POC-PRE technique with a  $T_{FWHM}$  pulse duration of 5 ps and a frequency spacing among the channels of 200 GHz by using EDFA (case of Figure 2.5a),
- (b) transmission with the POC technique with a  $T_{FWHM}$  pulse duration of 8.5 ps and a frequency spacing among the channels of 100 GHz by using EDFA (case of Figure 2.5b),
- (c) transmission with the POC technique with a  $T_{FWHM}$  pulse duration of 8.5 ps and a frequency spacing among the channels of 100 GHz by using hybrid Raman/EDFA amplification, and
- (d) represents the case of ideal propagation and it means that the lightpaths can have any length.

Practically, the case (d) could be achieved by means of either optical or electrical 3R regenerators. A preliminary numerical investigation on the 40 Gbit/s optical 3R regeneration based on the semiconductor optical amplifiers has been carried out. It has shown that by locating the optical 3R regenerators every 500 km a maximum propagation distance of 4500 km can be reached.



**Figure 2.21.** - Number of 40 Gbit/s lightpaths vs distance in links encompassing G.652 fibres, by using: a) transmission with the POC-PRE technique with a  $T_{FWHM} = 5$  ps,  $\Delta f = 200$  GHz and EDFA, b) POC technique with  $T_{FWHM} = 8.5$  ps,  $\Delta f = 100$  GHz and EDFA, c) POC technique with  $T_{FWHM} = 8.5$  ps,  $\Delta f = 100$  GHz and Raman/EDFA, d) ideal propagation.

The (b) technique is the simplest to be achieved since the chromatic dispersion control is less severe (absence of prechirp) but it is the one that requires more lightpaths since the signal can propagate over shorter distances. In case (c) the number of lightpaths can be much less than in (a) and (b) since the scheme Raman/EDFA allows the signal to propagate over longer distances. The

advantage of (d) with respect to (c), in terms of lightpaths, is not so relevant since, in the European network scheme, the paths longer than 2000 km that should carry heavy traffic, are rare. Furthermore, in case (d) the network complexity could increase very much due to the presence of the 3R regeneration - a technique that still requires several contrivances.

For the correct interpretation of the results it is clear that in case (c) the presence of Raman/EDFA amplifiers in all the links of the networks is not necessary but only in the links containing lightpaths that have a distance longer than 900 km. In the same way, in case (d) the presence of the 3R regeneration is required only in links containing lightpaths longer than 2000 km.

It is evident that at this moment it is difficult to make a comparison in terms of costs among the techniques reported in cases (a) to (d), especially for links employing Raman amplification and 3R regenerators. However, if we look at the distribution of the lengths of the required lightpaths, we could see that the number of lightpaths with distances longer than either 900 km or 1300 km is much less with respect to the total number of the lightpaths. It means that a limited introduction of Raman/EDFA amplifiers every 100 km in some links of the network could strongly reduce the number of the 40 Gbit/s Tx/Rx devices.

Furthermore, only a limited number of 3R regenerators need be located in links providing lightpaths longer than 2000 km, that, according to our network topology and traffic requirements, are very few.

In this work we have supposed the OXC as ideal devices. This means that the lightpaths in each node can be optically routed without any distortion and loss. This assumption would seem very strong, especially because of the presence of WDM multiplexing/demultiplexing devices that surely introduce losses. According to the experimental activities carried out in the framework of the IST ATLAS project, the OXC can be assumed as a device that introduces only a limited loss that can be compensated by means of an optical amplifier with limited gain and, as a consequence, with negligible ASE contribution. For instance, it has been experimentally demonstrated that no relevant penalty in terms of performance is introduced in the optical demultiplexing/multiplexing process of a channel in the middle of a link 500 km long with further wavelength conversion of the channel obtained by means of a Periodical Poled Lithium Niobate device [17]. This result allows us to regard our approximation on the OXC quite good. It is clear that any further degradation introduced by the OXC would result in reduction of the maximum propagation distance and an increase in the number of the lightpaths.

## 2.7 Cross Gain Saturation in Erbium Amplifiers

In WDM circuit-switching networks, the number of channels present in an EDFA may vary due to network reconfiguration or channel failure. This leads to cross-gain saturation in fibre amplifiers that in turn induce power transients in the surviving channels that can cause severe service impairment due to either inadequate eye opening or the appearance of optical nonlinearities. Even more serious bit error rate deterioration can arise in WDM packet-switching burst-mode networks, where no power is transmitted during empty slots. In [54] we have presented results of experimental and theoretical investigation of the effect of cross gain saturation on transmission of packetized Ethernet data at 10 Mbit/s over three wavelength channels in a cascade of five EDFAs. Our experimental setup is depicted in Figure 2.22.

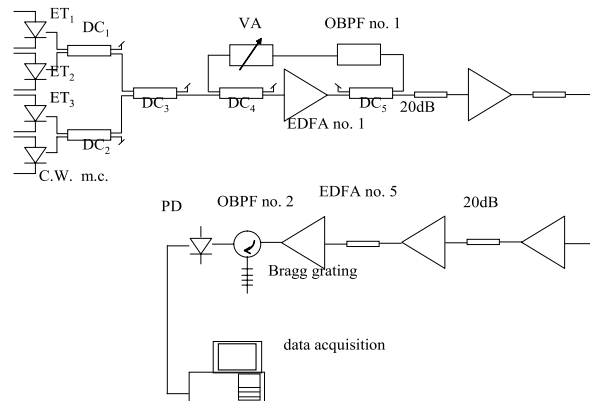


Figure 2.22. - Schematic diagram of the experimental setup.

Electrical signals at 10 Mbit/s from three Ethernet Hubs carrying the traffic of a local area network (LAN) were converted to optical signals in three Ethernet transmitters (ET). At least three personal computers were connected to each Hub. The traffic was generated by copying long data files between personal computers and servers, playing video files stored in another computer and transferring files with *ftp*. Optical signals from three Fujitsu DFB lasers tuned to 1549.1, 1551.1, and 1552.9 nm were combined in 3 dB directional couplers (DC) with a continuous wave monitoring channel power at 1556.3 nm and fed to the input port of EDFA no. 1. The continuous-wave signal was used to

monitor output power fluctuations due to the bursty nature of the LAN traffic and the cross gain saturation of EDFAs. Gain of the EDFA no. 1 may be clamped by an optical feedback loop. The lasing power generated in the gain clamped amplifier propagates through the cascade and absorbs input power fluctuations caused by the bursty nature of the LAN traffic. By adjusting the pump power, the gain of individual fibre amplifiers was set to exactly compensate the 20 dB inter-amplifier loss for input powers of  $P_{sig} = -14.8$  dBm/channel (for signal channels) and  $P_{cw} = -17$  dBm for the monitoring channel at 1556.3 nm. Due to the cross gain saturation of EDFAs, random and bursty low frequency modulation of the three signal channels at the packet level is reflected in the variation of EDFA gain and consequently in power fluctuations of the originally CW channel. An optical band pass filter (OBPF 2) implemented as a circulator and Bragg grating was used to select the 1556.3 nm wavelength power from the output spectrum of an EDFA. A PIN FET photo-detector and a data acquisition card were used to process the time variable monitoring channel signal. In order to acquire the fast power transients after amplifier no. 5 caused by the CW channel power over- and undershoots, the sampling rate of 1 MHz has been selected. Two second long samples of the CW channel power have been recorded and stored in a PC. The evolution of the cross gain saturation along the cascade was investigated by placing the OBPF no. 2 and a photo-detector at the output of the individual EDFAs.

Figures 2.23 and 2.24 show zoomed parts of the optical power recorded at the output of EDFA no. 3 and EDFA no. 5, respectively. It is seen that after EDFA no. 3, five output power levels occur. Due to inequality of gain at individual signal channel wavelengths, the "1 channel ON" output power level is split into two levels. This inequality of gain accumulates along the cascade and after EDFA no. 5 even the "2 channels ON" and "3 channels ON" levels are split in two levels so that the total number of discrete power levels is eight, see Figure 2.24. Also, characteristic power overshoots and undershoots occurred at the output of EDFA no. 3 and their amplitude increased after EDFA no. 5. The cause of these over- and undershoots is the same as in the surviving channel power fluctuations in the WDM add/drop experiments. The monitoring channel power excursions grow between EDFA no. 1 and no. 3, the difference between the "3 channels OFF" and "3 channels ON" levels increases from 3.96 to 5.09 dB, respectively. Due to the decrease in gain of EDFA no. 4 and 5 caused by accumulation of amplified spontaneous emission (ASE) power around 1530 nm along the EDFA cascade, monitoring channel power excursions after amplifier no. 5 falls to 3.80 dB, see Figure 2.25.

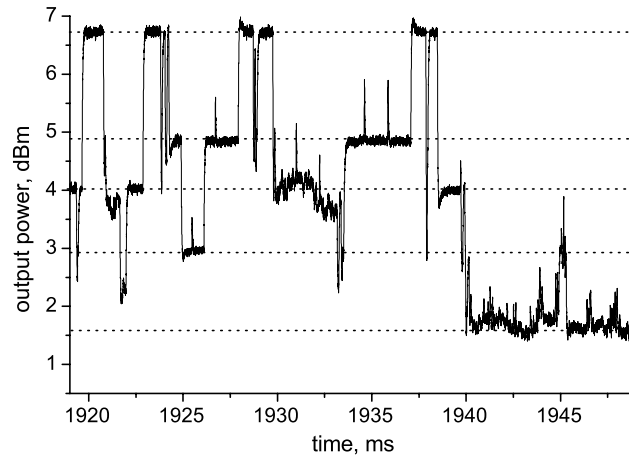


Figure 2.23. - Time evolution of monitoring channel power at the output of EDFA # 3 - experiment.

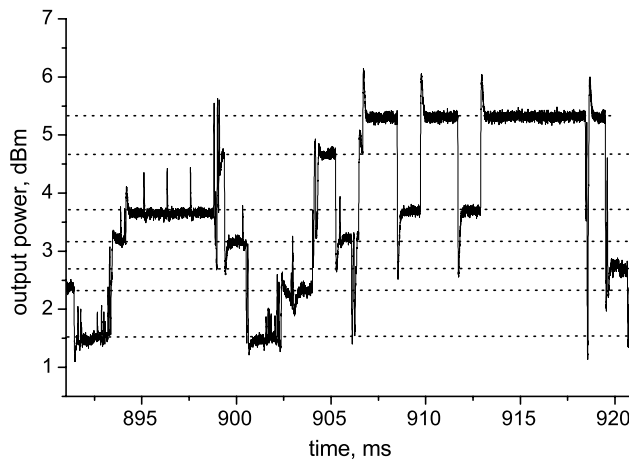
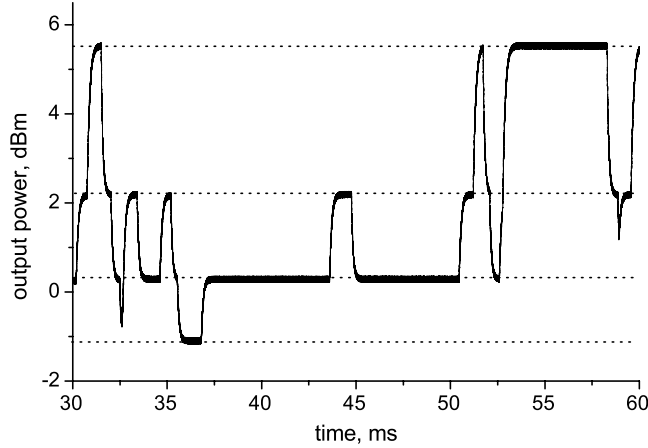


Figure 2.24. - Time evolution of monitoring channel power at the output of EDFA # 5 - experiment



**Figure 2.25. - Time evolution of monitoring channel power at the output of EDFA # 1 - simulation**

The model used for the simulations is based on the assumption of homogeneously broadened gain medium and absence of excited-state absorption [55]. The simulation of burst-mode WDM packet traffic was described in detail in [56]. The amplified spontaneous emission is taken into account in our model. The wavelength region from 1450 nm to 1650 nm is resolved into 200 bins of constant width  $\delta\lambda = 1\text{nm}$ . Dynamic behaviour of each doped fibre in the cascade is described by a single ordinary differential equation for time evolution of the length averaged metastable level population,

$$N_2(t) = \frac{1}{\rho L} \int_0^L n_2(z,t) dz, \text{ where } n_2(z,t) \text{ is the population density of the } {}^4I_{13/2}$$

metastable level averaged over the cross section of erbium-doped fibre (EDF),  $L$  is the length of the EDF, and  $\rho$  is the  $\text{Er}^{3+}$  ion density. The ON/OFF sources simulating the time-slotted burst-mode WDM traffic were randomly generated in a way similar to that described in [57]. The ON intervals represent the presence of packets; the OFF intervals represent idle periods. Using a random number  $U$  uniformly distributed on  $[0,1]$ , statistically independent ON and OFF intervals (in units of slots) were generated for each WDM channel with a

$$\text{truncated Pareto distribution via } T_{ON} = T_{slot} \left\lfloor \frac{1}{U^{1/\alpha_{ON}}} \right\rfloor, \quad T_{OFF} = T_{slot} \left\lfloor \frac{1}{U^{1/\alpha_{OFF}}} \right\rfloor$$

where  $\alpha_{OFF}$ ,  $\alpha_{ON}$  are parameters regulating the burstiness of the traffic, and  $\lfloor x \rfloor$

is the floor function. Time slot length of  $T_{slot} = 57.6 \mu\text{s}$ , corresponding to 65 bytes of 10 Mbit/s IP was selected to simulate the traffic used in our experiment. The monitoring channel power was sampled at 100 points per  $T_{slot}$ . The resultant time step for the evolution of the length-averaged metastable level population was  $\delta t = 0.576 \mu\text{s}$ . Simulation of the highly variable input traffic was performed over ten million time steps, corresponding to 5.76 ms of transmission. Figures 2.26, 2.27 show a short time segment of 30 ms of monitoring channel power evolution at the output of directional coupler DC<sub>5</sub> and at the output of EDFA no. 5, respectively.

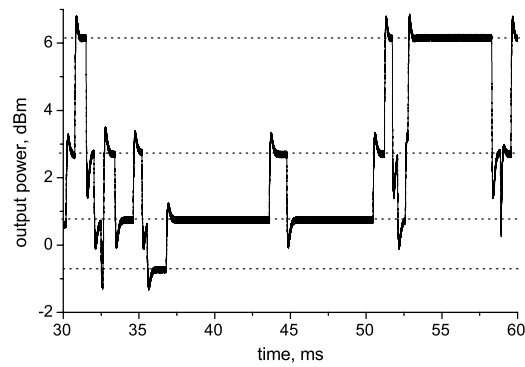


Figure 2.26. - Time evolution of monitoring channel power at the output of EDFA # 5 – simulation.

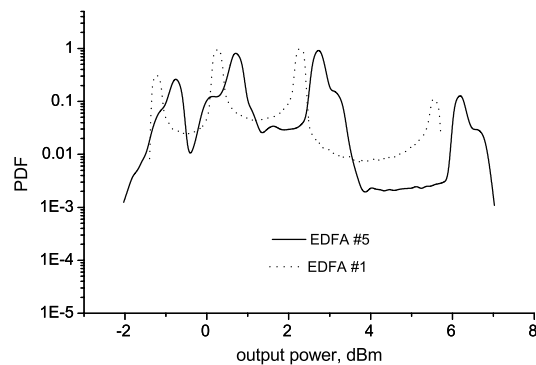


Figure 2.27. - Probability density function of  $P_{cw}$  after EDFA #1 and #3 – simulation



Power over- and undershoots develop after EDFA no. 3 at instances of packet appearance or disappearance in one of the signal channels. These over- and undershoots are responsible for peak broadening of the monitoring channel power probability density function (PDF) shown in Figure 2.27.

Our experimental results demonstrate that substantial swings in EDFA output power can be induced by cross gain saturation when packetized bursty traffic is transmitted through a cascade of fibre amplifiers. Amplitude of these swings grows along the cascade in the same way as surviving channel power fluctuations when channels are dropped/added in a WDM circuit-switching scenario. Although our experimental setup allowed us to transmit the packetized traffic with bit rate of 10 Mbit/s over three channels only, our numerical simulations predict that even in an 8 channel WDM system with 155 Mbit/s bit rate bursty traffic the probability of long empty slots is rather high and can give rise to large fluctuations of EDFA gain.

## 2.8 Conclusions

The investigation presented in this chapter had four aims: (1) to illustrate what are the maximum capacities that can be achieved by using different transmission techniques; (2) to show how the choice of the transmission systems depends on the traffic characteristics of the networks; (3) to show the role of future devices as optical 3R regenerators and (4) to explore the effect of bursty traffic on the quality of the transmitted optical signals.

The results have shown that the technique of the periodical compensation of the chromatic dispersion with prechirp is the one that permits to achieve the longest distances when short pulses are used. Conversely, if we are interested in shorter distances but with a larger number of channels, the use of long pulses is preferable, also because we can avoid the use of the prechirp. By using the hybrid Raman/EDFA amplifier, we can greatly improve the system performance with respect to the case when only EDFA amplification is used.

The choice of transmission technique strongly depends on the transport network characteristic both in terms of node distance and traffic requirement. By considering an environment given by an European networks in which the main node distances are between 1000 and 2000 km, the use of the dispersion management with Raman/EDFA permits to have a more efficient lightpath distribution without resorting to the 3R optical regeneration, limiting the number of the channels, with an important saving of 40 Gbit/s Tx-Rx equipment and optical amplifiers that can require a narrower bandwidth.

This work was focused on the G.652 fibres since they are the ones most installed in the world. Other available fibres are the G.653 (Dispersion Shifted, DS) fibres but these cannot be used in the C bandwidth since the four-wave mixing is very degrading due to the fact that the chromatic dispersion has the zero value just in this window. Therefore, the WDM systems could be installed operating in S and L bandwidths only. Furthermore, at moment other novel fibres have been specifically designed just to support WDM systems. They exhibit a small, but nonzero, chromatic dispersion. As shown by several papers, by using these fibres, known as nonzero DS fibres (G.655 recommendation), much higher capacities can be reached compared to systems using the G.652 fibres, however, at the moment, it is questionable whether the capacity requirements can justify the installation of new fibre cables.

## 2.9 References

- [1] G. P. Agrawal, "*Fiber-Optic Communication Systems*," Wiley-Interscience, John Wiley & Sons, INC. 1997.
- [2] J. P. Blondel, "Massive WDM Systems: Recent Developments and Future Prospects" *Proc. ECOC 2001*, Amsterdam 30 Sept.-4 Oct. 2001.
- [3] S. Bigo, W. Idler, J-C Antona, G. Charlet, C. Simonneau, M. Gorlier, M. Molina, S. Borne, C. de Barros, P. Sillard, P. Tran, R. Dischler, W. Poehlmann, P. Nouchi, and Y. Frignac, "Transmission of 125 WDM channels at 42.7 Gbit/s (5 Tbit/s capacity) over 12×100 km of TeraLight™ Ultra fibre," *Proc. of ECOC 2001*, Amsterdam 30 Sept. - 4 Oct. 2001.
- [4] M. Zitelli, F. Matera, and M. Settembre, "Single-channel transmission in dispersion managed links," *Journal of Lightwave Technology*, vol. 17, pp. 2498-2505 (1999).
- [5] E. Leonardi, M. Mellia, M. Ajmone Marshan, "Algorithms for logical topology design in WDM networks" *Optical Networks*, 1, 35-46 (2000)
- [6] G. P. Agrawal, "*Nonlinear Fiber Optics*," Academic, San Diego, Cal., (1995)
- [7] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, "*Nonlinear Optical Communication Network*," John Wiley & Sons, Inc. (1998).

- [8] D. Dahan and G. Eisenstein, "Numerical Comparison Between Distributed and Discrete Amplification in a Point-to-Point 40 Gbit/s 40 WDM-Based Transmission System With three Different Modulation Formats" in press on *Journal of Lightwave Technology* 2002.
- [9] A. Mecozzi, C. Balslev Clausen, and M. Shtaif, "System impact of intra-channel nonlinear effects in highly dispersed optical pulse transmission", *IEEE Photonic Technology Letters*, vol. 12, pp. 1633-1635, 2000.
- [10] A. Pizzinat, A. Schiffini, F. Alberti, A. Paoletti, D. Caccioli, P. Griggio, P. Minzioni, and F. Matera, "Experimental and numerical comparison of dispersion compensation techniques on different fibers" in press to *IEEE Photonic Technology Letters* of October 2002.
- [11] F. Matera and M. Settembre, "Comparison of the performance of optically amplified transmission systems," *Journal of Lightwave Technology*, vol. 14, no. 1, pp. 1-12, 1996.
- [12] S. Bigo *et al.*, "1.28 Tb/s WDM transmission of 32 ETDM channels at 40 Gb/s over 3×100 km distance," *Proc. of ECOC 2000*, Munich, Sept. 2000, vol. 4, pp. 17-19.
- [13] C. M. McIntosh, A. G. Grandpierre, D. N. Christodulides, J. Toulouse, and J-M P. Delaux, "Eliminating SRS Channel Depletion in Massive WDM Systems via Optical Filtering Techniques," *IEEE Photonics Technology Letters*, vol. 13, no. 4, pp. 302-304, 2001.
- [14] A. R. Chraplyvy and R. W. Tkach, "What is the Actual Capacity of Single-Mode Fibers in Amplified Lightwave Systems?" *IEEE Photonics Technology Letters*, vol. 5, no. 6, pp. 666-668, 1993.
- [15] A. Pizzinat, M. Santagiustina, and C. Schivo, "Impact of hybrid EDFA-distributed Raman amplification on a 4×40 Gbit/s WDM optical communication systems" submitted to *Photonics Technology Letters*.
- [16] [www.fub.it/atlas/](http://www.fub.it/atlas/): Milestone M311.
- [17] A. Schiffini *et al.* "In-Field demonstration of N×40 Gbit/s transmission with in-line Ti:PPLN all-optical wavelength in a link 500 km long encompassing both G.652 and G.655 fibres," Submitted at *OFC 2003*.

- [18] S. Bjørnstad, M. Nord, G. Hanssen, B. Slagsvold, and D. R. Hjelm, "Impact of FWM in polarisation multiplexed 12.5 GHz channel spacing WDM systems," *IEEE Lasers Electro-Optics Society (LEOS)*, (1), pp. 285-286, 2001.
- [19] M. Nord, "Reducing Four Wave Mixing Impairments in DWDM Optical Metro Networks," *Cost 266 Technical Presentation*, Zagreb, 2001.
- [20] A. Tzanakaki, I. Wright, and S. S. Sian, "Wavelength Routed Networks: Benefits and Design Limitations," *SCI2002*, Orlando, Florida, July 2002
- [21] A. Banerjee, J. Drake, P. Lang, and B. Turner, "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements," *IEEE Communications Magazine*, pp. 144-150, January 2001.
- [22] J. Lacey, "Optical cross-connect and add/drop multiplexers: technologies and applications", (Tutorial), WT1, *OFC2002*.
- [23] Chikama *et al*, "Photonic Networking Using Optical Add Drop Multiplexers and Optical Cross-Connects," *Fujitsu Sci. Tech., J.*, 35, 1, pp.46-55, July 1999.
- [24] I. Tomkos, M. Vasiliev, J. Rhee, M. Sharma, and M. Soulliere, "Transparent Ultra-Long-Haul Optical Networks Enabled by Broadcast and Select OADMs", *APOC*, 2002.
- [25] M. Vasilyev, I. Tomkos, R. June-Koo Rhee, M. Mehendale, B. S. Hallock, B. K. Szalabofka, M. Williams, S. Tsuda and M. Sharma, "Broadcast and select OADM in 80×10.7 Gb/s ultra-longhaul network," *IEEE Photonics Technology Letters*, vol. 15, no. 2, pp. 332 -334, Feb 2003.
- [26] A. Boskovic, M. Sharma, N. Antoniadis, and M. Lee, "Broadcast and Select OADM nodes: Applications and Performance Trade-offs", TuX2, pp. 158-159, *OFC 2002*.
- [27] V. Wan Parys *et al*, *OFC 2001*, TUT4, pp. 1-3.
- [28] P. Öhlén, "Noise and Crosstalk Limitations in Optical Cross-Connects with Reshaping Wavelength Converters," *Journal of Light-wave Technology*, vol. 17, no. 8, pp. 1294-1301, August 1999.
- [29] A. Tzanakaki, K. M. Guild, D. Simeonidou, and M. J. O'Mahony, "Error-Free Transmission through 30 Cascaded Optical Cross-

- Connects Suitable for Dynamically Routed WDM Networks,” *Electronics Letters*, vol. 35, no. 20, pp. 1755-1756, Sept. 1999.
- [30] H. Mouftah and J. M. H. Elmirghani, “*Photonic Switching Technology*”, IEEE Press, USA, 1999.
- [31] M. Sharma, M. Soulliere, A. Boskovic, and L. Nederlof, “Value of Agile Transparent Optical Networks”, (invited paper), TuV1, *LEOS annual meeting*, 2002.
- [32] H. Takanishi, K. Oda, and H. Toda, “Impact of Crosstalk in an Arrayed-Waveguide Multiplexer on N×N Optical Interconnection”, *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 1097-1105, 1996.
- [33] J.K. Rhee, I. Tomkos, M. J. Li, “A Broadcast and Select OADM Optical Network with Dedicated Optical Channel Protection,” *Journal of Lightwave Technology*, vol. 21, no. 1, pp. 218-228, Feb. 2003.
- [34] J. D. Downie, I. Tomkos, N. Antoniadis, and A. Boskovic, “Effects of Filter Concatenation for Directly Modulated Transmission Lasers at 2.5 and 10 Gb/s,” *Journal of Lightwave Technology*, vol. 20, no. 2, pp. 218-228, Feb. 2002.
- [35] R. Ludwig, W. Pieper, A. Ehrhardt, E. Jahn, N. Agrawal, H.-J. Ehrke, L. Küller, and H.G. Weber: “40 Gb/s Demultiplexing experiment with a 10 GHz all-optical clock recovery using a modelocked semiconductor laser” *Electron. Lett.* 32, pp. 327-8, 1996.
- [36] B. Sartorius, “All-optical clock recovery for 3R optical regeneration”, *OFC2001*, Anaheim, USA (invited).
- [37] B. Sartorius, C. Bornholdt, O. Brox, H.J. Ehrke, D. Hoffmann, R. Ludwig, and M. Möhrle: “All-Optical Clock Recovery Module Based on a Self-Pulsating DFB Laser”, *Electron. Lett.* 34, pp. 1664-1665, 1998.
- [38] H. Yokoyama, Y. Hashimoto, H. Kurita, and I. Ogura: “All-optical subharmonic clock recovery and demultiplexing”, *OFC 2000*, ThP5, Baltimore, USA.
- [39] C. Bornholdt, B. Sartorius, S. Schelhase, M. Möhrle, and S. Bauer, “Self-pulsating DFB-laser for all-optical clock recovery at 40 Gb/s,” *Electron. Lett.* 36, pp.327-328, 2000.

- [40] H. G. Weber, R. Ludwig, U. Feiste, C. Schmidt, E. Hilliger, C. Schubert, J. Berger, M. Kroh, and T. Yamamoto, "Optical Signal Processing for High Speed Communication Systems," *Proc. 10<sup>th</sup> European Conference on Integrated Optics (ECIO)*, 4.4.-6.4.01, Paderborn, pp. 147 – 152.
- [41] T. Tekin, C. Schubert, J. Berger, M. Schlak, B. Maul, W. Brinker, R. Molt, H. Ehlers, M. Gravert, and H.-P. Nolting, "160 Gbit/s error-free all-optical demultiplexing using monolithically integrated band gap shifted Mach-Zehnder interferometer (GS-MZI)", *Technical Digest Integrated Photonic Research*, paper ICW4, July 17-19, 2002, Vancouver, Canada.
- [42] B. Sartorius, C. Bornholdt, O. Brox, H.J. Ehrke , D. Hoffmann, R. Ludwig, and M. Möhrle: "Bit-rate flexible all-optical clock recovery", *OFC'99*, FB1, San Diego, USA.
- [43] S. Bauer, C. Bornholdt, O. Brox, D. Hoffmann, M. Möhrle, G. Sahin, B. Sartorius, S. Schelhase, B. Lavigne, and D. Chiaroni, "Ultrafast Locking Optical Clock for IP Packet Switching Applications", *OFC 2000*, TuF5, Baltimore, USA
- [44] D. Chiaroni, B. Lavigne, A. Dupas, P. Guerber, A. Jourdan, F. Devaux, C. Bornholdt, S. Bauer, B. Sartorius, and M. Möhrle: "All-optical clock recovery from 10 Gb/s asynchronous data packets" *ECOC 00*, Th 10.34.5., Munich, Germany
- [45] B. Lavigne, P. Guerber, D. Chiaroni, C. Janz, A. Jourdan , B. Sartorius, C. Bornholdt, and M. Möhrle: "Test at 10 Gb/s of an optical 3R regenerator using an integrated all-optical clock.", *ECOC 99*, Th A2.3, Nice, France
- [46] B. Sartorius, C. Bornholdt, S. Bauer, M. Möhrle, P. Brindel, and O. Leclerc, "System application of 40 GHz all-optical clock in a 40 Gb/s optical 3R regenerator," *OFC 2000*, PD 11, Baltimore, USA.
- [47] B. Sartorius, J. Slovak, C. Bornholdt, H.-P. Nolting, "Novel 3R Regenerator Architecture with Wavelength Flexible Output ", *ECOC 2002*, Copenhagen, DK.
- [48] C. Bornholdt, J. Slovak, B. Sartorius, "Novel all-optical 3R regenerator concept demonstrated at 40 Gb/s," *ECOC 2002*, Copenhagen (DK), 8-12 September, 2002, PD 4.8.

- [49] M. Möhrle, B. Sartorius, C. Bornholdt, S. Bauer, O. Brox, A. Sigmund, R. Steingrüber, M. Radziunas, and H.-J. Wünsche, "Detuned grating multisection-RW-DFB lasers for high-speed optical signal processing," *JSTQE* 7, no. 2, pp. 217-223, 2001.
- [50] H.-P. Nolting and B. Sartorius, "Fast Clocked Decision – A Novel Ultra High Speed Switching Technique," *ECOC 2002*, Copenhagen, DK.
- [51] M. Moreschini, F. Matera, M. Settembre, and R. Sabella, "Wavelength assignment in optical networks according to traffic requirements and transmission impairments," *Proc. of ONDM 2001*, Wien, 4-6 February 2001.
- [52] A. Banerjee *et al.*, "Generalised Multiprotocol Label Switching: An Overview of Routing and Management Enhancements," *IEEE Communication Magazine*, Jan. 2001.
- [53] S. Demaess *et al.* "Pan-European Optical Transport Networks: an Availability-based Comparison," Accepted for publication in *Photonic Network Communications* and will be published in 2003.
- [54] M. Karásek, M. Menif, and L. A. Rusch, "Output power excursions in a cascade of EDFAs fed by multichannel burst-mode packet traffic: Experimentation and modelling," *IEEE J. Lightwave Technol.*, vol. 19, pp. 933-940, July 2001.
- [55] M. Karásek, A. Bononi, L. A. Rusch, and M. Menif, "Gain stabilization in gain clamped EDFA cascades fed by WDM burst-mode packet traffic," *IEEE J. Lightwave Technol.*, vol. 18, pp. 308-313, 2000.
- [56] M. Karásek, A. Bononi, L.A. Rusch, and M. Menif, "Effectiveness of gain control in EFAs against traffic with different levels of bursty behaviour," *IEE Proc. Optoelectronics*, vol. 147, No. 5, pp. 355-362, 2000.
- [57] M. Karásek, M. Menif, and L.A. Rusch, "Output power excursions in a cascade of EDFAs fed by multichannel burst-mode packet traffic: Experimentation and modelling," *IEEE J. Lightwave Technol.*, vol. 19, pp. 933-940, July 2001

THIS PAGE HAS BEEN LEFT BLANK INTENTIONALLY



## Chapter 3

# Circuit/Wavelength Switching and Routing

*Didier Colle, Tibor Cinkler, Sophie De Maesschalck (editors), Piet Demeester, Christoph M. Gauger, Robert Inkret, Martin Köhn, Marko Lacković, Marije Ljolje, Xavier Masip-Bruin, Mario Mattiello, Christian Mauz, Branko Mikac, Mario Pickavet, Bart Puype, Sergio Sánchez-López, Dominic A. Schupke, Josep Solé-Pareta, Slobodanka Tomić, and Qiang Yan*

### 3.1 Introduction

This chapter deals with circuit-switched optical networks. Circuit-switching means in this context that the optical network nodes switch complete wavelength channels, without interpreting the content carried in these wavelength channels. Note however that this does not necessarily exclude opaque switching operation. This chapter focuses on the network level aspects of circuit switched optical networks.

The content of this chapter is divided in four main topics. The first topic concerns control plane architectures for IP over optical networks. Not only multi-layer but also multi-domain interworking schemes at the control plane are discussed. Also a modified version of the Private Network to Network Interface (PNNI) is studied as an alternative to the popular Generalized Multi-Protocol Label Switching (GMPLS) control plane solution.

The second topic concerns the design of the optical layer. In this context mainly Routing and Wavelength Assignment (RWA) problems are investi-

gated. RWA is not only studied under static traffic, but also under dynamic and uncertain traffic demands. In addition, this chapter studies how network recovery schemes can be incorporated in the RWA.

The third topic deals with the design of multi-layer IP over optical networks. More precisely, several techniques to build the logical network topology are investigated. Moreover, also the dynamic and asymmetric nature of the traffic in IP networks is considered.

The fourth topic presents a classification of OVPNs (Optical Virtual Private Networks) and some general methods for setting them up optimally with various additional aspects including protection and grooming.

### **3.2 Control Plane Architectures for IP-over-WDM Networks**

Current research is driving the evolution of optical networks towards Intelligent Optical Networks (IONs) by investigating the prospect of introducing intelligence in the (optical) transport networks in the form of a distributed control plane. This evolution towards IONs has already been reflected in important standardization activities. The ITU has been developing an architectural framework for Automatic Switched Optical Networks (ASONS) [1], while the OIF has been standardizing the User-Network Interface (UNI) [2]. The IETF has been working on the generalization of the Multi-Protocol Label Switching paradigm (GMPLS) [3] in order to control any kind of network technology making use of slightly modified, extended versions of current IP/MPLS routing and signalling protocols.

Subsection 3.2.1 discusses control plane models for IP over Optical networks and the corresponding pros and cons of integrating/separating the control planes of both layers. Subsection 3.2.2 investigates issues regarding dynamic provisioning over multi-provider interconnected GMPLS-capable networks. Finally, subsection 3.2.3 discusses whether a PNNI-based optical control plane could act as an alternative to an MPLS-based one.

#### ***3.2.1. Control Plane Architectures for Optical Networks***

In the context of the current evolution towards a distributed control plane for (optical) transport networks, the possibility has been investigated of integrating this control plane with the control plane of client networks (i.e., typically IP/MPLS networks). The peer and the overlay control plane integration models are the two most prominent extremes based on respectively full and no integration of both control planes. As a compromise between both extremes the

augmented model [4][5] has also been investigated. The goal of this section is to revise the pros and cons of these control plane models. A more in-depth study on control plane architectures for optical networks can be found in [6] and [7], from which this section is derived.

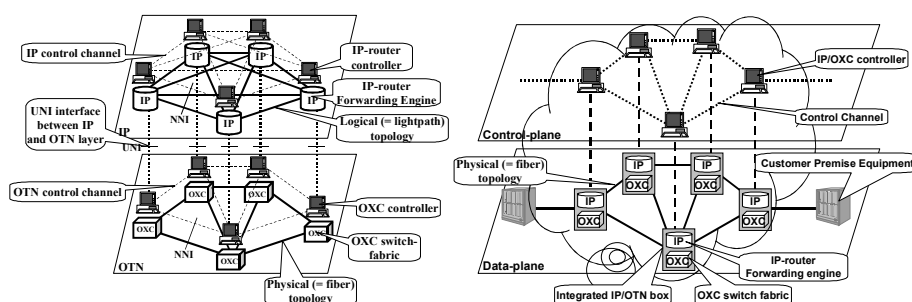
### *The overlay model*

The left side of Figure 3.1 shows the overlay model concept. In the overlay model both the IP and OTN layer have separate and independent control planes. The (control channel of the) UNI – a first version has been standardized by the Optical Internetworking Forum (OIF) – is the interface between the two control planes. Both control planes are completely independent from each other.

For coupling both layers, an IP over OTN translation protocol is provided in order to allow communication between the control planes of both layers. The IP over OTN translation protocol mainly provides address resolution between the layers and the signalling capabilities to initiate the connection setup/release, etc.

One of the drawbacks of the overlay model is the duplication of control functionality (e.g., in both layers a routing protocol has to run). Another disadvantage is the scalability problem: for each established lightpath a corresponding IP routing session has to be started. This was experienced as a problem in classical IP-over-ATM, compared to MPLS. Another drawback of the overlay model is that there is a clear client/server relationship, e.g. address resolution is required due to separate address spaces.

On the contrary, the advantage of separating both control plane instances in the overlay model is that any confidential information from the transport network is not disclosed/made accessible to any client network (operator) and that it is more straightforward to support multiple client layer networks and technologies.



**Figure 3.1. - The overlay (left) and peer (right) model**

*The peer model*

The right side of Figure 3.1 shows the peer model concept. In the peer model, a single control plane controls both the IP and OTN layers. The result is that IP router forwarding engines and OXC switch fabrics are treated as single logically integrated IP/OTN entities. So-called IP/OTN control channels are realized over the physical links between these logical IP/OTN entities. Lightpaths are treated as regular (optical) Label Switched Paths (LSPs) (in case GMPLS [3] is assumed) and thus do not result in a new peering session between the end-points (i.e., no control channel is established over the lightpath) [6].

The peer model has some important advantages. First of all, duplication of functionality is avoided. Secondly, the disadvantages of the client/server relationship between IP and OTN (e.g., problems with address resolution) do not exist anymore.

Although no additional peering session is required per established lightpath (which may solve some scalability problems: e.g., no processing of HELLO messages for each lightpath), the lightpath has to be advertised to the network as a so-called Traffic Engineering link in (G-)MPLS [8] and occupies a dedicated wavelength.

Clear drawbacks of the peer model are the following. The peer model is not applicable to all imaginable business models. For example, an Internet Service Provider (ISP) and a Transport Network (TN)-operator may not allow that the control over their network is taken over by the other one. The peer model is also limited to a single domain or Autonomous System, and thus there is no way to reduce the route computation time by dividing the network into sub-domains.

*The augmented model*

The augmented model is quite similar to the overlay model, in the sense that both layers may have their own control plane instance. However, some control information like reachability information may leak through the interface between both layers. Rephrased more practically, [4] states in what they call the “interdomain interconnection model” that the client layer reachability information is carried through the OTN, but OTN addresses are not propagated to the client network.

**3.2.2. Dynamic Provisioning over Multi-Provider Interconnected Networks**

As Automatic Switched Optical Networks (ASONS) will need to provide services of a global reach, i.e. global connectivity, its control framework will need to account not only for inter-layer (client IP-layer over server optical

layer) but also for inter-domain and inter-provider inter-working issues. To support automatic provisioning over multiple domains, all the involved domains have to have some notion of the available resources. This information can then be used for the route calculation. Also the signalling procedure, which in GMPLS works seamlessly over different layers, has to account for the fact that the path will span different domains. Again the question can be raised concerning the amount of information that needs to be exchanged in order to most optimally support the performance objectives of all the involved parties. In this context the overlay, peer and augmented control plane inter-working models, previously described, can also be considered. For the near-term requirements and the current (old) operation models, the overlay model that provides reachability information, is considered to be a satisfactory choice. On the other hand, in the long term, novel operational scenarios, e.g. such involving dynamic bandwidth trading, will leverage the flexibility of the augmented model, with the peer model being its most extreme type.

For such novel operation scenarios the flexibility of the interconnection architecture model is crucial. In the current interconnection architecture model the interconnections between domains are statically established. Applying the new paradigm of dynamic on-demand provisioning at the points of domains' interconnections, a novel interconnection architecture model can evolve. This section is derived from [9], [10], and [11] in which this novel architecture for GMPLS network interconnections is proposed and studied.

Figure 3.2 illustrates both static and extended dynamic interconnection architectures. In the static interconnection architecture the clients access the provider network at the provider edge (PE) nodes, and the domains inter-connect through domain border (B) nodes connected either directly by back-to-back links or through a cross-connect, in both cases over static configured interconnections. On the other hand, in a new Multi-Provider Edge (MPE) architecture, GMPLS-enabled cross-connects (MPE nodes) play the key role. Thereby, clients and domains connect to an MPE node, which can (as their peer and first-hop) take part in both the policy-based path calculation and the set-up of requested end-to-end connections.

Two key functionalities of an MPE node are (i) provisioning of on-demand interconnections and (ii) routing or control mediation. Thereby MPE, as a multi-layer node, may be able to provide interconnections on the different layers.

With the MPE overlay acting as a middle tear between providers and customers, and the providers themselves, we attempt to address already cited problems of providers reluctant to share their internal link information. In this case

MPE can act as a trusted mediator, collecting and customizing information. With MPE we also attempt to address the problem of inconsistent routing information, which arises, when provider domains are sharing already differently customized or aggregated information. Performance studies in [9], [10], and [11] illustrate how (i) the ability of MPE nodes to support dynamic interconnections between domains and (ii) the ability to invoke MPE services according to some policies, can improve network performance.

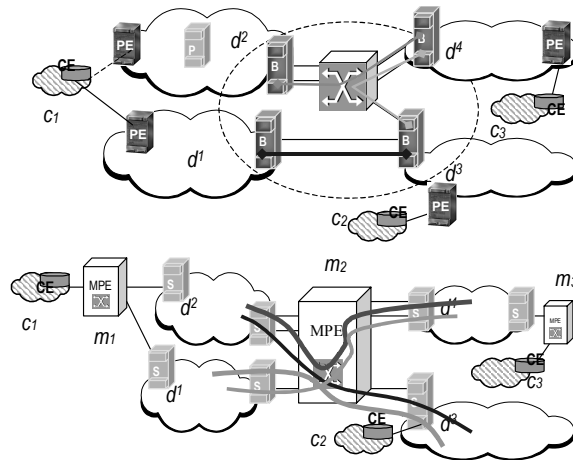


Figure 3.2. - Static Interconnection Architecture (top) and MPE model (bottom)

### 3.2.3. Comparing OPNNI versus GMPLS

GMPLS is being considered by the IETF as an extension to the MPLS Traffic Engineering control plane model to include optical networking. However, the discussion about an optical control plane based on the ATM Private Network to Network Interface (PNNI) [12] paradigm has started in standardisation bodies. PNNI is expected to be suitable for ASON after some appropriate modifications. As a mature technology, the PNNI can be very practical for a seamless migration from current transport networks to ASON.

Optical PNNI (O-PNNI) is an adaptation of the ATM PNNI to be an alternative control plane to the one provided by GMPLS. O-PNNI fits with the overlay model separating both the IP and OTN control planes. As mentioned above, no IP client can access the topology information from the transport network. Moreover, O-PNNI can be a control plane based on the augmented model, in which the client layer reachability information is carried through the OTN. In order to achieve this aim, the O-PNNI uses a modified PNNI Augmented Routing (PAR) [13]. The main modifications consist basically of

defining a set of PNNI Topology State Elements (PTSEs) to transport information of each client connected to the OTN.

GMPLS uses IPv4 and/or IPv6 addresses. However, O-PNNI uses Network Service Access Point (NSAP) addresses, so no common address space may be built with IP based networks.

O-PNNI supports Soft Permanent Virtual Connections (SPVCs), one of the ASON requirements. O-PNNI incorporates a Connection Admission Control (CAC) to indicate whether an optical node can admit a new connection. In GMPLS, the support of CAC does not seem to be well developed at this moment.

A hierarchical structure allows the O-PNNI to support multiple levels (up to 104) and therefore is scalable for very large networks. In GMPLS, OSPF and IS-IS only support two hierarchical levels.

ASON specifications recommend using separated control and data planes. GMPLS adds additional signalling information to the original MPLS signalling to support separated planes [14]. A similar effort has to be done in O-PNNI in order to adapt the PNNI, in which the control plane is merged with the transport plane, to the ASON requirements.

Service discovery is a concept introduced by the OIF. It consists in querying transport service characteristics before the optical UNI signalling establishes the connection. Neither O-PNNI nor GMPLS support service discovery.

In GMPLS, OSPF and IS-IS use Traffic Engineering specific extensions to propagate QoS information. These parameters are stored at each node, and a modified Constrained Shortest Path First (CSPF) protocol computes a path through the network. The O-PNNI routing protocol includes support for QoS characteristics to determine the optimal path. Moreover, O-PNNI incorporates a “crankback” mechanism to support rerouting around a failed component at the connection setup.

Finally, Table 3.1 presents a general comparison of GMPLS and O-PNNI according to the optical control plane requirements.

**Table 3.1. - Comparison of GMPLS and O-PNNI**

<b>CONTROL PLANE REQUIREMENTS</b>	<b>GMPLS</b>	<b>O-PNNI</b>
<b>Addressing</b>	Yes	Yes
<b>Connection Admission Control</b>	No	Yes
<b>Hierarchical structure</b>	2 levels	Up to 104 levels
<b>Separated control and data plane</b>	Yes	No
<b>Service Discovery</b>	No	No
<b>Path selection with QoS</b>	Yes	Yes

### 3.3 Designing Optimised Reliable WDM Networks

#### 3.3.1. Static Wavelength Assignment in WDM Networks with OXCs

##### *Introduction*

Wavelength assignment is a major dimensioning step, especially in the case of large and highly loaded WDM networks as the ones studied in this COST action. That is the reason for the growing need for algorithms and tools able to perform this operation.

This section will give an overview of the most efficient algorithms and the results they have provided when implemented in CANPC [15], described in chapter 6.3.

##### *Algorithms*

Two wavelength assignment algorithms have been considered. The first one, TABUCOL, has been conceived at TCOM, and is presented in [16] and [17]; the second one, Minimum Hop RWA has been proposed in [18]. Both of them perform routing and wavelength assignment separately. In this way, the exact minimum number of fibres, for a given routing scheme, can be calculated (for each link, it is enough to divide the number of route channels by the capacity, in wavelengths per fibre, and to take the ceiling integer). Thus, after the wavelength assignment, one can check whether the solution is optimal or not (i.e., if it uses the minimum number of fibres).

TABUCOL routing is based on an adaptive shortest path process, in which each link has an occupation weight that grows with the number of channels already routed on it. It performs wavelength assignment by means of a stochastic procedure based on the well-known graph colouring approach.

M.H.RWA searches for all the shortest disjoint paths, in terms of the number of nodes, for each demand, and then, recursively, chooses among them the one that provides the lower congestion (occupation of the most charged link). Then, it simply browses the demands in decreasing order of path length, and assigns them to the first free wavelength found.

##### *Results*

The main aim of the simulations has been to route the studied network keeping the total number of fibres as low as possible, and then to achieve an optimal wavelength assignment without using wavelength converters. Computations have been done on all the five COST 266 topologies (see chapter 1.3), and results for the basic topology are reported in the following table.



**Table 3.2. - Numerical results of the simulations on COST 266 basic topology**

			TABUCOL	M.H.RWA
16 wl/ fib	Total fiber length (km)		174568	168858
	Fibers per link	Max	8	7
		Avg	3.61*	3.48
		Min	1	1
	Fiber occupation		83.1%	86.2%
	Channels per link	Max	108	101
		Avg	47.99	47.98
Min		7	7	
32 wl/ fib	Total fiber length (km)		94188	95817
	Fibers per link	Max	4	4
		Avg	1.96	1.96
		Min	1	1
	Fiber occupation		77.2%	76.4%
	Channels per link	Max	105	100
		Avg	48.52	47.98
Min		13	7	
64 wl/ fib	Total fiber length (km)		59282	61241
	Fibers per link	Max	2	2
		Avg	1.20	1.22
		Min	1	1
	Fiber occupation		63.7%	61.4%
	Channels per link	Max	104	100
		Avg	48.71	47.98
Min		13	7	

\* Non-optimal solution (some fibers have been added to complete the assignment)

Both algorithms provide satisfying results, with a high fibre occupation. With one exception, they reach optimal solutions. Moreover, they seem to complement each other; in fact, TABUCOL performs better with a higher number of wavelengths per fibre (higher fibre occupation, lower number of fibres), while M.H.RWA performs better with a lower number of wavelengths per fibre.

Finally, it is very important to point out that, because of the optimal assignment, the use of wavelength converters would not have reduced the total number of used fibres.

### ***3.3.2. Routing in Optical Networks under Inaccurate Network State Information***

In an ASON, the control plane includes a lightpath control mechanism to efficiently set up and tear down lightpaths, which may be either centralized or distributed. In spite of the fact that adaptive routing mechanisms based on

global information perform better than the ones based on local information, they are only suitable for those networks where frequent network changes are not expected. Therefore, assuming distributed lightpath control in this scenario, maintaining precise global network state information on each node is almost impossible. Many factors, such as non-negligible propagation delay, frequent state updates and hierarchical topology aggregation, can affect the precision of the global network state information. Thus, when the lightpath selection process is performed based on outdated routing information, a connection blocking increment is produced.

The *BYPASS Based Optical Routing* (BBOR) [19] is a new adaptive source routing mechanism that computes dynamically explicit lightpaths in an ASON without conversion capabilities. Its aim is to reduce the connection blocking probability due to perform routing and wavelength assignment decisions under inaccurate global network state information.

The BBOR mechanism includes two main aspects, namely a triggering policy adapted to the RWA problem to reduce the routing signalling, and a bypass routing algorithm to counteract the effects of the routing inaccuracy produced by this routing signalling reduction.

Existing triggering policies are based on updating by either a periodical refresh or sending an updating message whenever there is a change in the network state. The BBOR mechanism introduces a new triggering policy based on a threshold value, which aims to include the network congestion, namely the available network resources. A network node triggers an updating message whenever a fixed number of  $N$  wavelengths change their status. By changing the value of  $N$ , we can evaluate the impact of different degrees of inaccuracy in the connection-blocking ratio.

The bypass routing algorithm dynamically reroutes the setup messages through an alternative pre-computed explicit route (bypass-path), when at any of the intermediate nodes the requested wavelength is found not available at the lightpath setup time. Three routing algorithms are derived from the BBOR mechanism named ALG1, ALG2 and ALG3. While ALG1 and ALG2 are developed to be applied to wavelength-selective networks [19], ALG3 is applied to wavelength-convertible networks [20].

The BBOR mechanism consists of three basic processes: (1) decide which wavelength of which link (bundle of  $B$  fibres) might be bypassed, (2) include these wavelengths as a parameter to be considered when selecting the lightpath and (3) compute the bypass-paths.

### 3.3.3. Protection in the Optical Layer for Mesh Networks

#### Introduction

There exist several possibilities to perform optical layer protection in mesh networks [21]. A distinction can be made with respect to the entity that is protected, namely path protection or span protection. Normally, path protection is more difficult to realize because, in case of a failure, a reconfiguration of all OXCs of affected working paths has to be performed. When the span (or link) is protected, only neighbouring OXCs have to be reconfigured. In the latter case, it is also much easier to detect the occurrence of a fault. It is even possible to reroute the complete optical multiplex section signal containing several wavelengths by simple fibre switching. On the other hand, the length of the lightpath may significantly increase. This results in a degradation of the signals and may require their regeneration.

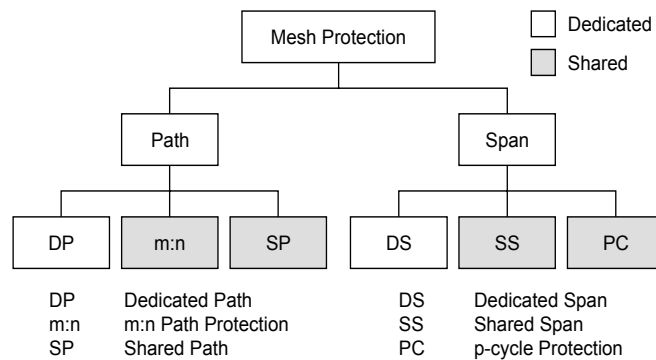


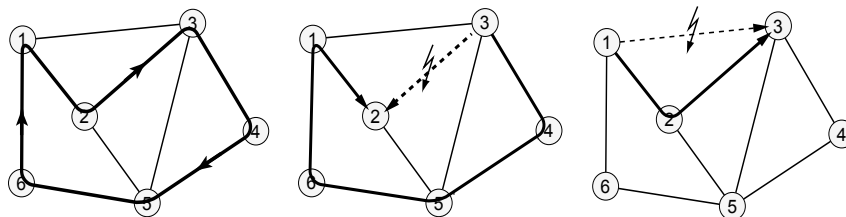
Figure 3.3. - Classification of different protection schemes in mesh networks.

In figure 3.3, a classification of common protection schemes in mesh networks is shown. A very simple approach is dedicated path protection (DP). In DP, for every working path, a protection path is reserved. This scheme is also denoted as 1+1 protection. When  $n$  working paths share one protection path, this is named 1: $n$  protection or, more generally, when  $m$  protection paths are available,  $m:n$  protection. When all protection paths can use a common pool of protection resources, this is called shared path protection (SP). In general, this is the most capacity-efficient protection scheme.

Similar to the path protection schemes, dedicated span protection (DS) and shared span protection (SS) exist.

A special case of span protection is the p-cycle protection scheme (see Figure 3.4), introduced by Grover and Stamatelakis [22]. The protection capacity is organized in cycles. Similar to a self-healing ring, when a failure of an on-

cycle link occurs, the signal passes the cycle in the counter-propagating direction. The difference to an SDH shared protection ring (SPRing) is that also the straddling links may benefit from the protection by the cycle. Therefore, p-cycle protection is much more capacity-efficient than protection schemes based on rings.



**Figure 3.4. - p-cycle protection. From left to right: network with established p-cycle, failure of an on-cycle and a straddling link.**

These schemes have in common that the protection resources are planned in advance and that for the occurrence of a failure a specific reaction is foreseen. In contrast to that, resilience procedures must start from the actual network state and the available resources in order to re-establish broken connections.

#### *Unified ILP-Formulation*

It can be shown that for all these mesh protection schemes, a unified formulation as an ILP-problem can be adopted [23]. Here the objective is to minimize the required capacity of the network. There must be sufficient spare capacity available to establish the intended protection scheme and to protect all entities. Often, the capacity requirements for a survivable network may be significantly reduced when the routing of the working paths is included in the optimization process of the protection resources. A pre-selection of the possible protection structures, e.g., a restriction to the  $k$  shortest possible protection paths for every working path or span, helps to reduce the complexity of the ILP-problem.

For dedicated protection schemes and for  $m:n$  path protection, it is possible to divide the problem into less complex sub-problems for which simpler exact algorithms than the ILP-formulation exist.

A case study for the different protection schemes has been performed using CPLEX 7.0 to solve the corresponding ILP-problems. The total traffic matrix for the year 2002 and a granularity of STM-16 per channel have been assumed. To compare the results, the required capacity of the different protection schemes has been normalized to the necessary working capacity of the networks when using shortest path routing.

For path protection, two failure scenarios are investigated, namely a link failure, which requires an edge disjoint (ed) protection path, and the occurrence of link or node failures, which may be counteracted by a node disjoint (nd) protection path. Note that a node disjoint path implies an edge disjoint path.

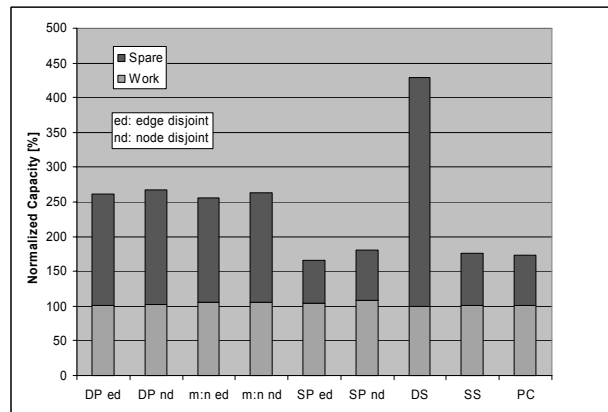


Figure 3.5. - Capacity requirements for the BT network.

The results for BT network (see chapter 1.3) are shown in Figure 3.5. For dedicated path protection, more than 160% additional capacity is required. To avoid trap-constellations, a slightly longer working-path than the shortest path has to be used. For node disjoint protection paths, 7% more capacity is needed compared to edge disjoint protection paths.

These figures are only slightly improved by  $m:n$  protection. In contrast to that, for shared path protection only 165% capacity has to be installed for working and protection traffic. Dedicated span protection needs over 329% protection capacity. Here, sharing among the spans reduces this to 74%, resulting in slightly more capacity than for shared path protection. For a  $p$ -cycle scheme, the required protection capacity accounts to 71%, due to the better implicit pre-selection of possible protection candidates. Thus, a better solution in the restricted search space may be found.

Because the RT topology has a lower mean nodal degree, one needs significantly more protection capacity than for the BT network. Especially, dedicated span protection performs badly. Here, a resilient network requires at least 285% total capacity when using shared path protection. For the TT network, the requirements are just the other way around. Due to the high meshing degree of the topology, only 47% capacity is required to make the network survivable. In this case, shared path and shared span protection perform equally well.

### *Heuristics*

When the nodes have no wavelength conversion capabilities, every path and protection structure gets coloured (gets a specific wavelength). In this case, the ILP approach is unfeasible even for small networks, and one has to rely on heuristics with lower complexity. Here, as an example, two such algorithms are explained.

In order to reduce the problem complexity, the planning of shared path protection for a given routing may be divided in two steps [24]:

- forming of a set of paths which may share protection resources, in the following referred to as PSG (Protection Share Group), and
- setting-up and allocating the needed spare resources for one PSG.

The first sub-problem may be mapped on a graph colouring problem. This is similar to the wavelength allocation problem. In fact, when only link failures are investigated, both problems are equivalent.

The second sub-problem may be solved by successively finding the shortest path in the network with special weights for the edges. These weights are adapted in each step. Another possibility is to use an ILP-formulation for each PSG. This is much simpler than to tackle the whole problem in one step.

In general, the obtained solution is sub-optimal. Therefore, the required capacity is slightly higher than that obtained by solving the ILP-problem. On the other hand, the properties of the PSG allow the simplification of the node-switching functionality.

When using  $p$ -cycle protection, the chosen wavelength assignment has a significant impact on the sharing of the protection resources [25]. When the wavelengths are chosen in an unfavourable way and a link consists of only one pair of fibres, sharing is essentially only possible among straddling links of the cycles. Therefore, it is of crucial importance to perform the wavelength allocation and the  $p$ -cycle search jointly in order to benefit the most from the capacity savings of  $p$ -cycles.

The developed algorithm is similar to the *First-Fit* assignment of wavelengths and works as follows. The working paths are ordered according to their hop-count and are assigned wavelengths one after the other. In each step, sufficient  $p$ -cycles for the chosen wavelength are added until all spans are protected. A scoring function is used to choose among different possible cycles.

In general, due to the wavelength continuity constraints, between 40% and 60% additional resources are required for a network without wavelength conversion. The reason is that sharing of protection resources may only occur

among paths with the same wavelength. The higher the mean nodal degree of the network, the lower will be this penalty. Another approach to avoid wavelength blocking with  $p$ -cycles is to use more fibres [26].

#### 3.3.4. Multi-Ring Networks

Today's transport networks are often organized in rings. Ring structures depend on simple optical add-drop multiplexers (OADM) and are easier to realize than mesh networks from a technological point of view. On the other hand, the planning of multi-ring-networks with several interweaved rings is very complex. In rings, the protection resources may be dedicated (dedicated protection rings, DPRing) or shared (shared protection rings, SPRing).

In [27], [28] an optimized algorithm to deduce an optimal ring-coverage is developed. Here, the rings are constructed out of faces, a concept coming from graph-theory. This allows to gradually change them.

From the ring-configuration, an effective graph of the network is constructed. It describes the effective logical topology, which results from the set of rings. With a fast routing and dimensioning algorithm, the rings are designed. The configuration is evaluated according to an objective function. Simulated annealing (SA) is used for the optimization. Beginning with a valid start configuration, in each step a neighbour configuration is derived from the actual configuration by choosing at random one of the operations:

- adding one face to one of the rings,
- creating a new ring with a random face,
- deleting a face of a ring (when there are no faces left, the ring is destroyed).

It is possible to incorporate boundary conditions, e.g., the maximum number of nodes in a ring, into the objective function of the optimization process. This algorithm is especially suited for large and highly meshed networks, where lots of possible cycles exist, and, therefore, the complexity of an approach that pre-computes these cycles is high.

#### 3.3.5. Uncertainty in the Planning Process

The uncertainty in the forecast of the future traffic, which is traditionally used for network planning, and its high dynamics require a systematic treatment of the vagueness of the traffic pattern. To model the vagueness of the future traffic, the entries of the traffic matrix are random variables with a distribution  $T_\alpha \sim D_\alpha(\mu_\alpha, \sigma_\alpha)$ . The expected value  $\mu_\alpha$  is set equal to the predicted traffic.

In this way, an ensemble of possible realizations of different traffic patterns may be treated simultaneously [29].

From the distribution of the traffic demands of a node pair, the probabilities for the realization of a specific connection may be calculated. Each request is routed through the given topology, resulting in a distribution for the required capacity of each link. For the dimensioning, the expectation value of the distribution on that link plus an amount in excess proportional to the variance is taken. This can be performed in a systematic way by choosing a confidence level, which is the probability that the link is able to carry the traffic.

By combining several statistically independent traffic sources, the relative variance and, therefore, the required capacity are minimized. In that way, it might be favourable to choose a longer path in order to share a link with other connections. By optimizing the routing, the best trade-off between longer paths and higher variances is searched. For the deterministic static case, the best solution for the routing is to assign the shortest path to each connection.

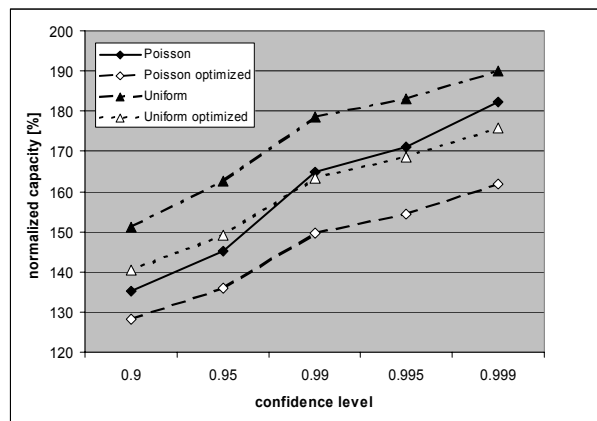


Figure 3.6. - Required capacity for different confidence levels.

In Figure 3.6, the results of a case study of a Pan-European network and different distributions are shown. The capacity is normalized to the deterministic case. Due to the uncertainty of the traffic pattern, a substantial excess of capacity in the network is necessary. With increasing confidence level, the required capacity rises. For example for a confidence level of 0.999, approximately twice the capacity of the deterministic case is needed. The higher relative variance of the uniform distribution results in a much higher required capacity for the same confidence level.

With this very simple optimization procedure, capacity savings of 5% to 10% may be achieved. The reductions are greater for the uniform case. In general,



the shortest path is assigned to the connections that occur with high probability, whereas the less likely ones take a longer path to share capacity with other connections.

This design approach tends to bundle the traffic and leads to “unbalanced” networks. Some links will carry a lot of traffic whereas peripheral links are less utilized. This might cause problems with protection and restoration.

### 3.4 Designing IP over WDM Networks

This section focuses on modelling, designing and dimensioning IP-over-WDM networks, including solutions assuming intermediate layers like SDH/SONET. Starting from a generic approach for IP-over-WDM network modelling, we present work on designing static WDM networks for transporting IP traffic and compare several data link layer topologies. Then, routing and grooming in IP transport networks are studied. Traffic engineering concepts for dynamic WDM networks and the influence of asymmetric IP traffic finalize this section.

#### 3.4.1. An Approach to an IP-over-WDM Transmission Network Modelling

##### *IP over WDM network structure*

The IP over WDM (Figure 3.7) network has been modelled using two main node types: *IPPoP* (IP Point of Presence) node, and *WDMPoP* (WDM Point of Presence) node.

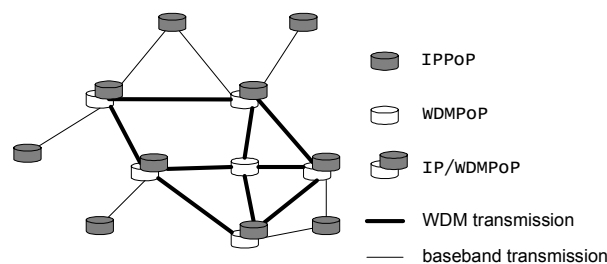


Figure 3.7. - IP over WDM network structure

An *IPPoP* node represents an aggregation point for the IP traffic, thus generating IP traffic for the WDM transmission network. It can be viewed as a backbone router, which aggregates IP traffic from many sources and lower levels of aggregation (e.g. edge routers in metropolitan areas). A WDM network is composed of *WDMPoP* nodes serving as ingress and egress points

for IP traffic entering and leaving a transmission network. They also provide switching functionality as intermediate nodes inside the WDM network. *WDMPoP* nodes are connected with optical fibres, carrying WDM signals.

#### *IP layer modelling*

An *IPPoP* node can directly contain two module types: IP node, and IP router, as shown in Figure 3.8. An *IPPoP* node serves as a source and destination of the IP traffic. An IP router aggregates and splits IP traffic destined for different *IPPoP* nodes. IP nodes and IP routers are connected using line cards. Each line card has two interfaces. One is an electrical interface and serves as an interface towards the IP network equipment. The other interface is optical and handles the signal from/to transmission media.

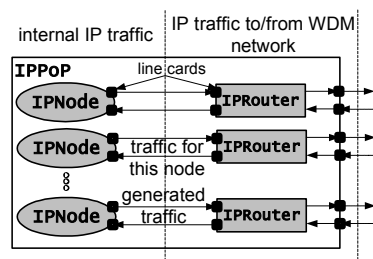


Figure 3.8. - *IPPoP* structure

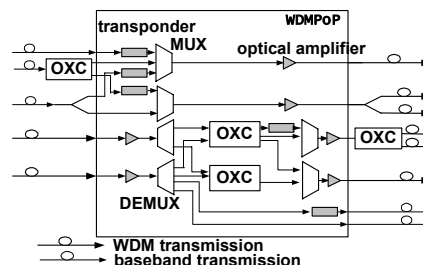


Figure 3.9. - *WDMPoP* structure – an example

#### *WDM layer modeling*

A WDM network consists of *WDMPoP* nodes connected with optical links carrying WDM signals. *WDMPoP* nodes contain several types of optical components. Optical components implement the functionality needed in the WDM network, including protection and restoration (P&R) mechanisms, and automatic channel switching. Figure 3.9 shows an example of a *WDMPoP* node, which includes all optical elements in all possible interconnections. The *WDMPoP* structure is determined by the IP traffic demands, the *WDMPoP* connections and the P&R mechanism. The P&R mechanism determines the WDM equipment redundancy in the *WDMPoP* nodes, as well as the link redundancy between *WDMPoP* nodes. Some mechanisms (restoration) require dynamic switching (dynamic change from impaired path to alternative), which requires OXC usage.

#### **3.4.2. Static Lightpath Routing in IP over WDM Networks**

The first IP-over-WDM networks that are emerging use OXCs to interconnect IP routers that communicate with each other over lightpaths [30]. Intermediate

multiplexing layers such as ATM or SDH are omitted, thus the whole available bandwidth of the lightpath is available [30]. Given the bandwidth demands between IP router sources and destinations, a feasible and efficient (static) configuration of the lightpaths is needed. This is referred to as finding an optimal virtual topology seen by the IP layer and a routing and wavelength assignment (RWA) on the WDM layer.

We split this problem into several sub-problems and solve these by heuristic algorithms. As laser sources currently constitute a large portion of the network element costs, we are interested in using as few sources (and thus lightpaths) as possible.

In the considered architecture, the OXCs are interconnected by fibre pairs, one for each direction. For each fibre the wavelength constraint has to be fulfilled. An OXC is further connected to at most one IP router, which is source and sink for at least one lightpath. We assume that the originated and terminated lightpaths can be selected from a set of wavelengths. As we do not consider wavelength conversion at the WDM nodes, a lightpath retains its wavelength.

Two routers become neighbours by a bidirectional virtual link (formed by two lightpaths, one for each direction) that offers the same amount of bandwidth capacity. The two contra-directional lightpaths do not necessarily need to have the same wavelength. Interconnected routers operate using a routing protocol. Today, routing protocols are shortest path based such as OSPF. As IP packets experience a delay at routers, the number of router hops should be restricted. We consider this by allowing only a maximum number of virtual links on the shortest path of a source-destination pair. Further constraints are discussed in [31].

Of course, the offered total network load has to be carried. We assume here that we are given a demand matrix with bandwidth entries. Each entry is lower or equal than the maximum capacity of a lightpath. In non-synchronous multiplexing environments which holds for IP we may add some margin to the demand entries.

In the following subsection we present the general solution approach and apply it to a network example. More details can be found in [31].

### *Solution Approach*

Due to the RWA problem alone, the whole problem becomes too complex to be solved by means of integer linear programs [30], [32]. Therefore we divide the problem into three sub-problems, and solve each by a heuristic algorithm.

In the first step we find a virtual topology such that the number of (virtual) links is minimized and the given IP demand is carried in a greedy manner [31]. Minimizing the virtual links corresponds to minimizing the lightpaths demanded from the WDM layer and thus the laser sources.

In the next step, we route the lightpath demands in the optical network using a shortest path algorithm.

In the last step we have to assign wavelengths to the paths taking into account the wavelength constraint in fibres and the wavelength continuity constraint in OXCs. In general, one is given a set of wavelengths which can be used network-wide, thus it is sufficient to obtain one feasible assignment. The wavelength assignment problem is related to the graph colouring problem [32]. Efficient graph colouring algorithms exist, which do not only perform a colour assignment but also minimize the number of colours (wavelengths). We used this optimization approach for the wavelength assignment to provide efficient wavelength utilization throughout the network. It must be noted that there may be cases where the minimum number of colours obtained exceeds the cardinality of the wavelength set given through the system. The approach for the wavelength assignment is the heuristic algorithm “DSATUR” [31].

#### *Case Study*

We used the described approach to configure the topology of the NSF network (14 nodes, 21 links) used in [33], where we associated each link with a fibre pair and each node with a router connected to an OXC. We also used the traffic matrix in [33] scaled such that the entry with the maximum value equals 10 Gbit/s, yielding a total network load of 120 Gbit/s. The bit rate capacity of a wavelength was set to 10 Gbit/s.

In the case without constraints on the number of hops we obtain 19 virtual links (this is half of the number of laser sources) and 4 wavelengths. The number of 19 virtual links is close to 13, the minimum number of links needed to connect the network. The utilization of the wavelengths is half of the capacity. This is still reasonable, since many demand-pairs in [33] carry much more traffic in one direction than in the other. Moreover, all nodes have to be connected, even those that do not source or sink much traffic.

Figure 3.10 depicts the number of virtual links and the number of wavelengths when restricting the maximum number of allowed IP hops in the virtual topology algorithm. Both functions are decreasing monotonously and from 1 hop to 2 hops, the function graphs drop down very much, since IP multiplexing becomes possible. More results are included in [31].

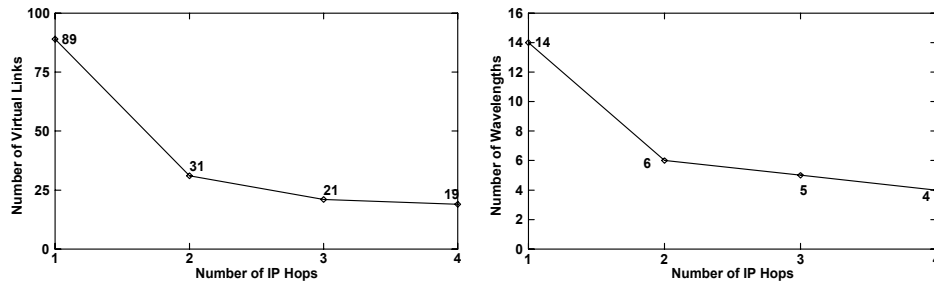


Figure 3.10. - The number of virtual links and the number of wavelengths over the number of IP hops

### 3.4.3. Design of Optimal Data Link Layer Topology for IP/WDM Networks

Now, we will focus on the connectivity of IP backbone routers, i.e. the layer 2 topology. Thus, we do not take into account the cost of WDM equipment [34]. In addition, the layer 2 topology does not necessarily follow the physical topology.

We assume the physical topology to be given, i.e. the length and placement of the optical fibres are known. The traffic matrix between layer 2 nodes is also known and traffic demands in layer 2 are routed along the shortest paths in layer 1 (i.e., the optical network layer). We finally assume that the number of fibres on the links is unlimited.

In [35], three scenarios to obtain a layer 2 topology are presented:

1. *1-to-1 mapping between layer 1 and layer 2 topology:*  
the layer 2 topology follows the physical topology; in [35] proposed for low traffic loads.
2. *full mesh layer 2 topology:*  
each router is connected directly to all other routers, and the minimum number of required interfaces equals the sum of the degree of all nodes; in [35] found to be the best solution for high traffic loads.
3. *optimal topology:*  
could give the best trade-off between the number of interfaces and their cost, here the focus is on finding the optimal topology.

For large networks, finding an optimal solution is an NP-hard problem. Hence, a genetic algorithm (GA) is used to obtain a near-optimal layer 2 topology.

We consider only two scenarios for mapping IP on WDM: IP/POS/WDM and IP/GbE/WDM. From our point of view, the application of the chosen scenarios is most probable in the near future [34].

*Generic Layer 2 P&R Capabilities*

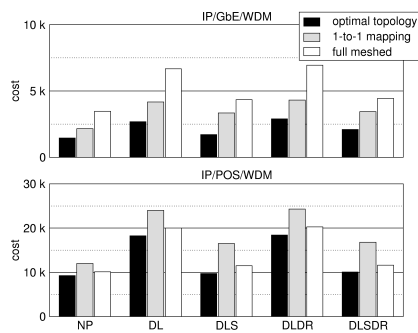
In order to simplify the network design, only some generic P&R scenarios obtained from the first two layers are introduced:

- A *Dual link:*  
router interfaces are doubled and placed onto two shortest independent physical paths in layer 1
- B *Dual link with dual router configuration:*  
all router interfaces are distributed over two independent routers instead of on a single router

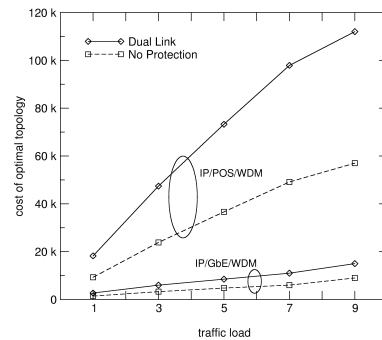
Additionally, traffic between interfaces can be shared, but in case of a failure, the load on the remaining interface increases. Here we assume that in case of the shared protection type, the traffic load is equally divided between the working and backup interface.

*Results and Comments*

The described algorithm has been applied to a network sample taken from the case study developed within COST 239 and COST 266 (see chapter 1.3). We assume that the traffic matrix is given in terms of required 1 Gbit/s channels between layer 2 nodes.



**Figure 3.11. - Cost relationship between proposed scenarios (load=1)**



**Figure 3.12. - Costs of optimal topologies for different traffic loads**

Topologies with dual link and dual link dual router type of protection are more expensive, but their performance is better in terms of potential packet loss. - shows the cost relationship between the scenarios described above (NP-No Protection, DL-Dual Link, DLS-Dual Link Shared, DLDR-Dual Link Dual

Router, DLSDR-Dual Link Shared Dual Router). It can be seen that in the IP/GbE/WDM scenario the optimal topology gives a substantial reduction in price compared to the other two proposed scenarios, whereas in the IP/POS/WDM scenario the cost of the optimal topology is approximately equal to the cost of a fully meshed topology. The sum of the primary lightpath lengths shows that topologies with shared protection (DLS and DLSDR) consist of less links of layer 2 than other types of protection.

Figure 3.12 shows the cost of an optimized topology for a given load. It can be seen that with growing traffic load, the cost of optimal topology increases almost linearly.

#### ***3.4.4. Dynamic Routing and Grooming in SDH/WDM Multi-Layer Networks***

As mentioned before, one possibility for an IP transport network is a dynamic SDH network on top of a dynamic optical network (IP/SDH/WDM). This is from the viewpoint of network evolution one of the most probable architectures for future IP backbones. Assigning low-bandwidth electrical connections to high-speed optical lightpaths, also known as traffic grooming, is an important aspect in such networks [36]. Routing on the electrical and optical layer is an essential component of every grooming strategy. Efficient transport of dynamic traffic demands requires optimised multi-layer routing and grooming algorithms. Non-integrated routing schemes treat both layers separately while integrated schemes try to improve the performance by combining both layers. Most approaches for integrated routing, e.g. [37] (SONET) or [38] (IP), take advantage of complete knowledge about topology and occupancy of both layers, which is not supported by the overlay network model [39]. However, the overlay model is favoured from an operator point of view as information on transport networks is very sensitive and should be kept secret.

In the following we assume the mean IP traffic to be given for each node-by-node connection. We have to groom and route this traffic through a dynamic SDH/WDM multi-layer network.

##### *Multi-layer Grooming and Routing Options*

The SDH/SONET-WDM-network approach leads to a multi-layer node model. This model comprises an optical cross-connect (OXC) on the optical layer, which is assumed to be non-blocking. The electrical layer mainly consists of a non-blocking electronic cross-connect (EXC), which is able to switch electrical connections at an arbitrary granularity. The EXC thus allows for an effective grooming of electrical connections. The EXC and OXC are connected by  $z$  transponders.

In a centralized routing scheme, the routing control centre has to choose a path within the network for each connection request. Four basic grooming options can be identified:

- A *single-hop grooming on existing lightpath*:  
Assign connection to one existing direct lightpath.
- B *multi-hop grooming on existing lightpaths*:  
Route on the electrical layer by using more than one existing lightpath.
- C *single-hop grooming on new lightpath*:  
Set up a new lightpath and route connection via this lightpath.
- D *combined multi-hop grooming on new and existing lightpaths*:  
Combination of A and C. Route connection by using a series of existing and new lightpaths.

Non-integrated routing schemes are capable of grooming on either existing or new lightpaths. Only with integrated routing does the routing control centre have enough information to also perform the combined grooming described in D. As a reference we will consider two different non-integrated routing schemes, namely PreferOptical and PreferElectrical (TLRC in [37]), which differ in the order they apply the different grooming policies until one succeeds (PreferOptical: A-C-B, PreferElectrical: A-B-C).

In contrast to the non-integrated routing schemes, WIR applies the different grooming strategies in parallel, including combined grooming. For a connection request, each possible path that results from a grooming strategy is rated according to a set of criteria. Finally, the path with the best rating is chosen. Finding paths with combined grooming is harder than applying grooming purely on the electrical or optical layer. In [37], Zhu and Mukherjee use a reachability graph capturing all possible links for a given connection request between any two nodes for their proposed SLRC scheme.

#### *Simulation Studies*

The presented simulation study was performed using a fictitious 9-node German network, the topology of which is shown in Figure 3.13. This network was introduced and dimensioned for static traffic demands [40] such that links contain a certain number of fibres each holding 8 wavelengths. For more details on the simulation scenario see [36].

For a multi-layer node, the number  $z$  of transponders is a crucial parameter, from a performance as well as a cost point of view. Particularly, for a given network topology, transponders are the major variable cost factor. Hence, we introduce the fraction of transponders installed as the absolute number of



transponders installed normalized by sum of all possible transponders in all nodes. For more details on this see [36].

Figure 3.14 shows the blocking probability for different schemes over the fraction of transponders installed. The non-integrated routing scheme PreferOptical outperforms the scheme PreferElectrical by almost one order of magnitude. The integrated WIR achieves a request blocking probability that is about one order of magnitude below the best non-integrated scheme and slightly below SLRC. There is a noticeable bend in all curves at a fraction of about 0.55. For lower fractions, the request probability is dominated by blocking at the transponders. For larger fractions, there are sufficient transponders available so that wavelength blocking dominates.

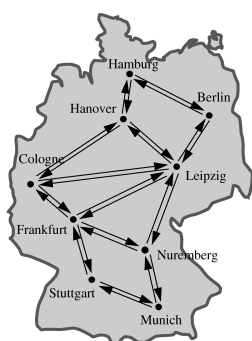


Figure 3.13. - Fictitious 9-node network of Germany

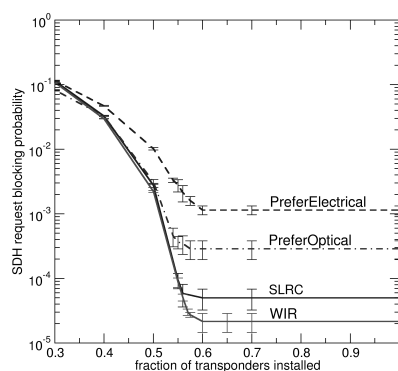


Figure 3.14. - Blocking probabilities in a SDH/WDM network

### Conclusions

The impact of routing and grooming algorithms on the performance of multi-layer networks is not negligible. They reduce the blocking probability in the network as well as increase the usage of the wavelength channels.

#### 3.4.5. Multi-Layer Traffic Engineering in Data-Centric Optical Networks

The logical IP topology in multi-layer data-centric optical networks is supported by lightpaths, which can be set up or torn down dynamically in Intelligent Optical Networks (IONs). While regular traffic engineering (TE) typically only routes traffic in the IP layer, the ability to reconfigure the logical IP topology on demand leads to cross-layer traffic engineering capabilities.

*What is Multi-layer Traffic Engineering and how does it work?*

Multi-layer Traffic Engineering allows reconfiguring the logical resources. To realize this however, a MTE process goes through three distinct phases [37]. Firstly, the traffic over the logical topology needs to be monitored, either physically on the router interfaces, or by keeping track of bandwidth specification included in signalling messages. Problems may be detected, optionally signalled through the network, so as to trigger the next step, the decision taking phase.

Here, the MTE strategy will find an appropriate reconfiguration of the logical topology, given the monitoring data. It will decide which logical IP links to set up and tear down, and also which traffic to attract to newly established links.

Lastly, once the reconfiguration has been decided on, corresponding MTE actions still need to be performed during the realization phase. The logical network layer will use signalling to request new lightpaths from the optical layer, or to remove existing lightpaths. It will also adjust IP routing to make sure the new logical topology is used efficiently.

*Issues in Designing Multi-layer Traffic Engineering Strategies*

Strategies can be classified according to a number of properties, some of them having analogies in classic TE. One aspect is the overall architecture of the network and strategy. Calculations may run either online or offline, and in the latter case, the strategy can be centralized, as opposed to being distributed. Also important is the distinction between peer (also called integrated) and overlay models in multi-layer networks. Furthermore, algorithms themselves can vary in complexity, because of the number of alternative actions considered, and the extent to which multi-layer actions interact with Single-layer Traffic Engineering (STE) actions (i.e. routing). A good TE strategy needs to integrate both MTE and STE aspects.

The specific objective of the strategy is another property. Not just what this objective is, but also how it is reached; either by keeping the network optimal at all times (a proactive strategy), or by only reacting to certain problems that have been detected first (a reactive strategy).

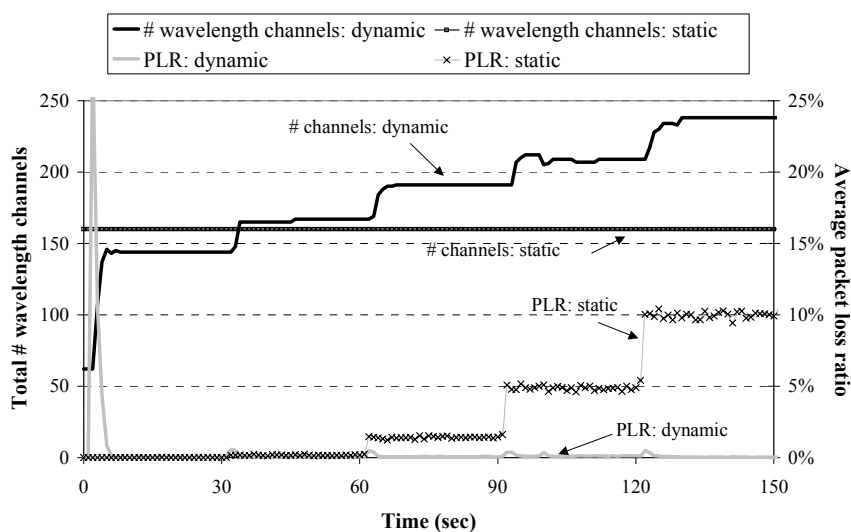
Further issues exist however, a certain amount of inertia may need to be introduced in the strategy to lower the number of lightpath operations, at the cost of a somewhat lower optimality of the logical topology. Care has to be taken then to avoid cumulative memory effects and other potential instabilities.

### *Benefits of Multi-layer Traffic Engineering*

Data from a case study highlights the benefits of MTE. A reactive strategy was chosen. It defines both lower and upper thresholds for bandwidth usage of an IP-link. Exceeding these bounds triggers the process which removes under-loaded links or sets up new links to remove congestion. The choice of the new link and also the amount of traffic attracted over it, are optimized so that load on routers is reduced as much as possible, and bandwidth utilization of both previously congested and newly established link approaches an ideal mean between both thresholds.

Figure 3.15 compares the performance of this strategy (i.e., the dynamic case) with the situation where the topology is statically dimensioned. Traffic volume is increased in steps, and from 60 sec. onwards, exceeds the nominal bandwidth volume for which the static case was dimensioned, leading to rising packet loss ratio's in this case. The reactive strategy however can keep PLR satisfactory small by provisioning additional lightpaths, at the cost of slightly higher optical bandwidth usage under nominal traffic (sec. 30 to 60).

More details on this case study and further studies can be found in [41], [42] and [43].



**Figure 3.15. - Analysis of the benefits of MTE**

### *Conclusions*

Multi-layer Traffic Engineering strategies deploy and profit from the additional flexibility provided by Intelligent Optical Networks. MTE algorithms

can be approached from a number of different perspectives, and once some design issues have been taken care of, they yield significantly better performance than strategies which use statically dimensioned logical topologies.

### 3.4.6. Influence of Traffic Asymmetry on the Network Design

IP traffic is, in contrast to telephone traffic, asymmetric. The degree of asymmetry depends of course on the exact location in the network (access versus backbone network, in the surrounding of a server farm or not, etc.). Today's backbone network is designed to support bidirectional services like SONET/SDH. When this bidirectional and symmetric traffic is replaced by IP traffic, large portions of the network bandwidth may sit idle, while the bandwidth at the opposite direction can, at the same time, be completely congested. Using unidirectional WDM line-systems instead of the currently in use bidirectional ones, could thus lead to a significant cost reduction. In this section, this cost advantage will be quantified for three of the COST266 optical backbone topologies discussed in the introductory chapter. More details can be found in [44], [45], [46] and [47], from which most of this section is extracted.

#### *Assumptions and definitions*

First the traffic model explained in chapter 1.3 has to be extended to include an asymmetric IP traffic demand. This was based on the number of Internet hosts in each region ( $U_i$ ) and the number of web sites stored in servers located in these regions ( $W_i$ ):

$$IP\ traffic_{i \rightarrow j} = \frac{2 * W_i^\alpha * U_j}{W_i^\alpha * U_j + W_j^\alpha * U_i} * IP\ traffic_{i-j}$$

$$IP\ traffic_{j \rightarrow i} = \frac{2 * W_j^\alpha * U_i}{W_i^\alpha * U_j + W_j^\alpha * U_i} * IP\ traffic_{i-j}$$

where  $\alpha$  allows to change the asymmetry of the traffic. We define the Asymmetry Factor (AF) of a traffic matrix M as

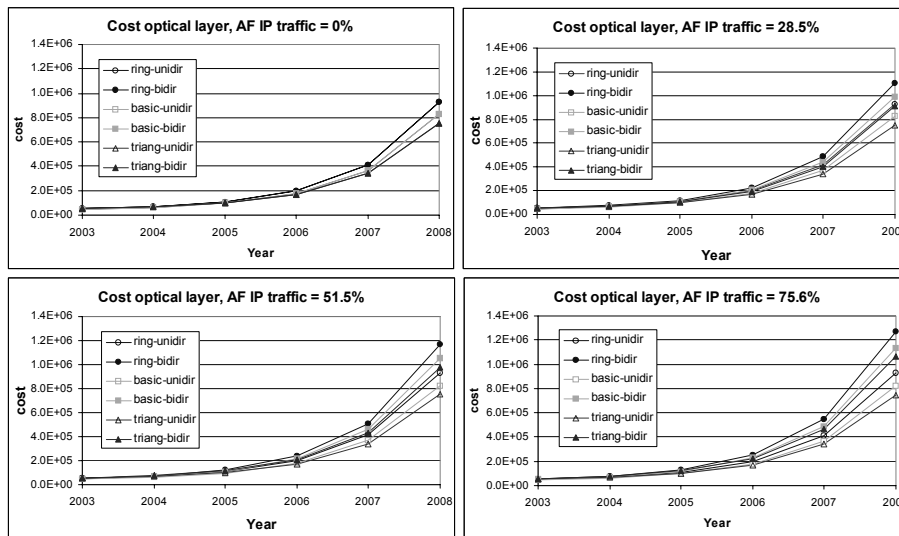
$$\frac{\sum_{i,j,i \neq j} |M_{ij} - M_{ji}|}{\sum_{i,j,i \neq j} |M_{ij} + M_{ji}|}$$

These equations allowed us to construct traffic matrices with different AFs for the time frame 2003-2008.

The capacity installation problem in the optical layer was then solved for (a) the case that the line-systems in the optical layer are bidirectional and (b) the case that these line-systems are unidirectional, and this for the BT, RT and TT. The cost of a bidirectional line-system was assumed to be twice the cost of a unidirectional one. The logical layer on top of these optical networks was in all cases assumed to be a full mesh.

*Results*

The result of these network designs is shown in Figure 3.16. Regardless of the AF of the traffic demand, the design for the ring network is always the most expensive one, while that of the triangular network is the least expensive. This is due to the used cost model that includes a proportional part of the digging cost in the link cost. The cost of a network design is then determined by the total length of used fibre, and not by the total duct length. Such a cost model would be used by a network operator that leases capacity from a carrier's carrier. The total length of used fibre is largest for the RT, as the connections have to follow on average a longer path between source and destination than in the BT and TT (shortest path in km.).



**Figure 3.16. - Cost of the unidirectional and bidirectional optical layer network design for varying asymmetry factor of the traffic demand**

When the AF of the IP traffic differs from 0, the designs with bidirectional line-systems are always more expensive than the design with unidirectional line-systems. The use of unidirectional line-systems instead of bidirectional

ones would be 36% cheaper for the TT, 38% for the BT and 43% for the TT. This is of course due to the fact that the average filling of the bidirectional line-systems decreases as the AF of the traffic demand increases: only one direction of the bidirectional line-systems gets properly filled. With unidirectional line-systems, less number of line-systems can be installed on one direction of a link than on the other, and all line-systems are thus properly filled up.

### 3.5 Providing Optical VPN Services

Instead of providing leased line services, or providing wavelength channels or even dark fibres, it seems more useful in certain cases to provide a *virtual private network*. The term *virtual network* denotes that the network is not built physically and separately, but is only an allocated part of the resources of a public network of a provider. It is *private* since it serves a closed group of users.

In general an OVPN (Optical Virtual Private Network) is a VPN built over optical channels, e.g., wavelength paths within a Multi-hop Wavelength Routing (WR) Dense Wavelength Division Multiplexing (DWDM) Network.

The open objective is how to set these OVPNs in such a manner that all user needs are satisfied while using as few network resources as possible.

#### 3.5.1. Overview

Figure 3.17 shows an example optical VPN, where the core is optical, with MPLS routers attached to it. For connecting certain MPLS Routers (E), wavelength paths established using optical cross-connects are used. Since these routers are not connected in a full mesh by wavelength paths, the optical edge switches (O1-O4) will have to provide MPLS router functionality as well between certain domains and users. For example, when connecting E1 to E4 and E2 to E3 the OVPN will consist of three lightpaths crossing some cross-connects and four MPLS routers in nodes O1, O2, O3 and O4.

Here we have exploited the advantages of traffic grooming, i.e., numerous MPLS traffic streams (LSPs) of a VPN can share a single wavelength path.

The objective of the optimisation is in all cases to reduce resource usage at higher (electrical) layers (i.e., to reduce the load of the virtual routers), subject to the constrained amount of capacity of each wavelength channel and limited number of wavelengths.

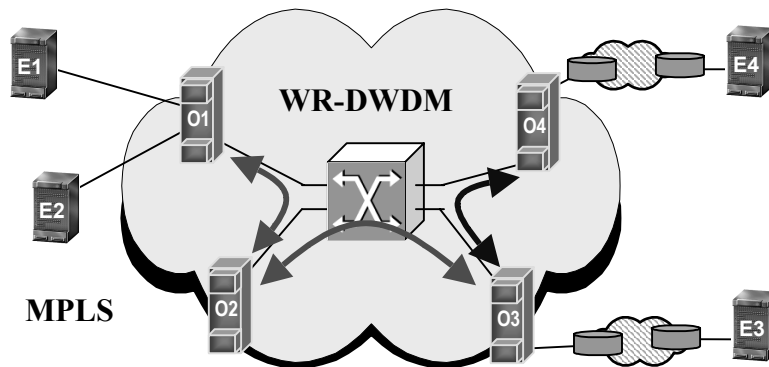


Figure 3.17. - Illustration of an OVPN

The deployment of Virtual Private Networks (VPNs) increases steadily. Reference [48] reports three IP VPNs by GlobalOne, Infonet and Worldcom/UUnet. The last two are based on the MPLS technology.

The deployment of VPNs grows in Enterprise networks according to [49]. They report on the use of managed VPNs by Research, Insight-Research and NovaStor and add that the use of VPNs results in significant revenue. MPLS VPNs can be used jointly with BGP to avoid limitations of addressing plans and to ensure security [50]. This paper also mentions provisioning end-to-end VPN services across multiple service providers and carriers. Reference [51] gives an excellent overview of technologies enabling VPNs and QoS support for VPNs (particularly DiffServ) and management alternatives.

A detailed overview of MPLS based VPNs can be found in [52] and [53].

VPNs share the link bandwidth and the node resources among each other. This idea has several advantages. We do not have to build separate private physical networks, but only to configure VPNs. This reduces costs and speeds up provisioning. Furthermore, the VPNs can be dynamically reconfigured or re-dimensioned in contrast to physical networks. This allows sharing resources between various VPNs.

A VPN-DiffServ solution is proposed in [54]. In [55] dynamic relations are in scope with capacity resizing and stochastic fair sharing, but without protection. The resource allocation in conjunction with the routing design has been analyzed in [56], [57] and [58] over multi-service networks with QoS constraints. Various tools are used, e.g., asymptotic approximations to reduce the complexity of the numerical calculations, multiplexing inside a VPN and introducing priorities between the traffic classes. Network dimensioning is

addressed in [59] and the methodology is presented for determining the sizes of VPNs.

An algorithm of very low complexity is presented in [60], however, the traffic streams are not handled separately (pipe model), but an aggregate traffic which has source or sink in one network node are handled jointly (hose model). The protection is not handled at all.

All above papers configure VPNs over a network capable of performing multiplexing, e.g., Frame Relay, ATM, MPLS or any kind of multiservice networks. However, there are few methods for establishing VPNs over optical networks, and even less for two or more layers, which are of particular interest for GMPLS and ASTN networks.

There are multiple papers discussing the architecture and configuration of VPNs [61][62]. However, very few papers deal with the protection of these VPNs.

### ***3.5.2. The Three Types of OVPNs Considered***

We differentiate 3 types of OVPNs denoted as T1, T2 and T3 based on the granularity of these VPNs:

- **T1:** "VPNs over WR-DWDM": Here we assume that we build conventional VPNs over an optical (Wavelength Routed-DWDM) network. Links of different VPNs can share a single wavelength path.
- **T2:**  $VP\lambda N$ : This is the most widely accepted type of OVPNs. Here a VPN link is a wavelength path, and one wavelength path serves a single VPN only. This wastes the resources in some cases, however the control is simple. If the capacity of a wavelength path is not sufficient to accommodate a link of the VPN we allow use of multiple (even parallel) wavelength paths. The advantage is that the traffic of different VPNs is separated (and isolated), which simplifies management and enhances security. Compared to T1 it also decreases the load of the electrical layer.
- **T3:**  $VP_{multifibre}N$ : This is the enhanced version of T2 OVPNs for multifibre networks, where we assume that a point-to-point demand is a wavelength path, while a link consists of as many wavelength channels of a certain wavelength  $\lambda$  as there are fibres that support that wavelength.

If we have demands of low capacity, T1 is preferred, while for point-to-point connections approaching the capacity of a wavelength T3 is preferred.

In all cases we assume traffic grooming [63], i.e., traffic streams of a VPN are groomed, i.e., multiplexed, e.g., in case of T2 into a single wavelength path,



and they are handled together. In all Virtual Router nodes these traffic streams can be re-multiplexed. This can be done at the electrical layer only since re-multiplexing, i.e., time division multiplexing is required. For this reason, taking not only the optical, but both, optical and electrical layer into account when configuring the system is demanded.

### ***3.5.3. Three Methods for Setting up OVPNs Optimally***

Many excellent papers deal with design, configuration and optimisation of WDM Networks. These methods can be generalised for the design of OVPNs, as well.

Our goal is to configure the VPNs and the lightpath system optimally without separating the network layers. This improves the quality of the results; however, the complexity of the problem grows.

Based on the level of decomposition we differentiate three methods denoted as M1, M2 and M3 for the OVPN configuration and they are described in detail in [64] and [65].

- **M1: Global:** First we formulate the problem formally as an Integer Linear Program (ILP) for the directed graph model in order to find the optimal configuration of all OVPNs simultaneously. Although this gives the globally optimal solution it is useless, due to its complexity. Therefore, less complex methods are needed that provide results close to the global optimum.
- **M2: VPN-by-VPN:** Here we decompose the problem to setting up VPNs one-by-one. The complexity is significantly lower; however, the quality of the results can be poor. For example setting up the first VPN can hinder setting up some other VPN by monopolising some of the common resources. Furthermore, this approach heavily depends on the order in which the VPNs are set up. This method is very useful if the demands for VPNs are not known in advance, appear one-by-one, and have to be satisfied without waiting for the other demands for VPNs that will arrive.
- **M3: Demand-by-Demand:** Here we decompose the whole problem to routing the demands one-by-one by a shortest path algorithm, e.g., that proposed by Dijkstra. This is the simplest method; however, the results might be poor. Here we need some heuristic tricks. A promising approach is to sort demands, and to start by those that need larger bandwidth, and are limited in length. Another possible heuristic approach is to decrease the cost of the links that are already used by the considered VPN when one new demand of that VPN has to be routed. The third heuristic decreases

the costs of those wavelengths that are already used by the considered VPN. Using these three heuristics jointly improves the obtained results significantly, [65] presents these results.

#### **3.5.4. Protection Alternatives for OVPNs**

We consider two protection alternatives for all the above OVPN types (T1-T3) and methods (M1-M3): protection at the higher (electrical) layer and at the lower (wavelength) layer. We refer to these methods as Internal and External OVPN Protection. In case of Internal protection, the OVPN users get enough resources not only to carry their traffic in a failure-free case, but also to switch to alternative paths in case of a failure. This will of course require adequate resource management and protection mechanisms. External protection is simpler for users, but more complex for the operators. In this case the operator must protect the failed spans of certain OVPNs, to "hide" the failures from the OVPN users.

The protection paths can be made either link or node disjoint with the working path. Furthermore, dedicated and shared protection can be differentiated.

### **3.6 Conclusions/Summary**

In this chapter four main topics in designing WDM-based networks have been discussed. The focus was limited to circuit-switched optical networks. The impact of the client network – in the future this will typically be an IP-MPLS network – and recovery capabilities were often taken into account due to their importance.

The first topic studied concerns control plane architectures for such networks. There exist several control plane models for IP over Optical networks. In the overlay model, both layers have an independent control plane. The peer model integrates both control planes into a single integrated control plane controlling both layers (which is mainly favourable from a technical perspective). The augmented model is a compromise between both extremes.

In order to enable more advanced and flexible provisioning scenarios of a global reach (e.g. bandwidth trading, dynamically provisioned re-configurable virtual networks of global reach), domains of different ASON providers will need to inter-work. For this purpose the novel MPE architectures, incorporating dynamically provisioned inter-domain interconnections, and leveraging flexibility of peer and augmented interconnection schemes, has been proposed and studied.

As an alternative to a GMPLS-based optical control plane, an Optical PNNI (OPNNI) control plane has been studied. Both have technical pros and cons, and both need many extensions and adaptations because neither of them completely supports all the optical control plane requirements. Apart from technical features and requirements, the final decision will depend on political and economical aspects.

The second topic studied concerns the routing and wavelength assignment (RWA) in WDM networks. In a first case study, two novel RWA algorithms have been compared with each other. TABUCOL searches the shortest path based on a routing weight that increases with an increasing number of occupied wavelength channels on a link. M.H.RWA searches all disjoint shortest paths and selects the less congested one. It was shown that the former one performs best in case of a rather high number of wavelength channels per fibre, while the latter performs best in case of less wavelength channels per fibre. It was also found that the use of wavelength converters would often not reduce the number of required fibres. In order to deal with dynamic traffic in the optical network, the BYPASS Based Optical Routing (BBOR) technique is proposed. This technique reduces the signalling overhead by flooding new link states after a fixed number of wavelength channels changed their status and provides bypass routes to circumvent potential blocking on links defined as being congested.

In a second part, the RWA-problem is extended to incorporate network recovery. It has been shown that the capacity efficiency of shared-span protection and p-cycles is comparable to that of shared path protection. p-cycles differ from shared protection rings in the sense that their backup resources are not only dedicated to protect on-cycle but also straddling links. The capacity efficiency of dedicated path protection and  $m:n$  protection are also comparable to each other but is significantly higher than that of the previously mentioned recovery schemes. Finally, dedicated span protection requires the highest amount of spare capacity. It was also found that the wavelength continuity constraint in network without wavelength converters severely impacts the capacity efficiency of p-cycles. Not only the impact of network recovery in meshed networks was investigated, but also a heuristic has been developed to design multi-ring networks.

In the last case study, it was shown that the impact of the uncertainty – modelled as a probability distribution function – on the expected traffic volumes may become very large, when a tight confidence interval is considered.

The third topic studied in this chapter differs from the second one in the sense that it studies the design of multi-layer IP-over-WDM networks. In such a network, the network can consist of an IP Point-of-Presence (IPPoP) node and/or of a WDMPoP node. Most of the work was dedicated to study techniques to optimize the logical network topology. A distinction is made between static and dynamic techniques.

For the static case, the first case study showed that limiting the number of hops along the shortest path in the logical IP network may significantly increase the number of logical links needed in the logical IP network and thus also the number of required wavelengths in the underlying optical network. In the second case study it was shown that optimizing the logical network topology might imply a significant reduction of the overall network cost in an IP/GbE/WDM scenario, while it does not improve the network cost – compared to a fully meshed topology – in an IP/PoS/WDM scenario.

For the dynamic case, the first case study – dealing with SDH/WDM networks – showed that when a connection request cannot be supported by an existing lightpath, setting up a new lightpath for the connection leads to a smaller blocking probability than routing the connection over a chain of existing lightpaths. By having an integrated routing strategy, allowing the connection to be routed over a chain of existing and new lightpaths, the blocking probability is reduced even further. This case also showed the important impact of the fraction of installed transponders on the blocking probability. In the second case study, the benefit of dynamically reconfiguring the logical IP network was illustrated: it allows keeping the packet loss ratio reasonably low, while the physical resources are only used in case they are really needed. In addition to that, some other issues concerning dynamic reconfiguration of the logical IP network have been discussed (e.g., the need to introduce some inertia, in order to avoid instabilities).

In the last case study, the impact of the typical asymmetric nature of the IP traffic has been investigated: the higher the traffic asymmetry, the more beneficial it becomes to deploy unidirectional line-systems, compared to bi-directional line-systems having the same amount of capacity in both directions.

The fourth topic studied in this chapter classified optical VPNs, and proposed methods for their configuration including protection alternatives. Since this problem is very complex, it must be decomposed into smaller sub-problems, which can be solved for networks of practical interest. One of the most useful decomposition strategies is to route the demands one-by-one, sorted according to their bandwidth requirements and distance, as well as to apply a link-cost scheme, that prioritizes the use of wavelengths and links that are already used

by demands of the considered VPN. The demand-by-demand decomposition allows using any method for routing and wavelength assignment or protection described in the preceding sections of this chapter. All these results are applicable to single layer static and dynamic networks such as the optical transport network (ITU-T OTN) with the automatic switched optical network (ITU-T ASON) control plane as well as to general multi-layer transport networks, with the automatic switching capabilities such as those proposed for the automatic switched transport network (ITU-T ASTN) control plane or a generalized multiprotocol label switched (IETF GMPLS) control plane.

### 3.7 Issues Open for Further Research

With respect to the first topic studied in this chapter, i.e. control plane architectures for IP-over-WDM networks, future research will need to address the development of novel algorithms (e.g., for traffic engineering), which use the protocols, currently under development and standardization, in an intelligent way.

Also until now, the conducted studies with the MPE model focused on the effects improving the blocking probability of the services (single-domain and multi-domain VPNs) in the 2-layers (TDM/WDM) network for different 2-layers routing and resource sharing strategies. Future research will need to focus on the control plane architectural components, their functionalities and the way they organize in the distributed MPE overlay to support different interconnection models. In this context modelling of the multi-layer exchange point, which is the basic building block of this architecture, needs to be given special attention. The aim is also to generalize this 2-layers model to a multi-layer hierarchy. At last, the work on provisioning aspects of global VPNs over the MPE architecture needs to be continued.

Future research should also focus on improving and concretizing the O-PNNI protocol to support all the optical control plane requirements. This improvement will consist of:

- Specifying which parameters to advertise within the proposed PNNI Topology State Element (PTSE) in order to disseminate optical network information throughout the hierarchical structure.
- Adding signalling information to support a separated control and data plane.

- Supporting service discovery to allow transport service characteristics to be queried before a connection establishment takes place.

With respect to the second topic studied in this chapter, i.e. the design of reliable WDM networks, future research should investigate how to extend the RWA algorithms TABUCOL and Minimum Hop RWA in order to incorporate network recovery techniques. It will be crucial that this is done in such a way that the performance of these RWA algorithms does not deteriorate.

Future research should also focus on extending the BBOR mechanism in order to make it applicable to networks with conversion capabilities: this may result in novel routing algorithms. Based on this, the impact of the conversion capabilities on the blocking probability will be examined: this may lead to the conclusion that the BBOR mechanism can act as potential alternative solution for improving the network performance to the solution that simply adds wavelength converters. Another research topic might concern analyzing the potential usage of the BBOR mechanism when proposing integrated routing strategies which take into account the collaboration of both the IP and optical layers.

Taking into account the available signal and switching capabilities of optics, it would be interesting to develop new protection schemes specifically tailored to the WDM layer. Efficient heuristics are needed when an ILP formulation is not tractable.

Another key issue for future research is the robustness of the network design when the dynamism of the traffic pattern in the transport networks increases and when traffic patterns become more unpredictable.

### 3.8 References

- [1] ITU-T Rec. G.8080, “*Architecture for the automatically switched optical network (ASON)*,” November 2001.
- [2] “*User Network Interface (UNI) 1.0 Signaling Specification*,” OIF2000.125.3, June 2001.
- [3] P. Ashwood-Smith, *et al*, “Generalized Multi-Protocol Label Switching (GMPLS) Architecture,” Internet-draft February 2001, <http://search.ietf.org/internet-drafts/draft-many-gmpls-architecture-00.txt>
- [4] B. Rajagopalan, *et al*, “IP over Optical Networks: Architectural Aspects,” *IEEE Comm. Mag.*, vol. 38, no. 9, pp. 94-102 (2000).

- [5] L. Raptis, *et al*, "Design and Experiments of an Automatic Switched Optical Network (ASON)," *Proc. of 27<sup>th</sup> European Conference on Optical Communications, ECOC 2001*, vol. 3, Amsterdam, Netherlands, Sept. 30 - Oct. 4, 2001, pp. 256-257.
- [6] D. Colle, M. Jaeger, *et al*, "IP-based control plane architectures for optical networks," *Proc. of 2<sup>nd</sup> Int. Workshop on All-Optical Networks, WAON 2001*, Zagreb, Croatia, 13-14 June, 2001, pp. 7-15.
- [7] D. Colle, M. Jaeger, *et al*, "Developing control plane models for optical networks," *Technical Digest 2002 Optical Fiber Communication Conference (OFC2002)*, ISBN 1-55752-700-8, March 17-22, 2002, Anaheim, CA, USA, pp. 757-759.
- [8] K. Kompella and Y. Rekhter, "LSP Hierarchy with MPLS TE", internet-draft March 2001, <http://search.ietf.org/internet-drafts/draft-ietf-mpls-lsp-hierarchy-02.txt>.
- [9] S. Tomic and A. Jukan, "Dynamic Provisioning over Multi-provider Interconnected GMPLS-enabled Networks," *Proc. of 7th IFIP Working-Conference on Optical Network Design and Modelling (ONDM 2003)*, Budapest, Hungary, Feb. 3-5, 2003.
- [10] S. Tomic and A. Jukan, "MPFI: The Multi-provider Network Federation Interface for Interconnected Optical Networks," *Proc. of IEEE Global Telecommunications Conference, GLOBECOM 2002*, Taipei, Taiwan, Nov. 17-21, 2002.
- [11] S. Tomic and A. Jukan: "Policy-based Lightpath Provisioning Over Federated WDM Network Domains," *Proc of Optical Fiber Communication Conference 2002, OFC 2002*, Anaheim, California, pp. 499 – 501, March 17-22, 2002.
- [12] ATM Forum, af-pnni-0055.002, "Private Network to Network Interface Specification Version 1.1", April 2002.
- [13] ATM Forum, "PNNI Augmented routing (PAR)", Version 1.0, af-ra-0104.000, January 1999.
- [14] P. Ashwood-Smith, *et al*, "Generalized MPLS – Signaling Functional Description," internet-draft, May 2001.
- [15] D. Tarongì, *et al*, "Computer Aided Network Planning Cockpit CANPC," *Proc. of 2<sup>nd</sup> Int. Workshop on All-Optical Networks, WAON 2001*, Zagreb, Croatia, 13-14 June, 2001.

- [16] C. Bungarzeanu, J.Ehrensberger, and D.Tarongi, "Planning Tool for Wavelength Assignment in Static WDM Transport Networks," *Proc. of IEEE ICT 2001*, Bucharest, Romania, June 4-7, 2001.
- [17] P. Strahm and C. Bungarzeanu, "Planning Tool for WDM Networks with All-Optical Packet Switching", *Proc. of 2<sup>nd</sup> Int. Workshop on All-Optical Networks, WAON 2001*, Zagreb, Croatia, 13-14 June, 2001.
- [18] S. Baroni and P. Bayvel, "Wavelength Requirements in Arbitrarily Connected Wavelength-Routed Optical Networks," *Journal of Lightwave Technology*, vol. 15, no. 2, Feb. 1997.
- [19] X. Masip-Bruin, *et al*, "An Adaptive Routing Mechanism for Reducing the Routing Inaccuracy Effects in an ASON," *Proc. of the Seventh Working Conference on Optical Network Design and Modelling (ONDM'03)*, Budapest, Hungary, pp. 333-349, Feb. 2003.
- [20] X. Masip-Bruin, *et al*, "Routing and Wavelength Assignment under Inaccurate Routing Information in Networks with Sparse and Limited Wavelength Conversion", *Proc. of IEEE GLOBECOM 2003*, San Francisco, USA, Dec. 2003.
- [21] G. Maier, *et al*, "Optical network survivability: Protection techniques in the WDM layer," *Photonic Networks Communication*, vol. 4, no. 3-4, pp. 251-69, July-Dec. 2002.
- [22] W. Grover and D. Stamatelakis, "Cycle-oriented distributed preconfiguration: Ring-like speed with mesh-like capacity for self-planning network restoration," *Proc. of the IEEE International Conference on Communications, ICC 1998*, vol. 1, pp. 537-543.
- [23] C. Mauz, "Unified ILP Formulation of Protection in Mesh Networks," *Proc. of 7<sup>th</sup> Int. Conf. on Telecommunications, ConTEL 2003*, Zagreb, Croatia, June 11-13, 2003, pp. 737-741.
- [24] C. Mauz, "Allocation of Spare Capacity for Shared Protection of Optical Paths in Transport Networks," *Proc. of the Third Int. Workshop on Design of Reliable Communication Networks, DRCN'01*, Oct. 2001.
- [25] C. Mauz, "p-cycle Protection in Wavelength Routed Networks," *Proc. of the Seventh Working Conference on Optical Network Design and Modelling, ONDM'03*, Feb. 2003.



- [26] D. Schupke, C. Gruber, and A. Autenrieth, "Optimal configuration of p-cycles in WDM networks," *Proc. of the IEEE International Conference on Communications*, pp. 5:2761–2765, 2002.
- [27] C. Mauz, "Optimized Ring Identification for an Arbitrary Network Topology and Traffic Matrix," *Proc. of 2<sup>nd</sup> Workshop on All-Optical Networks*, WAON 2001, Zagreb, Croatia, June 13-14, 2001, pp.17-24.
- [28] C. Mauz, "Mapping of Arbitrary Traffic Demand and Network Topology on a Mesh of Rings Network," *Proc. of the Fifth Working Conference on Optical Network Design and Modelling (ONDM'01)*, Feb. 2001.
- [29] C. Mauz, "Dimensioning of a Transport Network Being Robust to Changes of the Traffic Pattern," *Proc. of the Forth Int. Conference on Transparent Optical Networks (ICTON'02)*, Apr. 2002.
- [30] N. Ghani, S. Dixit, and T.-S. Wang, "On IP-over-WDM integration," *IEEE Communications Magazine*, vol. 38, no. 3, March 2000.
- [31] D.A. Schupke and D. Sellier, "Lightpath Configuration of transparent and static WDM Networks for IP Traffic," *IEEE International Conference on Communications ICC 2001*, Helsinki, Finland, June 11-15, 2001.
- [32] R. Ramaswami and K. N. Sivarajan, "*Optical Networks: A Practical Perspective*," Morgan Kaufmann Publishers, San Francisco, California, 1998.
- [33] R. Ramaswami, and K. N. Sivarajan, "Design of logical topologies for wavelength routed optical networks," *IEEE Journal on Selected Areas in Communication*, vol. 14, no. 5, June 1996.
- [34] M. Ljolje, R. Inkret, and B. Mikac, "Design of Optimal Data Link Layer Topology for IP/WDM Networks", *Proc. of 2<sup>nd</sup> Int. Workshop on All-Optical Networks*, WAON 2001, Zagreb, Croatia, June 13-14, 2001, pp.27-32.
- [35] R. Inkret, "*Protection and Restoration Strategies for IP over WDM Networks*," Short Term Mission Report, COST 266 "*Advanced Infrastructure for Photonic Networks*", France Telecom R&D, Lannion, France, 2000.
- [36] M. Necker, C. Gauger, and S. Bodamer: "A New Efficient Integrated Routing Scheme for SDH/SONET-WDM Multilayer Networks", *OFC2003*.

- [37] K. Zhu and B. Mukherjee: "On-Line Approaches for Provisioning Connections of Different Bandwidth Granularities in WDM Mesh Networks", *OFC 2002*, Anaheim.
- [38] M. Kodialam and T.V. Lakshman: "Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks", *INFOCOM 2001*, New York City, vol. 1, pp. 358-366.
- [39] E. Modiano, "Traffic Grooming in WDM Networks", *IEEE Communications Magazine*, vol. 39, no. 7, 2001, pp. 124-129.
- [40] J. Späth, "Dynamic routing and resource allocation in WDM transport networks", *Computer Networks*, vol. 32, No. 5, 2000, pp. 519-538.
- [41] J. Comellas, *et al*, "Traffic engineering beyond the IP-MPLS layer", Poster session at IST-LION/IST-OPTIMIST Workshop on "Intelligent High-Capacity Optical Networks: opportunities and challenges", Oct. 2002, Turin, Italy.
- [42] Q. Yan, *et al*, "Applying Intelligent Optical Networking Functionality in the deployment of the Next-Generation Internet", *LEOS Benelux 2002*, Amsterdam, The Netherlands, Dec. 2002.
- [43] B. Puype, *et al*, "Multi-layer traffic engineering in data-centric optical networks", *COST266/IST Optimist Workshop at ONDM 2003*, Budapest, Hungary, February 2003.
- [44] S. De Maesschalck, *et al*, "Influence of the IP Traffic Asymmetry on the Cost of the Optical Network Layer", *Photonic Network Communications*, vol. 4, No. 2, May 2002, pp. 133-150.
- [45] S. De Maesschalck, *et al*, "How the Asymmetric Nature of IP Traffic influences the Cost of the underlying Optical Layer", *Fiber and Integrated Optics*, vol. 22, No. 2, Feb. 2003.
- [46] S. De Maesschalck, *et al*, "Asymmetric IP Traffic and its Consequences for the Optical Layer", *ECOC 2001*, Oct. 2001, Amsterdam, The Netherlands.
- [47] S. De Maesschalck, *et al*, "Influence of the Asymmetric Nature of IP Traffic on the Cost of the Optical Layer," *Proc. IPoW*, June 2002, Paris, France.
- [48] "Global IP VPN Landscape Changing Rapidly," *IEEE Communications Magazine*, vol. 39 no.3, p. 64, March 2001.
- [49] "Necessity is the Mother of VPN Invention," *IEEE Communications Magazine*, vol. 39, no. 9, p. 52, Sept. 2001.

- [50] A. R. Moral, P. Bonenfant, and M. Krishnaswamy, "The Optical Internet: Architectures and Protocols for the Global Infrastructure for Tomorrow," *IEEE Communications Magazine*, vol. 39, no. 7 pp.152-159, July 2001.
- [51] T. Braun, M. Guenter, and I. Khalil, "Management of QoS Enabled VPNs," *IEEE Communications Magazine*, vol. 39 no. 5 pp. 90-98, May 2001.
- [52] J. Guichard and I. Pepelnjak, "*MPLS and VPN Architectures*," Cisco Press, 2001.
- [53] B. Davie and Y. Rekhter, "*MPLS - Technology and Applications*," Academic Press, 2000.
- [54] I. Khalil and T. Braun, "Edge Provisioning and Fairness in VPN-Diffserv Networks," *Proc. of the 9th International Conference on Computer Communication and Network (ICCCN 2000)*, Oct. 2000, Las Vegas, USA.
- [55] R. Garg and H. Saran, "Fair Bandwidth Sharing Among Virtual Networks: A Capacity Resizing Approach," *IEEE INFOCOM* March 2000, Tel-Aviv.
- [56] D. Mitra and K. G. Ramakrishnan, "A Case Study of Multiservice, Multipriority Traffic Engineering Design for Data Networks," *IEEE GLOBECOM 99*, pp.1077-1083, Dec. 1999.
- [57] D. Mitra, J. A. Morrison, and K. G. Ramakrishnan, "Optimization and Design of Network Routing using Refined Asymptotic Approximations," *Performance Evaluation*, vol. 36-37, pp.267-288, 1999.
- [58] D. Mitra, J. A. Morrison, and K. G. Ramakrishnan, "Virtual Private Networks: Joint Resource Allocation and Routing Design," *IEEE INFOCOM 99*.
- [59] N. Anerousis, "Dynamic Virtual Private Network Dimensioning in Cost-Sensitive Environments," *IEEE Globecom*, Rio de Janeiro, Brazil, Dec. 1999.
- [60] A. Kumar, R. Rastogi, A. Silberschatz, and B.Yener, "Algorithms for Provisioning Virtual Private Networks in the Hose Model," *SIGCOM'01*, San Diego, California, August 2001.
- [61] Y. Qing, Y. Sivalingam, and B. Li, "Architecture and Analysis for Providing Virtual Private Networks (VPN) with QoS over Optical WDM Networks," *SPIE Optical Networks Magazine*, vol. 2, no. 2, Mar./Apr. 2001, pp.57-65.

- [62] B. Ramamurthy and A. Ramakrishnan, "Design of Virtual Private Networks (VPNs) over Optical Wavelength Division Multiplexed (WDM) networks," *SPIE Optical Networks Magazine*, vol. 3, no. 1, Jan./Feb. 2002, pp.59-66.
- [63] E. Modiano and P. J. Lin, "Traffic Grooming in WDM Networks", *IEEE Communications Magazine*, vol. 39, no. 7 pp.124-129, July 2001.
- [64] M. Maliosz and T. Cinkler, "Methods for Optical VPN Design over Multifiber Wavelength Routing Networks," *Proc. of the Workshop on Optical Network Design and Modelling, IFIP ONDM 2003*, Budapest, Hungary, Feb. 2003.
- [65] B. Megyer, T. Cinkler, and Z. Szombat, "Setting up oVPNs with Traffic Grooming and Protection," *Proc. of the Workshop on Optical Network Design and Modelling, IFIP ONDM 2003*, Budapest, Hungary, Feb.2003, pp 985- 1003.

## Chapter 4

# Optical Burst Switching and Optical Packet Switching

*Steinar Bjørnstad, Christoph M. Gauger, Martin Nord (editors), Elise Baert, Franco Callegati, Davide Careglio, Walter Cerroni, Jan Cheyns, Didier Colle, Piet Demeester, Chris Develder, Dag R. Hjølme, Gabriel Junyent, Mirosław Klinkowski, Marian Kowalewski, Marko Lacković, Marian Marciniak, Mario Pickavet, Carla Raffaelli, Josep Solé-Pareta, Norvald Stol, Erik Van Breusegem, and Paolo Zaffoni*

### 4.1 Optical Burst Switching and Optical Packet Switching Concepts

Most existing wide area telecommunication networks (WAN) have an SDH based, electronically circuit switched transport core. Connection set-up or tear down may require days or weeks and multiplexing/demultiplexing always require complex optical/electro/optical (O/E/O) conversions. Nowadays, the operators and vendors are working on an optical control plane, which should control set-up and tear down of connections. Work on automatically switched optical network (ASON) and generalised multi-protocol label switching (GMPLS) takes place within ITU and IETF, respectively. Resulting optically circuit switched (OCS) networks can offer explicit transfer guarantees, since circuit establishments are confirmed. However, this generates a delay equal at least to the round-trip time, typically several ms. Even though OCS networks will offer more flexibility than today's solution, the access to the optical bandwidth will still be provided with fibre/wavelength granularity.

Future networks should be able to serve a client layer that includes packet-based networks, such as the Internet, which may have a highly dynamic connection pattern with a significant portion of bursty traffic between the communicating pairs. In this case, OCS transport may not be flexible enough. It would require over-dimensioning of the number of connections and of the bandwidth reservation of each connection, to avoid excessive delay and extensive buffering at the ingress router. Here is when Optical Packet Switching (OPS) and Optical Burst Switching (OBS) come into play, with the goal of reducing delays and improving the utilisation of the network's resources through statistical multiplexing. This comes at the expense of not being able to offer explicit transfer guarantees. However, suitable node design and proper dimensioning of network resources may enable support of most services over the same network. OPS and OBS logical performance in relation to contention resolution and quality of service (QoS) differentiation will be discussed in 4.3 and 4.4. Moreover, hybrid schemes are possible where OPS and OBS share the WDM layer with an OCS scheme, serving applications with the need for explicit transfer guarantees.

Europe has been very active in OPS research, especially through projects like RACE ATMOS [1], ACTS KEOPS [2], IST DAVID [3], IST STOLAS [4] and COST 266. OBS was proposed very recently [5], but different research groups in Europe have already made significant contributions to this field, as later sections in this chapter show.

Inevitably, there are some differences in terminology within the research community; we here explicitly describe some concepts and terms used in this chapter. Both *optical packet switching* and *optical burst switching* are based on the idea of separating forwarding from switching in the network nodes. Forwarding decisions are taken by means of a burst control packet (BCP) or packet header that undergoes O/E conversion and electronic processing at the nodes. On the other hand, the burst/packet payload is optically switched, thus avoiding the costly O/E/O conversion, and in principle simplifying the interface cards to the optical fibres. OBS and OPS, as intended here, are hence different from *all-optical* packet- and burst switching approaches, where the headers are processed in the optical domain, thereby controlling a switch. Due to their increased optical complexity, we consider that such concepts are further away from implementation.

Commercial transponders performing O/E/O conversions are optimised for a specific signal bitrate and transmission format. Networks based on optical switching may avoid transponders, which opens up for network *transparency*, here meaning design of a network that readily handles any signal format and

bitrate. This is often cited as an attractive property of optical switching, together with the potential for high bit-rate systems (40 Gbit/s and above). Nevertheless, realisation of such *fully transparent* networks requires a high number of adaptive components, as discussed in [6]. There is hence a trade-off between the flexibility and the complexity/cost in network design. *Semi-transparent* networks, here meaning optical networks optimised for a certain signal format and bitrate may therefore be attractive. Furthermore, such a network, due to the fixed signal format, may allow O/E/O conversion, e.g. to perform 3R regeneration, wavelength conversion and buffering; whenever a function is less costly to perform in the electronic domain. A comparison of optical and electronic buffering is made in section 4.3.

#### 4.1.1. Packet/Burst Handling Schemes

Both in OBS and in OPS the switching matrix is required to be reconfigurable on the burst/packet time scale to allow the packets/burst to efficiently share node and fibre resources. As will be discussed in 4.2, the switching operation is hence more demanding in OPS than in OBS. The basic principles for switching architectures and functionalities are independent of the packet/burst handling scheme [6]. In Figure 4.1 we report the four potential packet/burst handling schemes, classified according to synchronisation and size of the data units.

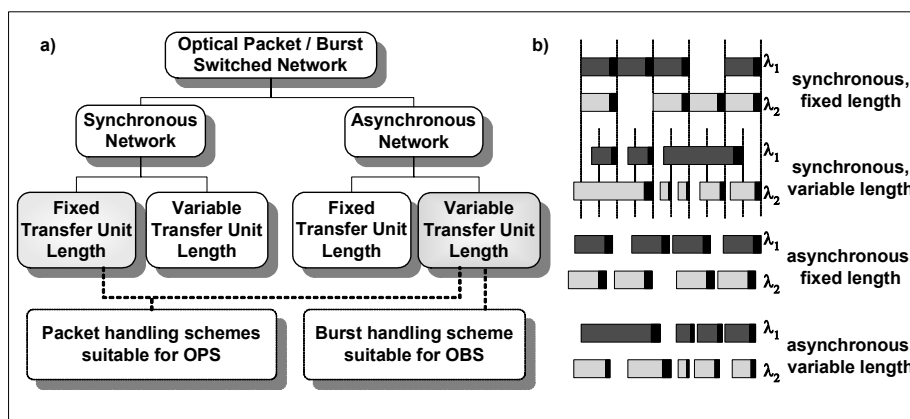


Figure 4.1. - Potential packet/burst handling schemes in OPS and OBS

Figure 4.1 shows that asynchronous and variable length data units are considered more suitable for OBS. This is motivated by decreased complexity at the optical layer, and OBS will also benefit from a large degree of freedom in the burst assembly mechanism. This is the case most widely studied in the literature and to which we refer in the following.

On the other hand, the best scheme for OPS is a matter of debate. The most studied case is that of synchronous, fixed length packet (FLP) [1][7][8], but more recently, work on variable length packets (VLP) have also been studied, with both asynchronous [9][10][11][12][13] and synchronous operation (typically *trains* of packets) [14]. We here discuss aspects relevant for the choice of packet handling scheme in OPS.

Synchronous operation requires optical synchronisers at the switch interfaces and a global network clock for practical realisation. Using FLP requires fragmentation of client packets and padding to fill the optical packet. Both factors increase the overhead, i.e. bandwidth consumption for successful transmission through the OPS network (c.f. 4.2.2), and thus the blocking ratio for the same client load. In addition, fragmentation calls for reassembly of client packets, which increases egress node complexity. Fibre delay line (FDL) based buffer design and management is simpler for FLP than for VLP, and it is simpler to maintain the packet order. The complexity of other contention resolution methods may be independent of the packet length, as discussed in 4.3. In general, a switch matrix operating in synchronous, FLP mode will have less contention than when operating in asynchronous mode [15]. Furthermore, in this mode, a re-arrangeably non-blocking may have equal performance to a strictly non-blocking switch matrix [16].

In most cases, the synchronous, FLP mode or the asynchronous, VLP mode is considered to be the better choice for OPS. This is since the former minimises blocking, enables simpler FDL based buffers and only requires re-arrangeably non-blocking switch matrices. The latter has less overhead and avoids the need for a network global clock, optical synchronisation and reduces edge node assembly time. However, at a given technology status, and in a given network context (client layer characteristics and service requirements), a performance and cost evaluation is needed to answer which packet handling scheme is more suitable.

#### **4.1.2. OPS/OBS Properties and Comparison**

We consider OPS and OBS for application in a mesh-based WAN or “core network”, context. Mesh networks minimise switch matrix sizes and propagation distance, at the same time enabling flexible load balancing and link-protection, whilst avoiding single-points of failure. Ring based OPS for metro area networks (MAN) are considered in 4.5. Considering transmission systems, high-capacity WDM systems are considered, and the interplay between channel count and network performance is discussed in section 4.3.



### Node functionalities

The network nodes combine edge router and core router functionalities (Figure 4.2). Edge (ingress/egress) routers modules are capable of communicating with the electrical layer, and may also form the interfaces to adjacent OPS/OBS layers. The main functionalities of an ingress router are to aggregate electrical packets into optical packets/bursts as defined by the assembly algorithm. Note that the aggregation process at an ingress node will shape the traffic flowing into the core, which may be beneficial to network performance. To some degree, aggregation may reduce the degree of self-similarity [17][18], although not at larger time-scales [19]. Its routing table and scheduling algorithm determine what control information to encode and when to send the optical packets/bursts. The egress router reads the optically encoded control information, defragments arriving packets/bursts, and forwards the client packets, after reassembly if required, to the client layer.

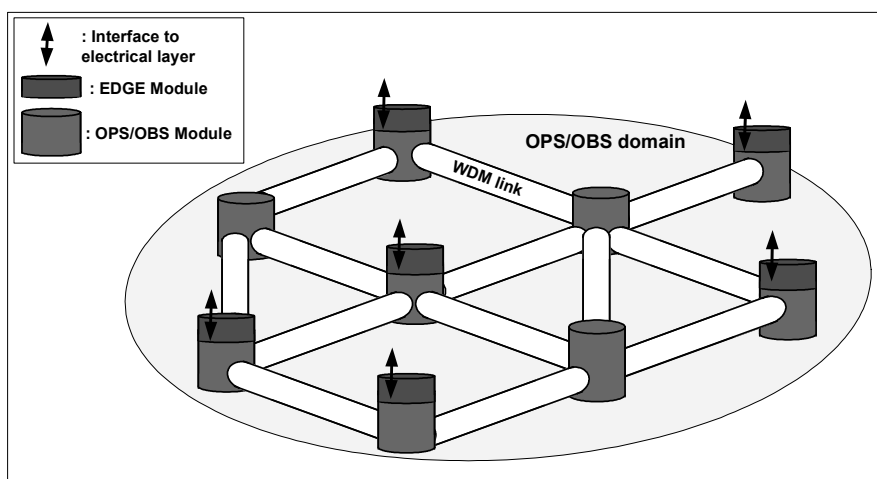


Figure 4.2. - Network consisting of nodes with edge and core routers functionalities

A core router module reads the control information after O/E conversion. It then processes this information electronically to configure the switch matrix and contention resolution resources. When required, control information should be re-written, otherwise the old header continues with the packet.

### Comparison of concepts

OPS and OBS concepts were compared in [6][13][20]; these studies identified some fundamental differences, briefly discussed here. Differences in node realisations, contention resolution and QoS differentiation are discussed in the next chapters.

Most OPS proposals assume packet durations in the 0.1-1  $\mu$ s range. We consider client packets representative of today's Internet traffic, which is in the 40 to 1500 bytes length [21]. In that case, OPS typically requires a very low degree of client layer packet assembly, at a bitrate of 10 Gbit/s. However, the level of aggregation needed for the same packet duration, increases with the bitrate. The payload length and relation to packet format has an impact on the node design, as discussed in section 4.2. Furthermore, control information is typically encoded "in-band". Here, in-band means that the header is transmitted on the same wavelength and either overlaps the payload in time (as for sub-carrier modulation), or it is transmitted just ahead of the payload (serial header).

OBS assumes more extensive burst aggregation, to realise bursts with payloads typically carrying tens of kbytes. OBS uses out-of-band encoding of control information, which is realised by sending BCPs on a common BCP wavelength on that link. Each BCP is transmitted ahead of its corresponding burst, and contains the timing information on the burst. The BCP information and the time offset, in combination with delayed reservation (DR) principle, enable advanced burst scheduling. This can optimise bandwidth usage and enable QoS differentiation methods as discussed in 4.3.

#### **4.1.3. Summary**

OPS and OBS are optical transport networks architectures that have a finer granularity than OCS and are expected to better support dynamic traffic patterns. One main advantage is the improved utilisation of network resources that can be expected. OBS assumes asynchronous operation with variable burst length, whilst OPS typically operate in either synchronous FLP mode or in asynchronous VLP mode. Different packet and burst granularities, control encoding and appropriate scheduling mechanisms are hence the main differentiators of OPS and OBS concepts. We develop these aspects in later sections, and discuss their impact when it comes to node design (section 4.2), contention resolution (section 4.3) and QoS differentiation mechanisms (section 4.4). Application of the OPS concepts to a metro environment is discussed in section 4.5. In the last section (4.6), we draw a conclusion based on this study.

## **4.2 Design and Analysis of OPS Nodes**

This section describes the main functionalities of OPS and OBS nodes and makes some design considerations. A brief overview of technology status is

given before presenting OPS node architectures proposed within COST266. In general, the level of details in work on node design varies greatly. Some work is closer to implementation and considers e.g. realisation of clock recovery and control unit, whilst other focus on packet/burst switch architectures to ensure correct forwarding, without details of realisation.

#### 4.2.1. OPS and OBS Nodes

The OPS/OBS node design space is broad. However, in the general case, an OPS or OBS node consists of four main building blocks. These are illustrated in Figure 4.3, through the example of an OPS node in slotted operation.

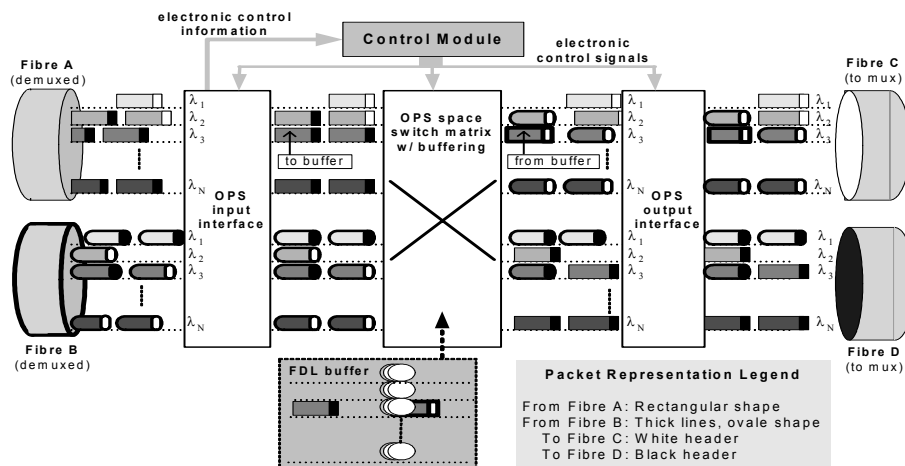


Figure 4.3. - Generic OPS node with FDL buffers in slotted operation. Packets may be followed as they transit through the switch

The input interface detects a preamble/synchronisation pattern, marking the packet arrival and enabling clock recovery. It monitors incoming signals and conditions the incoming packets as required, e.g. through power equalisation and alignment of packets in slotted operation. Furthermore, it retrieves control information, encoded in the packet headers or in the BCPs, and transmits it to the control unit in electronic form. To accommodate processing delay, packets may be delayed using fixed length FDLs. The electronic control unit makes look-ups in the routing table, and is responsible for implementing the scheduling policy. It considers the control information and current resource configurations to identify a suitable output port/wavelength, or buffer allocation. If needed, the control unit identifies new control information and updates timing information (in OBS). The switch matrix, buffers and interfaces respond to the electronic reconfiguration signals from the control unit. The switch matrix

influences the node performance by its switching time, maximum throughput, internal blocking properties and signal degradation. The output interface updates the headers or BCPs and conditions the signal by e.g. 3R regeneration, if required.

#### 4.2.2. General Design Considerations

The technology required to realise OPS/OBS nodes is currently not mature. Therefore cost considerations are hard to make at present; developments within technology are described in 4.2.3. We will in this paragraph focus on relevant performance parameters and bandwidth efficiency.

The OPS/OBS network should be designed to offer its client layers throughput in the Tbit/s range to meet expected increases in traffic. The throughput of an internally strictly non-blocking switch matrix is the capacity switched per time unit in the ideal case, when input and output channels are fully loaded. In realistic network scenarios, however, external blocking or *contention* occurs, leading to loss of packets/bursts, which depends on the packet/burst switch architecture/designs, including contention resolution resources. Typically, simulations are used to study different designs' logical performance as a function of node degree, load and traffic characteristics. In addition to acceptable loss rates, the packet/burst switch should support QoS differentiation, as discussed in section 4.4. The delay of an OPS/OBS node is in most cases negligible compared to the transmission delay in the network. However, if jitter caused e.g. by buffering is significant, it may lead to reordering of packets/bursts, which increases egress node complexity.

In a packet/burst switched network paradigm, each network layer encapsulates higher layer packets, thereby adding an overhead. In the OPS/OBS layer, successful processing and switching typically dictates packet fields for control information, synchronisation pattern(s) and optical guard bands (OGBs). OPS/OBS networks should be designed to handle a certain load offered by the client layer, but the overhead creates a need for the OPS/OBS network to actually be designed for a higher *optical* load.

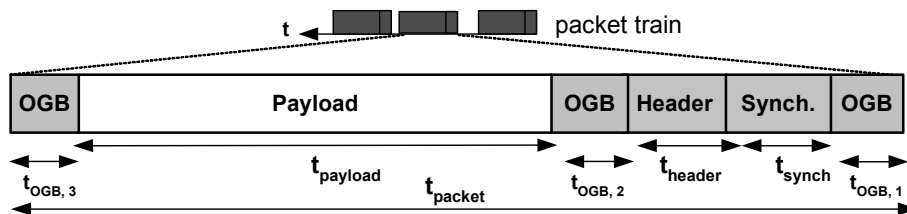


Figure 4.4. - Example of optical packet format

Considering a single node, the overhead, defined in equation 4.1, describes what proportion of time the switch matrix spends settling the switch and transmitting non-payload fields relative to the payload duration. The durations of these fields are technology dependent. E.g. the header duration depends on the header encoding method, and the packet synchronisation/preamble field must contain a pattern long enough to allow a stable clock-recovery with unambiguous start-of-packet detection. OGBs are required to accommodate jitter in e.g. the header insertion process and between packets in a packet train. In addition comes the switching time, during which the considered switch matrix path cannot be exploited.

$$Overhead_{OPS} = \frac{t_{packet} + t_{switch} - t_{payload}}{t_{payload}} = \frac{(t_{header} + t_{OGB} + t_{synch} + t_{switch})}{t_{payload}} \quad (4.1)$$

Even with 1 ns switching time, it is very hard to limit overhead to e.g. 10 % at 10 Gbit/s payload bitrate, when payload length is in the lower region of typical distribution of IP packet length. The interplay between packet format and switching time on overhead is further discussed in [22]. To reach this threshold typically requires some aggregation at the OPS ingress nodes, to minimise the fraction of optical packets with payloads in the 40-200 bytes range. For payloads around 1500 bytes, switching times around 100 ns can in general be accepted.

In OBS, each burst on a data wavelength needs a fraction of the control wavelength bandwidth to transmit its BCPs, similar to the header field in OPS. OGBs are still needed to accommodate finite switching times, but the impact decreases due to increased payload durations.

#### 4.2.3. Optical Technology Status

Most optical packet/burst switch designs require advanced technology for realisation. Among the essential challenges are control unit processing and optical switching technologies, which are briefly discussed below. Detailed studies of all enabling technologies have not been the focus of the COST 266 action, and are beyond the scope of this document, overview of key technology realisation can e.g. be found in [2].

The control unit should maintain an accurate routing table. It can typically be updated on a relatively slow time-scale, based on control plane information on topology changes and load balancing factors. On the other hand, electronic processing of packet headers must be performed for a high number of packets simultaneously. The operation typically includes table look-up to identify

output port, scheduling of output wavelength or buffer input in agreement with packets' service class. As a maximum bound, the sum of processing time and switch fabric reconfiguration time should not exceed the packet duration. Otherwise a data flow bottleneck at the switch input occurs. Hence, electronic processing may become a bottleneck in optical networks. Increasing packet duration gives more time for processing, so that OBS eases the requirement on total processing time. However, since the scheduling is expected to be more advanced, e.g. to implement a just enough time (JET) based QoS differentiation scheme; there will still be stringent requirements on the speed [23] and/or processing parallelisation degree of the control unit.

Switch fabrics for OPS and OBS can be categorised as *space switches*, *broadcast-and-select* and *wavelength routers*. This review considers maximum switching times for OPS of 100 ns, and is based on [22]. OBS may accept larger switching times, and may have more alternatives for switching fabrics. However, considering burst durations up to 100  $\mu$ s, the switching time should not exceed 10  $\mu$ s, excluding e.g. the use of the relatively mature micro electro-machine systems (MEMS) switches. All architectures considered are strictly non-blocking, unless otherwise stated.

$N \times N$  *space switches* are configured by setting basic solid-state optical switches according to the switch architecture. *Crossbar* switches are based on  $2 \times 2$  switches, and suffer from poor cross-talk properties. They are wide-sense non-blocking, thus an intelligent switch path selection can completely avoid internal blocking. *Router-Selector* switches use  $1 \times 2$  switches and exploit this architecture's excellent crosstalk properties. Crossovers and bends in the interconnection shuffle may however give differential loss and cause crosstalk. In these architectures the number of basic switches scales as  $N^2$  and  $2N^2$ , respectively. Reported matrices with less than 100 ns switching time are rare, and these matrices are limited to port counts of  $16 \times 16$ .

The standard *Broadcast-and-Select* switch is based on passive  $1/(N \cdot W)$  splitting of the  $N$  WDM input signals, each with  $W$  wavelengths, followed by active selection at each of the  $N \cdot W$  outputs. A high splitting ratio gives SNR reduction, and the loss is compensated for by EDFAs which further decrease signal quality by adding the ASE noise. The total number of SOA gates needed is  $W^2 \cdot N + W \cdot N^2$ . For a given size  $(N \cdot W)$ , SOA count is minimised when  $N = W$ . An integrated board compatible with a  $16 \times 16$  space switch with  $W = 16$  (throughput of 2.56 Tbit/s) has been reported [24]. 512 of the total 8192 SOAs would be in the on-state at any moment, and power consumption becomes an issue. Using sub-equipped versions of this board, an OPS experiment has demonstrated a throughput of 640 Gbit/s in asynchronous operation

[25]. An adaptation proposed in the context of OBS is called Tune-and-Select and has been shown to scale to an effective throughput in the range of Tbit/s [26].

*Wavelength-routers* are based on a passive fabric, usually array waveguide gratings (AWGs), with preconfigured input-output paths depending on the input wavelength. Reported in [27] is e.g. a demonstration of use of such a 42×42 AWG for packet switching. In general, spectral filtering and channel cross-talk may cause signal degradation, but AWGs have a very good potential for use in OPS switches.

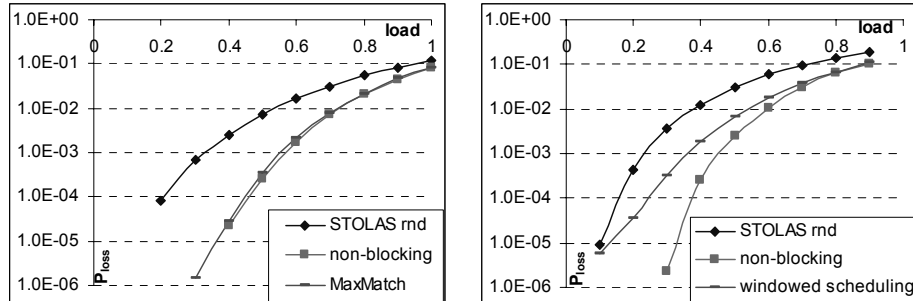
#### **4.2.4. Packet/Burst Switch Architectures**

In general, except for the control information encoding and control unit scheduling, OPS node architectures for asynchronous operation will also be suitable for OBS.

##### *Example of a blocking OPS architecture*

A very basic node design consists of having the demultiplexed WDM signals connected to an AWG via tuneable wavelength converters (TWCs). This is very simple switch architecture, however, it has the drawback that it is blocking, since no wavelength conversion is performed at the output of the switch, the converters at the input can only use the same wavelength set as in the input fibre. This has as a consequence that this switch architecture is internally blocking.

Important in such a switch design is how the output ports of the AWG are combined into the output fibres [28]. If this is done properly and intelligent choices on the wavelength conversion are made at the input, the performance of this blocking node can approximate the performance of a non-blocking node. In asynchronous mode, however, it is a lot harder to emulate a non-blocking node using this architecture. The problem is that once a decision is taken for a TWC this cannot be reverted, but when a future packet arrives it might become clear that another choice would have been better. Thus a possible way of improving performance is using a windowed scheduling mechanism [29]. This scheduling in the switch increases the time separation of header and payload by an extra FDL at the input. In this way there is a form of prediction of which packets will block each other, so that the converter decision for these overlapping packets can be coupled, which will result in a lower blocking probability. We show simulation results for this switch design in Figure 4.5.



**Figure 4.5. - Packet loss simulations for a blocking optical packet switch (3 fibres in and out, 15 wavelengths per fibre) in slotted (left) and asynchronous mode (right). A non-blocking node is shown as reference, STOLAS rnd is performance using no specific TWC assignment algorithm. In slotted operation, MaxMatch TWC assignments allow to emulate a non-blocking node. In asynchronous mode windowed scheduling improves the performance, but there still exists a serious gap with the ideal non-blocking node.**

#### *Example of non-blocking OPS architecture*

We here describe a switch design, suitable for asynchronous packet switching, shown in Figure 4.6, proposed in [11]. The input WDM signal of each fibre is demultiplexed to its corresponding wavelengths and fed to the input of the TWCs. The outputs of each TWC are then fed to the AWG inputs. By tuning the TWCs output wavelength, packets can be sent to either of the AWG outputs. The packet will be sent directly to the scheduled output, if vacant. If no output with correct destination is available, the packet will be sent to one of the buffer inputs, if a vacant buffer input can be found. If not, the packet will be dropped.

Buffered packets are clocked out of the buffer and sent back to an AWG input as soon as a wavelength output to the destination becomes available. At the buffer output, the wavelength, and thus AWG output, is set by tuning a tunable laser. This type of architecture is called a feedback design, and has the benefit of supporting packet priority, also when FDLs are used for buffering [30], [31]. When a packet is leaving the AWG for the output, the signal is converted to the desired wavelength before it is multiplexed onto the correct output fibre.

The design is suitable for the given COST scenario described in Chapter 1 of this report, where node degree is set to typically a maximum of 5. However, a drawback with this design is that it does not scale well to a high node degree. The total number of switch inputs  $n$ , is given as  $n = N \cdot w$ , where  $N$  is the number of input/output fibres and  $w$  is the number of wavelengths in each fibre. The total number of channels needed in the AWG is given as  $n+k$ , where  $k$  is the number of buffer inputs. An AWG with size  $(n+k) \times (n+k)$  is therefore



required. Since  $n$  increases both with the number of fibres and the number of WDM link wavelengths, the maximum switch size is limited by the size of the AWG, which is currently reported to be a maximum of 400 channels [32]. However, a scalable design based on the same principles, scaling to a very high number of wavelengths, and a high node degree, can be found in [11].

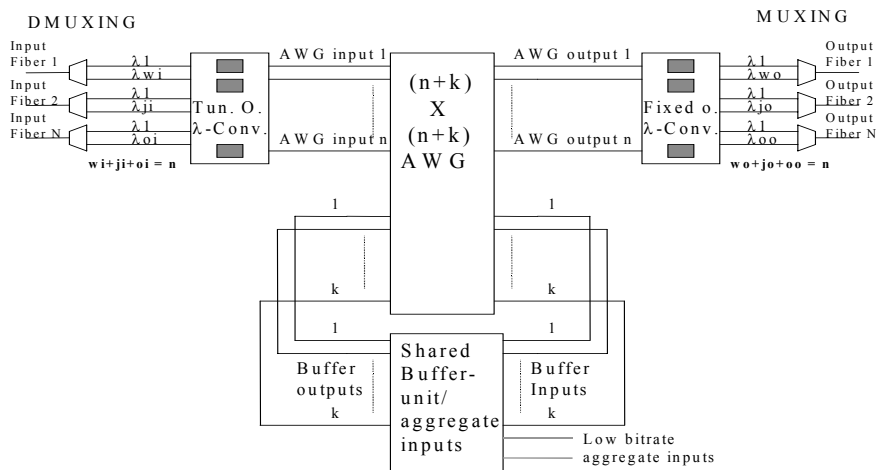


Figure 4.6. - An optical packet switch with shared buffers and low bit rate aggregate inputs

4.2.5. Summary

This section introduced the topic of node design in OPS/OBS networks, and pointed out important design factors. Currently, node design for OPS/OBS is focused on design principles and logical performance. For implementation in networks, the goal is to optimise the network’s performance/cost ratio. Hence, one should also consider impact from client layer requirements, technology status and the existing WDM layer’s topology and transmission system. OPS/OBS node design requires several performance/complexity trade-offs that will be further studied in relation to contention resolution and QoS differentiation, in sections 4.3 and 4.4, respectively.

4.3 Contention Resolution

4.3.1. Motivation and Overview

Optical burst and packet switching inherently rely on statistical multiplexing in order to achieve good utilisation in presence of bursty traffic. As a conse-

quence, temporary overload situations called contention situations occur and have to be resolved. A reservation or transmission conflict, which leads to burst or packet loss, exists if the wavelength on the designated output fibre is blocked by a different burst or packet. Such a contention situation can be resolved in one or several of the following three domains:

- **Wavelength domain:** By means of wavelength conversion, a burst or packet can be transmitted on a different wavelength channel of the designated output fibre. Thus, all wavelength channels of an output fibre can be considered a single shared bundle of channels.
- **Time domain:** A burst or packet can be delayed until the contention situation is resolved by applying a buffer. In OPS or OBS, either simple FDL buffers or electronic buffers can be used. While FDL buffers only provide a set of fixed delays an electronic buffer can provide virtually unlimited delay and random access. Complexity issues of FDL and electronic buffers will be discussed in the next subsection.
- **Space domain:** In deflection routing, a burst or packet is sent to a different output fibre of the node and consequently on a different route towards its destination node. As contention is not resolved locally in a single node but by rerouting over-load traffic to neighbouring nodes, this scheme depends heavily on network topology and routing strategy. Deflection routing results in only limited improvement for variable length bursts or packets [33] and has not been investigated within this COST Action. Furthermore, as a consequence, packets can arrive out of order at the egress node.

Space domain can be exploited differently in case of multi-fibre networks, i.e. several fibres are attached to an output interface. In this case, a burst can also be transmitted on a different fibre of the designated output interface without wavelength conversion.

In the following sections, contention resolution in wavelength and time domain is discussed and results of a joint comparative performance evaluation are presented.

#### ***4.3.2. Wavelength Conversion***

WDM not only provides increased transmission capacity but also allows for highly effective contention resolution. If wavelength converters are employed, all wavelengths on a fibre (or within a certain waveband) can be considered a bundle of channels shared by all bursts or packets to be transmitted over this fibre (waveband). In teletraffic theory, it is well known that a bundle of  $n$  parallel servers each with capacity  $c$  has a smaller blocking probability and thus a

higher utilization than a single server of capacity  $n \cdot c$ . This is called economies of scale.

Therefore, most node designs for optical burst and packet switching assume full wavelength conversion, i.e. every incoming or outgoing wavelength is equipped with a wavelength converter. However, as wavelength converters are technologically complex and rather expensive, shared converter pool concepts have been proposed and investigated [34][35]. In this case, the ratio of the number of wavelength converters in the pool and the number of wavelength converters in case of full conversion is referred to as the conversion ratio  $r_c$ .

#### 4.3.3. FDL Buffer Architectures

Since traditional queuing is not feasible in all-optical burst or packet switches, contention resolution in the time domain may be provided by using Fibre Delay Lines (FDLs), which imitate conventional queuing by delaying packets that are forced to go through an optical fibre of a given length. In literature different kinds of FDL buffer architectures (either in a single or multistage configuration) have been proposed [36], which may be basically classified into

- *feed-forward* buffer, where a packet coming out of the buffer goes directly to one of the output ports of the switch;
- *feedback* buffer, where a packet coming out of the buffer either goes to the output or re-enters the delay lines.

Since it is here assumed that the output wavelength is reserved at packet arrival, both the above buffer configurations are equivalent to an output queue.

On the other hand, in a DWDM network contention resolution may also exploit the wavelength domain, by sharing the wavelength pool of a fibre and then by transmitting contending packets on different wavelengths.

As a consequence, when a packet needs to be forwarded to an output fibre specified in the routing table, the Wavelength and Delay Selection problem (WDS) arises. In fact, these two actions are somewhat correlated, because the need to delay a packet is related to the availability of the wavelength selected. The WDS problem becomes also more complex in case of asynchronous, variable-length optical packets, since some gaps may appear between queued packets inside the FDL buffer due to the discrete number of available delays [9]. In order to solve the WDS problem under these traffic assumptions, a few resource allocation policies have been proposed [37][10]. Here we consider the following ones:

- *Random Non-Full queue* (RNF): the wavelength is chosen randomly excluding those that will be busy beyond the maximum available delay (full queues).
- *MINimum Length queue* (MINL): the wavelength that will be free as soon as possible (the shortest queue) is chosen.
- *MINimum Gap queue* (MING): the choice this time falls on the wavelength that introduces the smallest gap between the current packet and the last queued one.

In case all queues are full, no choice is made and the packet is lost.

In optical burst switching, JET reservation scheme offers the flexibility to reserve newly arriving bursts in gaps left by already reserved bursts. Thus, JET provides another solution for the problem of gaps induced by FDL buffers [38]. Also, regarding the WDS, the sequence in which wavelength conversion and buffering are applied can be exploited to trade off wavelength converter and FDL buffer usage [39].

#### **4.3.4. Electronic Buffering**

Optical memory with random access in the time-domain is known to still be immature, and FDLs are therefore used as optical buffers [30]. Since FDLs give fixed delays, random access is not possible, i.e., in contrast to electronic memory, FDLs cannot provide access to a specific data packet at an arbitrary access time.

First in first out (FIFO) operation of the buffer is desirable in order to avoid reordering of packets on their way to the destination. Although this can be achieved using FDLs, storing variable length packets brings up the need for buffering all, or most of the, packets using many different delays and thereby a high number of FDLs. As an alternative to FDLs, the use of simple electronic FIFO memory with few opto-electronic interfaces is suggested in [11]. When using electronic memory, fast random access with respect to time in a FIFO buffer can be obtained. A random access in space to a random storage unit is more complicated since addressing the storage unit before readout of the data is then needed. However, since FIFO buffering is used, access to a random storage unit in the buffer is not required.

Like when using FDLs, data-format transparency is obtainable in electronic memory, however bit rate transparency is more complicated since clock recovery circuits recognizing the bit rates is then needed.

### 4.3.5. Comparative Performance Evaluation

In a common evaluation scenario, the impact of different WDS algorithms and individual FDL delays in an FDL buffer is compared for OBS and OPS both assuming asynchronous operation and variable length bursts or packets. Then, both approaches are compared to OPS with electronic buffers based on the number of buffer interfaces.

Bursts and packets arrive at a node with 4 input and output fibres according to a Poisson process with an offered load of 0.8. Burst and packet length is negative exponentially distributed with mean 100 kbit (bursts) and 4 kbit (packets) which translates into an average transmission time  $h = 10 \mu\text{s}$  (bursts) or  $h = 0.4 \mu\text{s}$  (packets) for a 10 Gbit/s line-rate. Unless stated differently, 16 wavelengths are assumed on each fibre and FDL. For the FDL buffer, the length of FDL  $i$  can be calculated as  $i \cdot b$  with respect to a basic delay  $b$ . For OBS, JET is applied for fibre and FDL reservation. FDL reservation is performed at time of burst arrival (*PreRes* in [38]).

Figure 4.7 compares the three WDS policies for OPS with 8 FDLs in the buffer (circles) and OBS with 4, 6 and 8 FDLs in the buffer (triangles). Burst/packet loss probability is plotted as a function of the basic buffer delay  $b$  normalized to the average packet length.

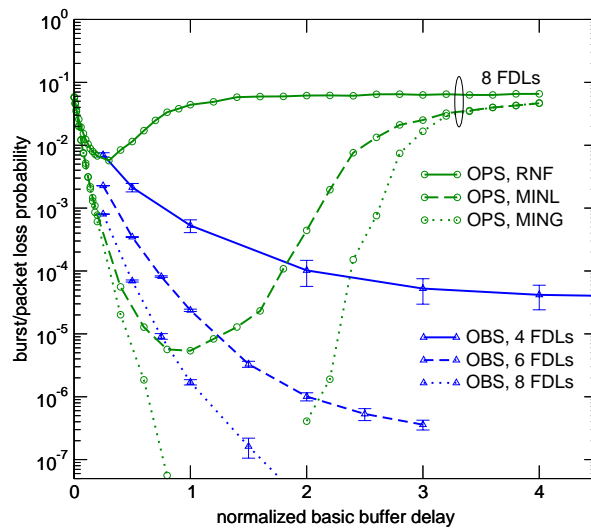
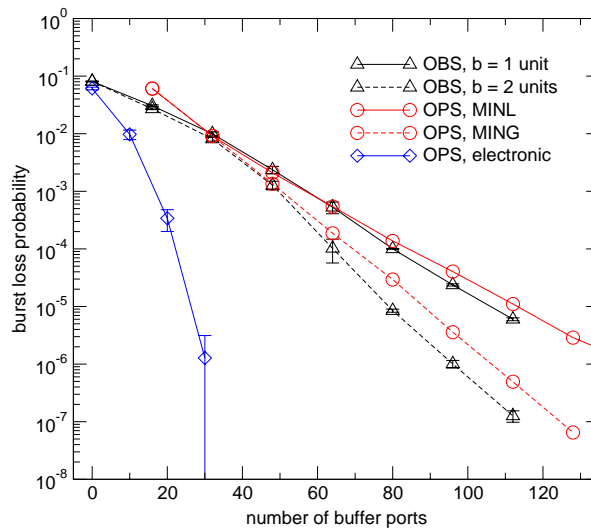


Figure 4.7. - Burst/Package loss probability versus normalized buffer delay

The RNF choice gives the worst performance since such policy does not detect the wavelengths immediately available, which do not insert gaps in the buffer. More intelligent mechanisms, such as MINL and MING, provide a strong

improvement. In particular MING outperforms MINL because it aims, first of all, at reducing the gaps, leading to a more efficient buffer utilization and therefore to shorter queues overall and better performance. All three curves clearly show a minimum of the packet loss rate as a function of the delay unit, with the position of such minimum depending on the wavelength selection policy adopted. These results demonstrate that a smart WDS policy together with an accurate dimensioning of the buffer parameters allow to achieve very good performance with only a limited number of FDLs.

For OBS, increasing the delay significantly reduces loss probability for different number of FDLs in the buffer. In contrast to OPS and due to the more flexible but also more complex JET reservation mechanism, no minimum appears but a saturation effect can be observed from a basic delay of 2-3 times the mean burst length on.



**Figure 4.8. - Burst/Package loss probability versus number of buffer ports**

Figure 4.8 depicts the impact of the number of buffer ports for FDL and electronic buffers. For FDL buffers, the number of buffer ports is the product of number of wavelengths and number of FDLs in the buffer. Basic delay of FDL buffers is chosen to be the optimal one in case of OPS (Figure 4.7), and one and two times the mean burst transmission time in case of OBS, respectively. It can be seen that increasing the number of buffer ports greatly reduces blocking for all scenarios. While all results for FDL buffers show comparable trends, electronic buffers need a significantly smaller number of buffer inter-

faces. However, as electronic buffers need O/E/O interfaces the number of buffer ports has to be minimized. For a node with 8 input and output fibres and different numbers of wavelengths, Figure 4.9 shows that a decreasing number of buffer ports are needed for an increasing number of link-wavelengths.

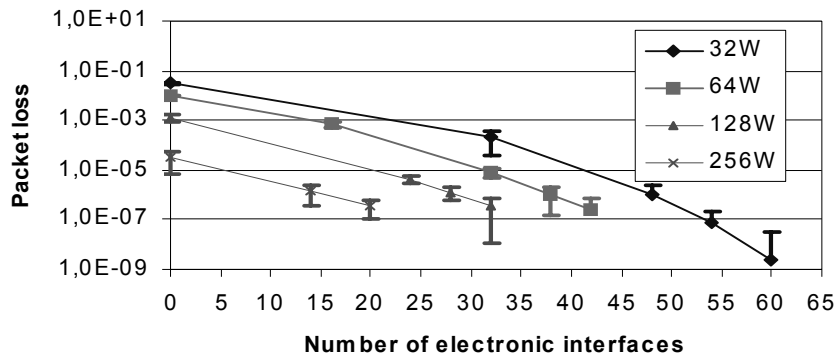


Figure 4.9. - Packet loss when using electronic buffering, 32, 64, 128 or 256 wavelengths (W)

#### 4.3.6. Summary

In this section a performance evaluation based on simulations from several partners in COST 266 has been presented. Different approaches for OPS/OBS contention resolution in time, assuming contention resolution also in the wavelength domain is shown. We conclude that both FDL and electronic buffers can significantly reduce loss probability when used together with wavelength conversion. The comparison shows that electronic buffers need fewer but potentially more expensive O/E/O interfaces to reach the same loss rates. For all the compared buffering schemes delay is negligible compared to typical end-to-end delays.

### 4.4 Quality of Service in OBS/OPS

#### 4.4.1. Introduction

Introduction of multimedia applications in the Internet, which may have strict real time and information loss demands like a high quality video component, have increased the need for service quality differentiation. We expect IP to be the converging protocol layer, but the IP protocol itself does not support QoS differentiation. When implementing an OBS or an OPS layer, the quality of the service offered will be influenced by the amount of resources, like buffering and wavelength converters, spent in the network nodes. An OPS layer

should therefore be able to support QoS differentiation, preventing over-dimensioning of the network nodes and delivering QoS differentiation to the IP-layer.

ITU-T Recommendation Y.1541 [40] defines some provisional IP network QoS class definitions and end-to-end network performance objectives. The highest demand of any class with respect to IP packet transfer delay is 100 ms, with an allowed delay variation of 50 ms. As shown in the analysis in this report's "contention resolution" chapter, delay in OPS and OBS networks is negligible. The lowest specified IP packet loss ratio for any class is  $10^{-3}$ . One should however notice that this value is for an end-to-end IP relation. When specified for real-time services, it is based solely on observation of high quality voice applications and voice codecs (see e.g. [41] for a laboratory study of four VoIP gateways demonstrating that even an IP packet loss ratio of 0.1 may be acceptable). It is however made very clear that some of the values in Y.1541 are too relaxed. Appendix IX (Informative) states: "The Classes in Table1/Y.1541 are intended to cover a broad range of applications for which the transport requirements are known. Examples of applications not covered by these classes are broadcast TV distribution, program audio, Digital Cinema, and compressed HDTV transport, where very low loss may be needed and possibly low network delay". It is also clear from other sources that some applications based on MPEG2 video coding can not tolerate IP packet loss ratios above  $10^{-5}$ , e.g. in [42] a packet loss ratio of  $10^{-6}$  is used for the highest priority QoS class. We therefore expect that a packet loss through an OBS/OPS node better than  $10^{-6}$  should be sufficient to service even the most demanding video-services.

#### ***4.4.2. Quality of Service in Optical Burst Switching***

In order to provide service differentiation directly in the optical layer several approaches have been proposed and investigated for optical burst switching. They take advantage of burst reservation, burst assembly or a combination of both and can be classified based on their key mechanism as follows [6].

- Differentiating offset values,
- Preemption (composite burst switching),
- Intentional dropping of (low priority) bursts,
- (Re-)scheduling in core nodes, and
- Access control and bandwidth reservation.



Offset-based schemes rely on the concept of delayed reservation, i.e. the burst control packet is separated from the data burst by an offset time. In the JET reservation scheme the exact start and end times of burst transmission are considered for reservation. Due to this detailed reservation information, bursts can be reserved in between two already reserved bursts.

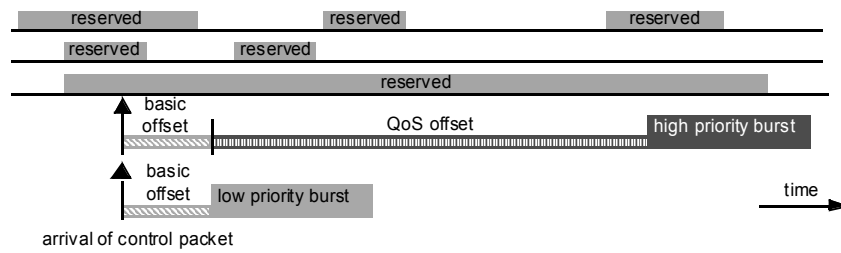


Figure 4.10. - Reservation scenario for bursts of different classes

Service differentiation is achieved by allowing early reservation of high priority bursts by assigning an extra offset time [43] - called *QoS offset*. Therefore, high priority bursts make their reservation in a rather lightly loaded system and have a smaller loss probability while low priority bursts experience the total system load and have a higher loss probability. Figure 4.10 illustrates this effect for three wavelength channels on which some bursts are already reserved, and high priority and low priority bursts with different QoS offsets arrive at the same time.

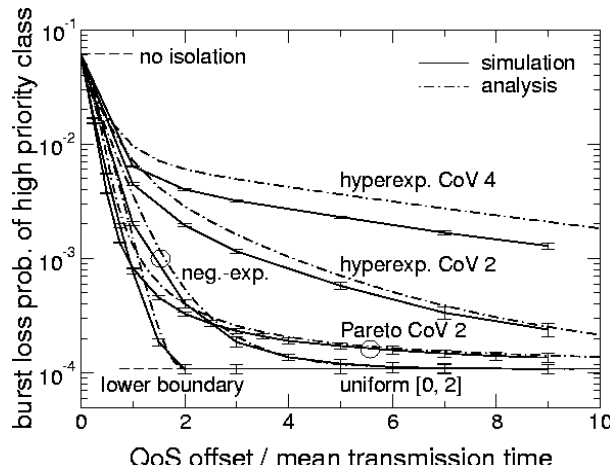


Figure 4.11. - Impact of QoS offset on burst loss probability of high priority bursts

The impact of QoS offsets on differentiation in loss probability has been analysed in [43] and [44]. Figure 4.11 depicts the impact of the QoS offset on the

burst loss probability of the high priority class. As the mean and the distribution of low priority bursts have significant impact, the QoS offset is normalized by its QoS mean burst length, and different burst length distributions of the low priority class are included.

Offset-based QoS has the same total blocking probability with or without service differentiation, i.e. the overall performance is not reduced significantly ([43] reported a slight increase in overall loss rate for low loads). However, this scheme has two severe drawbacks [44]: First, the loss probabilities of high and low priority classes are highly dependent on burst characteristics. Second, basic offset adaptation in core nodes can change the differentiation in offset, which leads to undesirable subclasses as they introduce unfairness [45].

#### 4.4.3. *Quality of Service in Optical Packet Switching*

##### *Fibre Delay Line buffering*

It is important to underline that the QoS management techniques in optical packet switches must be kept very simple due to the delay-oriented nature of FDL buffers. In particular, it is not possible to change the order of packets coming out of the delay lines, thus making pre-emption based techniques not applicable. Therefore, mechanisms based on a-priori access control of packets to the WDM buffers are necessary [46]. The intent here is to improve the WDS policies mentioned in 4.3.3 in order to differentiate the QoS by allowing different degrees of choice to different policies. The objective is to apply some form of reservation of the resources managed by the WDS policies, i.e. the available wavelength and delay, in order to privilege one traffic class over the other. The following alternatives have been investigated:

- *Time-threshold-based technique*: the resource reservation is applied to the time domain, and a delay threshold  $T_{low}$  lower than the maximum available delay is defined. The WDS policy for low-priority packets cannot choose delays that are above threshold; therefore, a low-priority packet cannot be accepted if the current buffer occupancy is greater than  $T_{low}$ , leaving the remaining buffer space to high-priority packets which see the whole buffer capacity. This causes packets belonging to different classes to suffer different loss rates.
- *Wavelength-based technique*: the resource reservation is applied to the wavelength domain. The WDS algorithm for high-priority packets can send packets to any wavelengths of a fibre, while low-priority packets are allowed to access only a subset ( $w_{low}$ ) of the wavelength resources, which in any case is shared with high-priority packets. Low-priority packets are

expected to suffer higher congestion and typically to experience higher delays than high-priority ones.

These two concepts have been applied to the MING WDS policy discussed in section 4.3.3, leading to two new QoS-oriented policies named MING-D and MING-LIM which use the time-threshold-based and the wavelength-based technique respectively. As an example, in Figure 4.12 the performance of the MING-D policy is shown for the same node configuration as in section 4.3.3, providing a good separation between the high-priority class (grey curves) and the low-priority class (black curves).

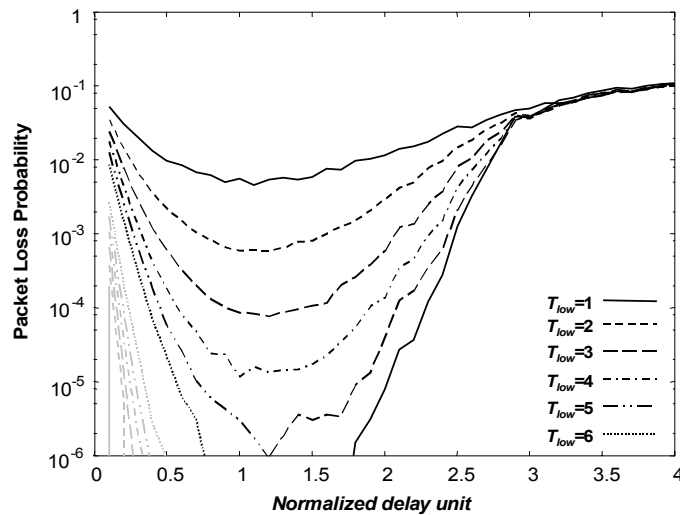
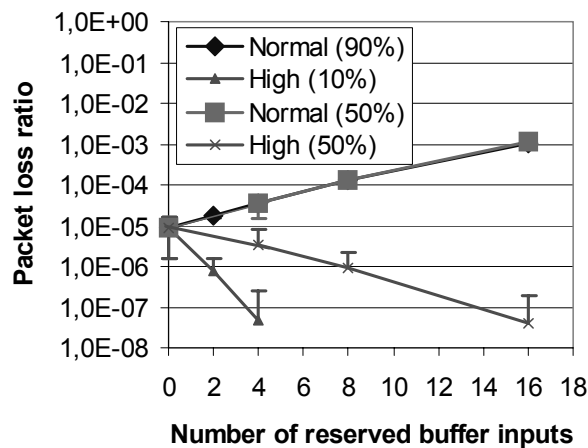


Figure 4.12. - Service differentiation with the MING-D policy

#### *Electronic buffering*

Using the OPS switching architecture described in 4.2.4, packet loss ratio differentiation dependent on the service classes is achieved by reserving parts of the buffer inputs for specified service classes. As argued in [47], we expect two service classes to be sufficient for service differentiation in an optical packet switched network. Therefore we have chosen to evaluate the packet loss when the buffer resources are divided into two different blocks of inputs, allowing two service classes. If the packet belongs to the High Class Transport bearer service (HCT), any available buffer input can be used. If the packet belongs to the Normal Class Transport bearer service (NCT), only a limited number of buffer inputs can be used, if one of them is available. If no buffer input is available, the packet will be dropped.

To evaluate the described principle, we have done simulations quantifying to which extent buffer inputs should be reserved when two traffic classes are assumed. The share of traffic belonging to the HCT class will be set to 10 and 50 % respectively, and the number of wavelengths in the links to 32.



**Figure 4.13. - Packet loss as a function of number of buffer inputs reserved for HCT packets, 32 available wavelengths, load 0.8, 8 input/output fibres, variable length Poisson packet arrival. A total of 42 buffer inputs are available. Curves are given for 10 % HCT (High) and 90 % NCT (Low) traffic and the two traffic classes having equal share of traffic (50 %). The two curves for the NCT traffic are coinciding and are therefore seen as only one curve. The error bars marks the limits within a 95 % confidence interval. Where only the upper limit is given, the lower limit is lacking. Higher precision can be achieved, making simulation time excessively long.**

In Figure 4.13, the total number of buffer inputs is set to 42. Number of reserved buffer inputs is varied from 0 to 16. The share of HCT traffic is set to 10 % and 50 % of the total traffic load, while the rest of the traffic consists of NCT traffic. The figure shows clearly that reserving buffer inputs gives a decrease in packet loss ratio for the HCT packets, while the NCT packets pay the price with a higher packet loss ratio. It is also confirmed that when the fraction of HCT traffic is increased to 50 %, the number of buffer inputs reserved for the HCT traffic has to be increased significantly in order to obtain the same packet loss ratio as for the case when HCT traffic load is 10 %.

When the number of reserved buffer inputs is set to 4 and the fraction of HCT traffic is set to 10 %, packet loss ratio is approximately three orders of magnitude higher for the NCT traffic than for the HCT traffic. The obtained PLR of  $< 10^{-7}$  satisfies the demands for the HCT class, while PLR of  $< 10^{-4}$  satisfies the demands for the NCT class. Assuming a 50/50 split of the traffic load

between the two QoS classes, PLR's that satisfies the QoS demands is obtained by reserving 8 buffer inputs.

#### ***4.4.4. QoS in MPLS Optical Networks***

In the first phase of building optical packet networks, QoS mechanisms can be based on MPLS, which offers resource reservation with proper control algorithms to support Class of Services and Traffic Engineering enhancing network efficiency.

##### *MPLS aspects of optical packet networks*

MPLS is a suitable technique for packet routing in optical packet networks. Optical packet nodes use information in packet labels to decide how to forward a packet. Within the optical node, the optical packet label is read and compared with a look up table. The payload is then transparently routed in the optical domain to the appropriate output port with a new label attached. Core MPLS optical nodes can process labels more rapidly than traditional address headers (i.e. IP headers), therefore network performance may be improved. This is especially desirable in optical packet networks, due to the weak buffering capabilities.

##### *MPLS control unit*

An MPLS controller should perform the following functions: (i) Building and maintaining the Label Information Base (LIB), (ii) MPLS signalling with support for CR-LDP and RSVP Tunnels for Label Distribution, (iii) forwarding which includes processing of incoming packets, making of forwarding decisions, packet shaping and delivering the packets to the outputs. Edge nodes should also be equipped with adaptation functions for incoming/outgoing traffic.

QoS tasks in an MPLS controller consist of packets classification in edge nodes, differentiated packets servicing in core nodes and Traffic Engineering that performs operations on traffic vectors such as merging, comparing, summarising and subtracting flows as well as LSP admission control and traffic load management.

##### *Support for QoS*

In order to address the QoS issue, the ability to introduce connection-oriented forwarding techniques to connectionless optical packet network is necessary. MPLS offers such functionality by establishing the label switched paths (LSPs). In effect, this allows optical network to reserve resources, such as buffers or wavelengths over predetermined paths for service differentiation,

providing QoS guarantees. Currently two main MPLS features for supporting QoS are used, differentiated servicing [48, 49] and Traffic Engineering.

#### *Differentiated Servicing – QoS in nodes*

The MPLS Class of Service (CoS) feature enables support for differentiated types of services across an MPLS optical network. The differentiated services model defines a variety of mechanisms for classifying traffic into a small number of service classes. Once packets are classified at the edge of the network, specific forwarding rules are applied in each node. This combination of packet marking and specific servicing procedure results in a scalable QoS solution for any traffic flow.

#### *Traffic Engineering – QoS in networks*

MPLS Traffic Engineering is a process of routing data traffic in order to balance the traffic load on the various links and nodes in the network. It assures e.g. better utilization of available bandwidth (e.g. automatically increasing or decreasing the bandwidth reserved for an MPLS TE tunnel based on measured traffic load), accommodation of high class traffic load to buffering capabilities in nodes, routing around failed links/nodes (reliability) and capacity planning. All these functions improve QoS of the networks.

#### **Summary**

In burst switching, service differentiation is achieved by e.g. allowing early reservation for high priority bursts through assigning them an extra offset time - called *QoS offset*. Therefore, high priority bursts make their reservation in a rather lightly loaded system and have a smaller loss probability while low priority bursts experience the total system load and have a higher loss probability. However, drawbacks of this principle are that the loss probabilities of high and low priority classes are highly dependent on burst characteristics and basic offset adaptation in the core nodes can change the differentiation in offset and lead to undesirable subclasses.

In OPS, when using FDL's, the QoS management techniques must be kept very simple due to the delay-oriented nature of the FDL buffers. In particular, if a feed-forward FDL buffer is used like in [8], it is not possible to change the order of packets coming out of the delay lines, thus making the pre-emption based techniques not applicable. Therefore, mechanisms based on a-priori access control of packets to the WDM buffers are necessary. Two techniques have been studied: 1) The *time-threshold-based technique*, where the resource reservation is applied to the time domain, and a delay threshold  $T_{low}$  is defined. In contrast with the high priority packets, the low-priority packets cannot

choose delays that are above the threshold, thereby allowing these packets to access only parts of the buffer. 2) The *wavelength-based technique* where the resource reservation is applied to the wavelength domain. Low-priority packets are allowed to access only a subset of the wavelength resources, while high-priority packets can access all the wavelength resources.

As opposed to using FDLs, if electronic buffers are used in combination with OPS, advanced techniques for QoS differentiation are feasible. Still, if efficient QoS differentiation can be achieved using simple principles, the number of components and thereby the costs can be reduced. In this report, a principle based on QoS differentiation in an optical packet switch with a limited number of electronic buffer interfaces has been studied. Differentiation on packet loss is achieved by giving high priority packets access to all buffer inputs, while low priority packets only can access parts of the buffer inputs. Obviously, this approach can be easily adapted to optical buffering with FDLs.

OPS can use MPLS for support of control and QoS differentiation. MPLS core nodes can process labels more rapidly than the traditional address headers (i.e. IP headers); therefore, network performance may be improved. This is desirable especially in optical packet networks, due to their weak buffering capabilities. By establishing label switched paths (LSPs), resources can be reserved in order to provide QoS guarantees, such as buffers, or wavelengths over pre-determined paths.

It is shown that all the described reservation techniques for QoS differentiation have the capability of supporting the PLR performance and the QoS differentiation suggested in the introduction. However, the burstiness of the traffic pattern will influence the performance of the suggested techniques. The design of the OPS/OBS nodes and of the network, as well as of the reservation parameters, will have to be decided according to the traffic pattern.

## 4.5 Optical Packet Switching in Metro Networks

Nowadays, the optical equipment vendors show increasing interest in enhancing existing and developing new packet-based technologies. The trend in networking is migration of the packet-based technologies from access networks to Metropolitan Area Networks (MANs). The special attention paid to MANs is a result of the rapidly increasing data traffic volume in the metro networks, which is challenging the capacity limits of the existing transport infrastructures such as SONET/SDH and ATM. In such a situation, packet-based transport technology, a natural fit with the now ubiquitous IP protocol,

appears to be one of the best choices for meeting the cost/efficiency trade-off in the metro networks. In this context, researchers world-wide are making great efforts to investigate new MAN architecture solutions (e.g. Resilient Packet Ring [50]).

OPS based on packet switching at the optical level that can also provide high throughput appears to be a good candidate for future MAN architectures (Optical MANs, O-MANs).

In this subsection we identify the main requirements for O-MANs and use them to compare five different O-MANs architectures: three currently under development in different research/university centres ([51], [52] and [53]), and two solutions investigated within the COST 266 Action [54] and IST project DAVID [3].

#### ***4.5.1. O-MAN Properties***

In the current networking context, a MAN (and consequently an O-MAN) has to meet the following requirements:

- *Flexibility*. It must be able to handle different granularities of bandwidth and to support a wide range of protocols.
- *Cost-effectiveness*. It must beat decisively the current technologies both in CAPEX and OPEX costs.
- *Upgradability*. It has to be able to incorporate new technologies in an easy and non-disruptive manner.
- *Scalability*. It must be possible to remove and add network devices in an easy and non-disruptive way
- *Efficiency*. It must provide high throughputs and short delays.
- *Fairness*. Starvation of nodes must be avoided through a regulation of bandwidth usage.
- *Multicasting*. It must allow multicasting in order to efficiently support applications such as videoconferences or distributed games.
- *Quality of Service*. It must have rapid provisioning capabilities and provide service guarantees to mission-critical data and delay-sensitive applications.
- *Reliability*. The network elements must offer a high degree of reliability. This mandates that critical sub-systems be fully protected and capable of in-service upgrade.



Clearly, these requirements are better met by packet-based technologies than by circuit-based technologies such as SONET/SDH rings.

HORNET (Figure 4.14a), which stands for Hybrid Opto-Electronic Ring NETwork, is a WDM time-slotted ring developed at the Stanford University [51]. HORNET has a tunable transmitter, fixed receiver design (TTFR), and the nodes use a CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) MAC protocol to govern access to the wavelengths (see [51] for more details on the access protocol). HORNET can use either slotted or variable-length packets- a characteristic which can provide flexibility. It can be based on two counter-rotating rings to offer cut-fibre protection; nevertheless, a failure will result in halving the available bandwidth. Multicast can be provided via node-by-node re-transmission, but no protocol is included and evaluated to handle multicast traffic and incoming traffic at the nodes. Neither QoS strategies nor fairness mechanisms are implemented. No performance evaluations to assess the merits of such architecture are available.

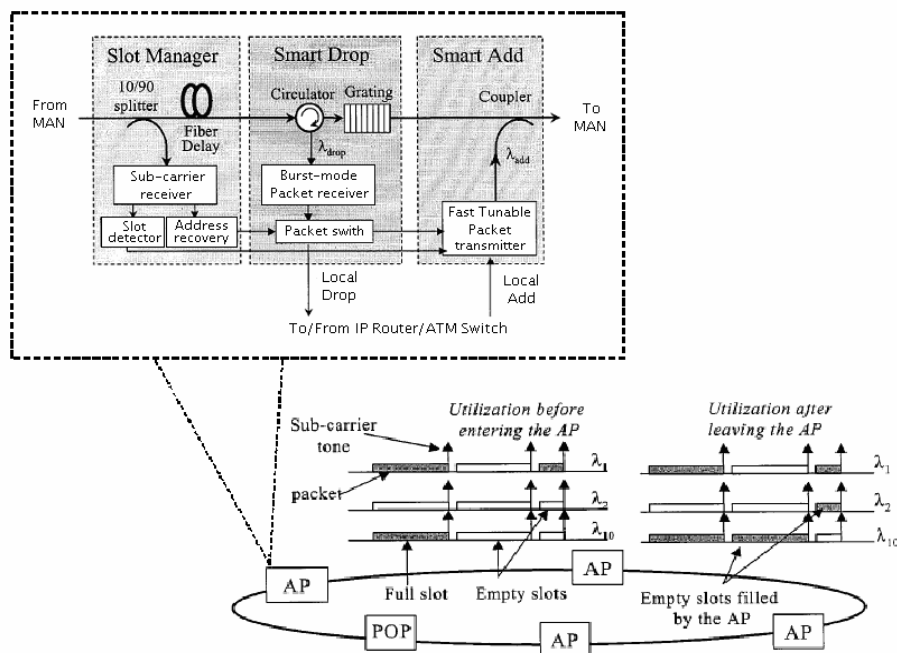


Figure 4.14.a. - O-MAN architectures: HORNET

RingO (Figure 4.14b) is a WDM time-slotted ring developed at the Politecnico di Torino [52]. The peculiarity of the network is that the number of nodes is

equal to the number of wavelengths. In this way, each node is equipped with a laser array (for supporting multicast traffic), and a fixed receiver operating on a given wavelength that identifies the particular node. This wavelength is extracted from the ring by the same node: in order to communicate with node  $k$ , a node uses  $\lambda_k$ . The MAC protocol is based on a generalisation of the empty-slot approach, where each node is able to check the state of the wavelength occupancy on a slot-by-slot basis. No other features are currently implemented in the test-bed, nevertheless a fairness mechanism based on a generalization of the SAT token, a QoS strategy based on differentiated service, and several methods for supporting variable-length packets have been suggested and evaluated by simulations. No protection mechanisms have been implemented and, currently, the scalability is largely affected by the imposition of the number of wavelengths equal to the number of nodes.

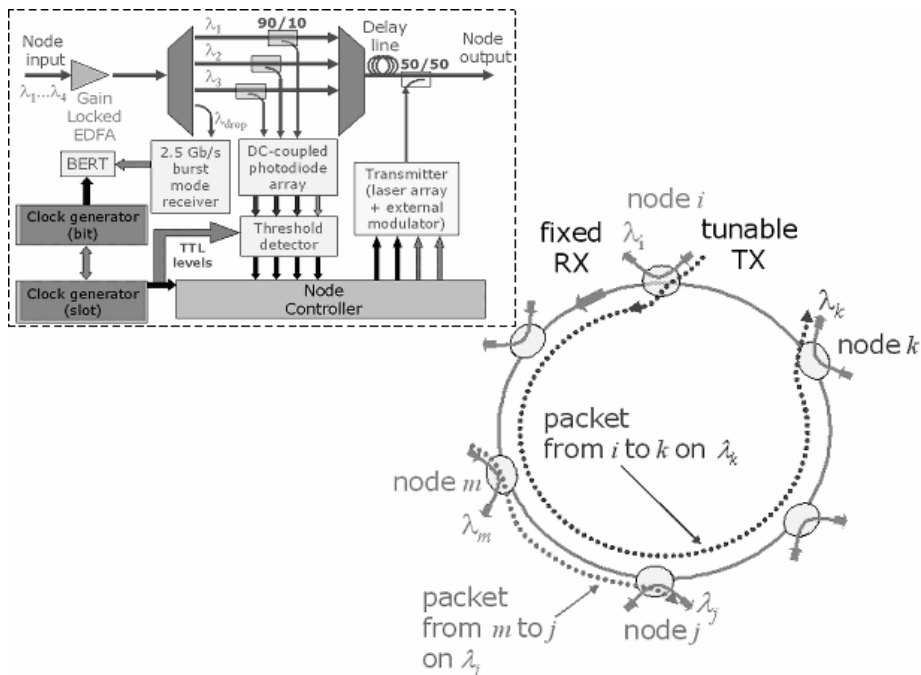


Figure 4.14.b. - O-MAN architectures: RingO

DBORN (Figure 4.14c), which stands for Dual Bus Optical Ring Network, is based on a unidirectional fibre ring organized around a Hub developed at Alcatel Research & Innovation [53]. This architecture uses a spectral separation of upstream and downstream flows from/toward the Hub, forming a dual logical bus structure. Nodes dynamically read data on the downstream bus and

write on the upstream bus, while the Hub interconnects the buses through a wavelength conversion. The spectral separation avoids the use of erasing functionality at the nodes increasing the nodal cascadability. A simple collision avoidance MAC is implemented through power detection utilizing a photodiode and a fixed-length delay line. The network can support any client packets and makes it easy to add/remove nodes to/from the network. Some cost studies have shown the benefits of this architecture [53]. A performance evaluation is not available; neither QoS strategies nor fairness mechanisms have been implemented. The current protection mechanism is based on duplicating the network components.

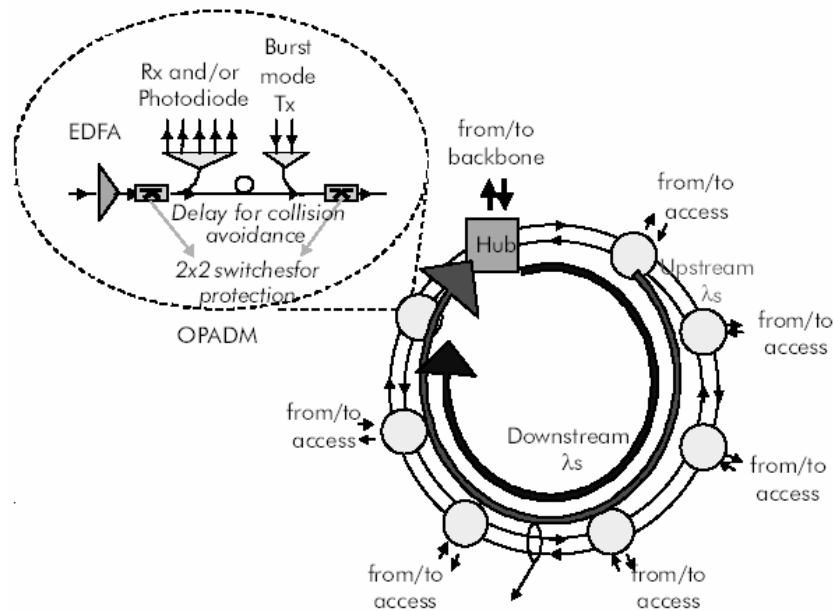
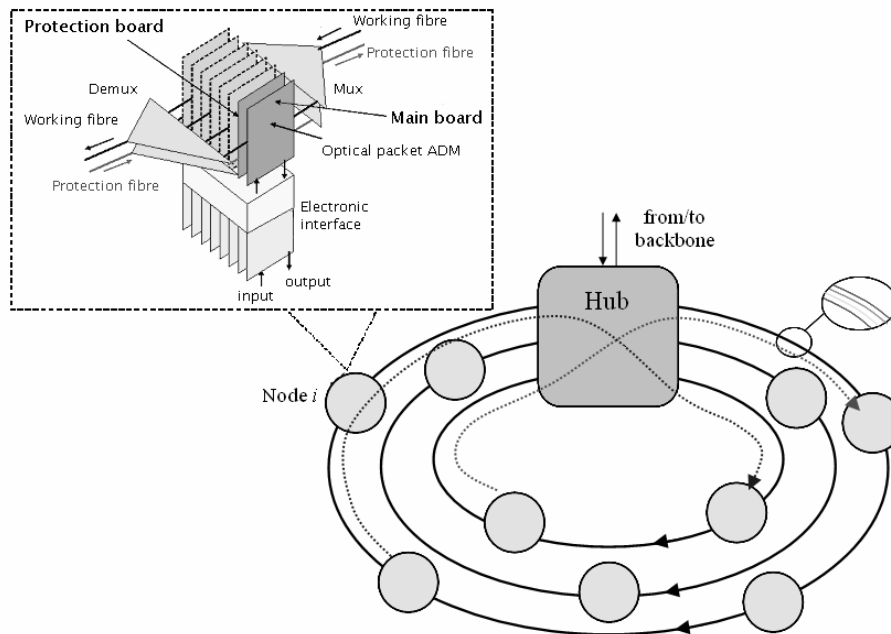


Figure 4.14.c. - O-MAN architectures: DBORN

All the above proposals are based on simple ring architectures where the overall throughput is limited to hundreds of Gbit/s and the access networks attached to the metro nodes are based on copper technologies. A Fibre-To-The-Home scenario would significantly increase the traffic demand up to a few Tbit/s throughputs. Therefore, our research has been concentrated on studying and evaluating novel advanced optical architectures based on multiple rings or multiple trees able to achieve more than 1 Tbit/s throughput.

We have theoretically studied the architecture shown in Figure 4.14d developed within the IST project DAVID. It consists of several optical slotted

packet unidirectional rings interconnected by an optical packet switch called Hub [3]. The performance of such architecture shows good performance in efficiency (both throughput and end-to-end delay) and effective fairness [55]. Moreover, several QoS mechanisms have been proposed and evaluated based both on differentiated service scheme and on guaranteed service provisioning. Some protection schemes have been designed and compared [3].



**Figure 4.14.d. - O-MAN architectures: DAVID**

The O-MAN architecture shown in Figure 4.14 e) consists of several slotted optical packet unidirectional trees interconnected by an array wavelength grating (AWG) which provides a static wavelength routing [54]. A Network Controller (NC) manages the network resources through a proper scheduling algorithm. This architecture recalls the well-known SONATA network. Two different scheduling solutions have been evaluated by simulation [56] [57], and a QoS provisioning scheme providing differentiated service has been proposed in [57]. A protection mechanism based on coupling an AWG with a Passive Star Coupler and a multicast support using optical splitters at the output of AWG have been suggested and evaluated in [58]. Cost study including scalability, upgradeability and reliability issues are in progress. Preliminary results show that this solution seems less costly than the DAVID solution [59].

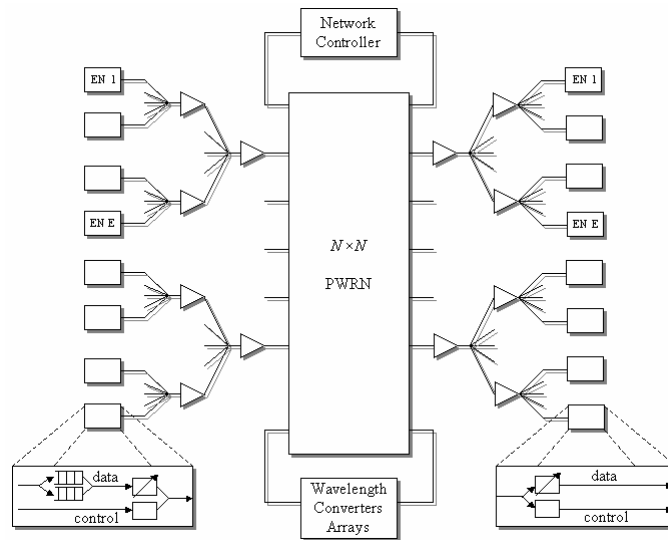


Figure 4.14.e. - O-MAN architectures: SONATA

Table 4.1 shows comparison of the above-described O-MAN architectures. The symbol – indicates that the O-MAN does not implement and/or consider that feature.

Table 4.1. - Comparison of O-MAN architectures

Feature	HORNET	RingO	DBORN	SONATA	DAVID
Flexibility	medium	medium	high	medium	medium
Cost-effective	high	medium	high	high	medium
Upgradability	high	high	medium	–	–
Scalability	medium	low	high	medium	medium
Efficiency	medium	medium	low	high	high
Fairness	–	high	–	high	high
Multicasting	medium	high	–	high	–
QoS	–	medium	–	medium	high
Reliability	medium	–	–	medium	medium

### 4.6 Conclusion and Outlook

In COST 266, optical burst and packet switching have been studied with respect to performance and complexity. Node designs, contention resolution strategies and QoS architectures in OBS have received special attention.

The work done in COST 266 and recently published papers on OPS/OBS show an emerging trend towards asynchronous packet and burst switching. Comparative concept studies show that packet and burst switching share several properties, but also have some important differences like in the control signalling schemes, the granularity and the overhead.

Key design parameters, as well as a detailed node design of OPS/OBS networks, are described in this chapter. In order to achieve a sufficiently low burst or packet loss rate in the described node designs, contention resolution schemes must be employed, both in OBS and OPS. In combination with the wavelength dimension, both FDL and electronic buffering are compared for OBS and OPS in a common scenario. It is shown that both buffering technologies are able to significantly reduce loss probability. While electronic memory has the potential of using fewer buffer interfaces, O/E/O interfaces are expected to be a major contributor to the cost of the buffer.

For OBS, a classification of several QoS schemes is given, and the offset-based QoS is studied in more detail. For support of control and QoS differentiation, OPS can use MPLS reserve resources and provide QoS guarantees, such as buffers, or wavelengths over predetermined paths.

Finally, metropolitan area networks (MAN) have been investigated. In the MAN environment, the advantages of OPS solutions are highlighted, and the main requirements have been identified and described. These requirements are used to compare optical metro network architectures that are currently under development in research/university centres with two new advanced architectures that have been studied within the COST 266 action and the IST DAVID project.

The performance of a node depends on the available resources, which are the results of several trade-offs between performance and complexity. In addition, client layer requirements, technology status, as well as the topology and transmission systems of existing WDM networks influence the network design.

Further work should therefore include network and end-to-end client layer performance studies, comparing the OCS and OBS/OPS techniques, with respect to performance and CAPEX/OPEX. Scalability is an issue gaining increased importance as the need for large networks increases. Further work on node and network design should take into account the need for network scalability, both with respect to throughput, signal quality and cost efficiency. Topics suitable for further studies are e.g. end-to-end QoS schemes, contention resolution schemes and their influence on the general performance of

OPS/OBS networks, both in the backbone and metropolitan environments. How these schemes can be combined with a GMPLS control scheme should also be investigated.

#### 4.7 References

- [1] F. Masetti *et al.*, "High speed, high capacity ATM optical switches for future telecommunication transport networks," *IEEE Journal of Selected Areas in Communications (JSAC)*, 14(5), June 1996.
- [2] C. Guillemot *et al.*, "Transparent Optical Packet Switching: the European ACTS KEOPS project approach," *IEEE Journal of Lightwave Technology (JLT)*, 16(12), pp. 2117-2134, Dec. 1998.
- [3] L. Dittman *et al.*, "The European IST Project DAVID: a Viable Approach towards Optical Packet Switching," accepted for *IEEE JSAC*, 2003.
- [4] K. Vlachos *et al.*, "An Optical IM/FSK Coding Technique for the Implementation of a Label-Controlled, Arrayed Waveguide Packet Router", to appear in *IEEE JLT*, Nov. 2003.
- [5] M. Yoo, M. Jeong, and C. Qiao. "A High Speed Protocol for Bursty Traffic in Optical Networks," *SPIE's All-Optical Communications Systems*, vol. 3230, pp. 79-90, Nov. 1997.
- [6] M. Nord, S. Bjørnstad, and C. Gauger. "OPS or OBS in the Core Network?" *Proc. of Optical Network Dimensioning and Modeling (ONDM) Conference*, Feb. 2003, Budapest.
- [7] I. Chlamtac *et al.*, "CORD: contention resolution by delay lines," *IEEE JSAC*, 14(5), pp. 1014-1029, June 1996.
- [8] F. Callegati *et al.*, "Architecture and performance of a broadcast and select photonic switch," *Optical Fiber Technology*, 4(3), pp. 266-284, July 1998.
- [9] F. Callegati, "Optical buffers for variable length packets," *IEEE Comm. Letters*, 4(9), pp. 292-294, Sept. 2000.
- [10] L. Tančevski *et al.*, "Optical routing of asynchronous, variable length packets", *IEEE JSAC*, 18(10), pp. 2084-2093, Oct. 2000.
- [11] S. Bjornstad, D. R. Hjelm, and N. Stol, "An optical packet switch design with shared electronic buffering and low bit rate

- add/drop inputs,” *Proc. of Int. Conference on Transparent Optical Networks (ICTON) 2002*, vol.1, pp. 69-72, Warsaw, Poland, 2002.
- [12] S. Bjornstad, D. R. Hjelm, and N. Stol, ”A scalable optical packet switch for variable length packets employing shared electronic buffering”, *Proc. European Conference on Optical Communication, ECOC 2002*, Copenhagen, Sept. 2002.
- [13] C. Develder *et al.*, ”Node architectures for optical packet and burst switching,” *Proc. Int. Topical Meeting on Photonics in Switching (PS2002)*, (invited) paper PS.WeA1, pp. 104–106, Cheju Island, Korea, July 2002.
- [14] F. Callegati and W. Cerroni, ”Time-wavelength exploitation in optical feedback buffer with trains of packets,” *Proc. SPIE Opti-Comm 2002*, vol. 4874, pp. 274-285, Boston, USA, July 2002.
- [15] P. B. Hansen, S. L. Danielsen, and K. E. Stubkjaer, ”Optical packet switching without packet alignment,” *Proc. ECOC 1998*, pp. 591-592, Madrid, Spain, 1998.
- [16] Z. Haas, ”The 'staggering switch': An electronically controlled optical packet switch,” *JLT*, May 1993, pp. 925-936.
- [17] A. Ge, F. Callegati, and L. S. Tamil, ”On optical burst switching and self-similar traffic”, *IEEE Comm. Letters*, 4(3), Mar 2000, pp. 98 -100.
- [18] F. Xue and S.J. Ben Yoo, ”Self-similar traffic shaping at the edge router in optical packet-switched networks”, *Proc. International Conference on Communications, ICC 2002*, vol. 4, pp. 2449-53.
- [19] G. Hu, K. Dolzer, and C.M. Gauger. ”Does burst assembly really reduce the self-similarity?” *Proc. of Optical Fiber Communication (OFC) conference 2003*, Atlanta, USA, 2003.
- [20] Technical Committee Telecommunications COST 266, Workgroup 2, Advanced Infrastructure for Photonic Networks, Optical Packet and Burst switching Report. (Ed:S. Bjørnstad) Ch. 2. (S. Vandweerd *et al.*) ”Optical burst and packet switching research”, [www.ure.cas.cz/dpt240/cost266/](http://www.ure.cas.cz/dpt240/cost266/), 2002.
- [21] K. Claffy, G. Miller, and K. Thompson, ”The nature of the beast: Recent traffic measurements from an Internet backbone,” *Proc. the annual meeting of the Internet Society 1998. (INET'98)*.



- <http://www.caida.org/outreach/papers/1998/Inet98/index.xml>  
Caida 1998.
- [22] Technical Committee Telecommunications COST 266, workgroup 2, *Advanced Infrastructure for Photonic Networks, Optical Packet and Burst switching Report*. (ed. S. Bjørnstad) Ch 4. (M. Nord) "Analysis of required switching times and identification of suitable switching technologies in optical packet switching," [www.ure.cas.cz/dpt240/cost266/](http://www.ure.cas.cz/dpt240/cost266/), 2002.
- [23] S. Junghans and C. Gauger "Architectures for Resource Reservation Modules for Optical Burst Switching Core Nodes," *Proceedings of the 4<sup>th</sup> ITG Workshop on Photonic Networks*, Leipzig, May 2003.
- [24] D. Sahri *et al.* "A Highly integrated 32-SOA gates optoelectronic module suitable for IP multi-terabit optical packet routers," *Proc. OFC'01* PD32, Anaheim, USA, 2001.
- [25] F. Masetti *et al.* "Design and Implementation of a Multi-Terabit Optical Burst/Package Router prototype," *Proc. OFC'02*. vol.1, FD11 - FD13. Anaheim, USA, 2002.
- [26] H. Buchta, C.M. Gauger, E. Patzak, and J. Saniter, "Limits of effective throughput of optical burst switches based on semiconductor optical amplifiers," *Technical Digest of the Optical Fiber Communication Conference (OFC 2003)*, Atlanta, March 2003, pp. 215-217.
- [27] J.Gripp *et al.* "Demonstration of a 1.2 Tb/s Optical Packet Switch Fabric (32 \* 40 Gb/s) based on 40 Gb/s Burst-Mode Clock-Data-Recovery, Fast Tuneable Lasers, and a high-performance NxN AWG," *ECOC 2001*, PD A.1.7, Amsterdam, the Netherlands, 2001.
- [28] J. Cheyns, C. Develder, E. Van Breusegem, E. Baert, A. Ackaert, M. Pickavet, and P. Demeester, "Routing in an AWG based optical packet switch," *Photonic Network Commun. (PNET)*, Kluwer Academic Publishers, 5(1), pp. 69 - 80., January 2003.
- [29] J. Cheyns, E. Van Breusegem, C. Develder, A. Ackaert, M. Pickavet, and P. Demeester, "Performance improvement of an internally-blocking optical packet/burst switch," accepted for ICC 2003, , Anchorage, Alaska, USA, 11-15 May 2003.

- [30] M. C. Chia *et al.* "Packet Loss and Delay Performance of Feedback and Feed-Forward Arrayed-Waveguide Gratings-Based Optical Packet Switches With WDM Inputs-Outputs," *IEEE JLT* 19(9), September 2001.
- [31] C. Develder, M. Pickavet, and P. Demeester, "Strategies for an FDL based feed-back buffer for an optical packet switch with QoS differentiation", *Proc. Int. Conf. on Optical Internet (COIN2002)*, paper COIN.TuD1, pp. 114–116, Cheju Island, Korea, 21–25 Jul. 2002.
- [32] Y. Hida *et al.* "400-channel 25-GHz spacing arrayed waveguide grating covering a full range of C and L-bands," *OFC'01*, vol. 3, WB2, 1-3, Anaheim, USA, 2001.
- [33] Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit, "All-optical packet-switched networks: a study of contention-resolution schemes in an irregular mesh network with variable-sized packets," *Proc. OptiComm 2000*, (Dallas), October 2000.
- [34] S. L. Danielsen, B. Mikkelsen, C. Joergensen, and T. Durhuus, "WDM packet switch architectures and analysis of the influence of tuneable wavelength converters on the performance," *IEEE JLT*, vol. 15, pp. 219–227, Feb. 1997.
- [35] V. Eramo and M. Listanti, "Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters," *IEEE JLT*, vol. 18, pp. 1818–1833, Dec. 2000.
- [36] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in Optical Packet Switches," *IEEE JLT*, 16(12), pp. 2081-2094, Dec. 1998.
- [37] F. Callegati, W. Cerroni, and G. Corazza "Optimization of Wavelength Allocation in WDM Optical Buffers," *Optical Networks Magazine*, 2(6), pp. 66-72, Nov. 2001.
- [38] C. M. Gauger, "Dimensioning of FDL buffers for optical burst switching nodes," *Proc. ONDM 2002*, Torino, 2002.
- [39] C. M. Gauger, "Performance of converter pools for contention resolution in optical burst switching," *Proc. SPIE OptiComm 2002*, Boston, July 2002.
- [40] ITU-T Recommendation Y.1541: Series Y: " *Global information infrastructure and internet protocol aspects. Internet protocol aspects – Quality of service and network performance: Network*

- performance objectives for IP-based services,*" Prepublished Recommendation (05/2002).
- [41] Black, Uyles, "Advanced Internet Technologies," Prentice Hall, pp. 123-128, 1999.
- [42] R.J. Gibbens *et al.* "An Approach to Service Level Agreements for IP networks with Differentiated Services," Statistical Laboratory, University of Cambridge and Internet and Data Networks, British Telecommunications plc. (Article submitted to Royal Society).
- [43] M. Yoo, C. Qiao, and S. Dixit, "QoS Performance in IP over WDM Networks", *IEEE JSAC- Special Issue on the Protocols for Next Generation Optical Internet,*" 18(10), pp. 2062-2071, Oct. 2000.
- [44] K. Dolzer, C. Gauger, "On Burst Assembly in Optical Burst Switching Networks – A Performance Evaluation of Just-Enough-Time," *Proc. 17th International Teletraffic Congress (ITC 17),* Salvador/Brazil, 2001.
- [45] C. Gauger, "Determining offset times in optical burst switching networks", Presentation at COST 266 MCM, Zagreb, June 2001.
- [46] F. Callegati, G. Corazza, and C. Raffaelli, "Exploitation of DWDM for Optical Packet Switching with Quality of Service Guarantees," *IEEE JSAC*, 20(1), pp. 190-201, Jan. 2002.
- [47] S. Bjørnstad, N. Stol, and D. R. Hjelm. "Quality of service in optical packet switched DWDM transport networks," *Proc. Asia Pacific Optical and Wireless communications conference and exhibition 2002 (APOC 2002, SPIE conference),* Shanghai, China, 2002.
- [48] E. Horlait and N. Rouhana, "Differentiated Services and Integrated Services Use of MPLS," *Proc. 5th IEEE Symposium on Computers and Communications (ISCC 2000),* pp 1-6, 2000.
- [49] M. Klinkowski and M. Marciniak, "Services differentiation in MPLS photonic packet networks," *Proc. ONDM 2003,* Budapest, Hungary, Feb. 2003.
- [50] IEEE 802.17, Resilient Packet Ring Working Group, <http://grouper.ieee.org/groups/802/17/>

- [51] K.V. Shrikhande *et al.*, "HORNET: a packet-over-WDM multiple access metropolitan are ring network," *IEEE JSAC*, 18(10), Oct. 2000.
- [52] R. Gaudino, *et al.*, "RINGO: a WDM ring optical packet network demonstrator," *Proc. ECOC 2001 Th.L.2.6*, Amsterdam, Netherlands, Oct. 2001.
- [53] N. Le Sauze, *et al.*, "A novel, low cost optical packet metropolitan ring architecture," *Proc. ECOC 2001*, PD, Amsterdam, Netherlands, October 2001.
- [54] D. Careglio, J. Solé-Pareta, S. Spadaro, and G. Junyent, "Performance Evaluation of Interconnected WDM PONs Metro Networks with QoS Provisioning," *Proc. 2<sup>nd</sup> Workshop on All-Optical Network (WAON 2001)*, Zagreb, Croatia, June 2001.
- [55] D. Careglio, J. Solé-Pareta, S. Spadaro, and G. Junyent, "Evaluación de prestaciones del la red metropolitana del proyecto DAVID", *12<sup>th</sup> Telecom I+D*, Madrid, Spain, Nov. 2002 (in Spanish).
- [56] D. Careglio, J. Solé-Pareta, S. Spadaro, and G. Junyent, "Performance Evaluation of Interconnected WDM PONs Metro Networks with QoS Provisioning," *Proc. ONDM 2003*, Budapest, Hungary, Feb. 2003.
- [57] D. Careglio, A. Rafel, J. Solé-Pareta, S. Spadaro, and G. Junyent, "Quality of service strategy in an optical packet network with a multi-class frame-based scheduling," *Proc. HPSR 2003*, Turin, Italy, June 2003.
- [58] C. Fan, M. Maier, and M. Reisslein, "The AWG||PSC network: a performance enhanced single-hop WDM network with heterogeneous protection," *Proc. of IEEE Infocom 2003*, San Francisco, CA, April 2003.
- [59] A. M. Hill, D. Careglio, J. Solé-Pareta, and A. Rafel, "Relative costs of WDM rings and PONs for metro optical packet networks," *Proc. PS 2003*, Paris, France, Sept. 2003.

## Chapter 5

# Availability Analysis

*Robert Inkret (editor), Didier Colle, Sophie De Maesschalck, Marko Lacković, Marije Ljolje, and Branko Mikac*

### 5.1 Introduction

Due to almost exponential growth of traffic in telecommunication networks, among different scenarios of transport network evolution, the most certain one is deployment of WDM based all-optical networks. In a network where components carry huge amount of traffic, it becomes clear that the aspect of resilience to the various failures is of outmost importance. As a kind of generic term, the level of resilience to failures is called survivability. Hence, survivable networks are those with the required level of resilience to failures.

When designing all-optical network architecture, it is very important to pinpoint optical components, devices and even architectural choices that are critical to the network survivability. This is accomplished in the analysis process during which various network architectures or concrete networks are compared. Such a comparative analysis requires a quantitative survivability measures. These can be classified into three groups; probability-based (availability), traffic-based (e.g. expected loss of traffic), and topology-based measures (e.g. connectivity, diameter, etc.).

Within the framework of COST 266 project, two teams dealt with the reliability and availability analysis of all-optical networks. The team from University

of Zagreb mainly dealt with the probability based survivability measure, namely availability, while team from the University of Ghent dealt with traffic based survivability measures. This chapter presents reports from both teams. In the first part, an overview of availability model together with the results of application of the model to the COST266 case studies is presented. In the second part results from the team from University of Ghent are presented.

## 5.2 Availability Performance Analysis

In order to analyse network availability performance, one has to define considered architectures, the topologies onto which the architectures were applied, and the models used for performance analysis. The section starts with the overview of considered network architectures, together with the protection switching schemes. More can be found in [3] and [4]. Next, availability model and measures are explained. COST 266 case study topologies presented in the chapter 1 of this report were used. Finally, results for the availability performance analysis applied to the considered topologies are presented. The section ends with comments and conclusions regarding presented analysis and results.

### 5.2.1. Description of Analysed Network Architectures

WDM optical network is composed of nodes and WDM links. A generic node structure depends on the type of optical network, while for an actual node structure it depends also on the network topology and design options. Each link is composed of a number of fibres ( $F$ ), each of which can carry up to  $W$  wavelength channels.

Two types of optical networks are considered in the model; passive optical WDM networks (PWN), without the possibility of automatic switching or rerouting of wavelength paths, and automatically switched optical WDM network (ASWN). Note that the terms passive and active are used in conjunction with node structure only.

A WDM network provides circuit switching service to the upper layer through the notion of either wavelength path (WP) or virtual wavelength path (VWP). Both WP and VWP are achieved by the concatenation of wavelength channels on optical links. The difference between the two is in wavelength continuity constraint in the case of WP. The constraint states that WP should be assigned to the same wavelength on all links in a path between the source and destination. Such constraint is not imposed in the case of VWP. Clearly, in the WP case we have a WDM network without the possibility of wavelength conversion.

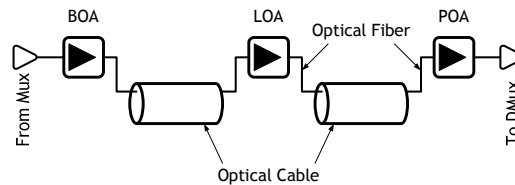
In the case of PWN, concatenation is implemented by the simple splicing (or by using optical connectors) in optical nodes, while in the case of ASWN concatenation is implemented by appropriate switching in optical switches.

In the presented WDM networks, all communications are assumed symmetrical, i.e., lightpaths between any two nodes in both directions use the same links and nodes.

A structure of optical link depicted in Figure 5.1 is identical for all considered network architectures. Here, distinction between three applications of optical amplifiers is made. After the signal has been processed in the optical node, it is amplified to the appropriate power level in a booster optical amplifier (BOA). At periodic length intervals, due to attenuation in optical fibre, the signal has to be re-amplified. This is achieved in line optical amplifiers (LOA). The number of LOA's,  $N_{OA}$ , depends on the length of the link  $L$  and the line amplifier span, i.e., the distance between optical amplifiers,  $d_{OA}$ ,

$$N_{OA} = \left\lceil \frac{L}{d_{OA}} \right\rceil - 1.$$

An additional component of the optical link is optical cable which could be viewed as a simple container of optical fibres.



**Figure 5.1. - Optical link structure**

### *Passive WDM Networks*

#### PWN without Protection

The optical network is composed of WDM multiplexers, de-multiplexers, optical amplifiers, cables and fibres. An entry point to the network is optical transponder (TP), as depicted in Figure 5.2, which translates incoming optical signal from a non-ITU-WDM grid wavelength to an ITU-WDM grid wavelength. Circuits are realized by manual connection of identical wavelengths in “optical nodes”, as depicted in Figure 5.2.

Clearly, in such networks, protection and/or restoration mechanisms should be implemented in higher layers.

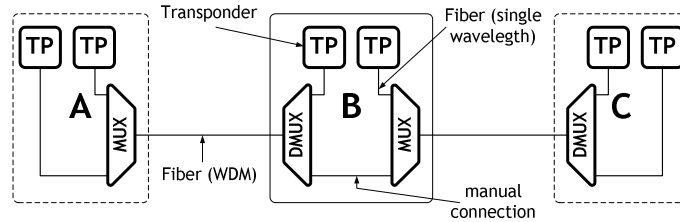


Figure 5.2. - PWN without protection

#### PWN with 1+1 Protection

A simple introduction of protection mechanisms into PWN is the use of 1+1 protection. In such protection, for each circuit between source and destination, two wavelength paths are used for both directions, totalling to 4 lightpaths (see Figure 5.3, which shows only one direction). The transmitting transponder sends signal on both lightpaths, termed primary and spare, while the receiving transponder, according to the level of input power, chooses the proper incoming lightpath.

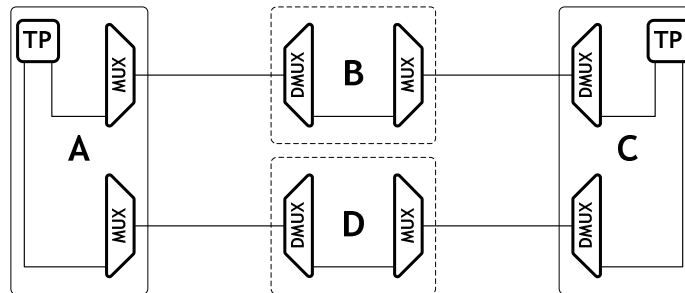


Figure 5.3. - PWN with 1+1 protection

In order to have such network capable of surviving any single failure, elements of any two lightpaths should not share physical elements. It is clear that for this case it is sufficient to place lightpaths onto two link-independent paths (there is no conditional probability between link failures on these paths).

#### Automatically Switched WDM Network

##### ASWN without Protection

If optical space switches and appropriate network management is used instead of manual switching, the optical network becomes capable of supporting automatic switching, i.e., we have an automatically switched WDM network (ASWN).



In the case of the WP scheme, one of the possible generic node structures is depicted in Figure 5.4. Note that the figure does not show the control and management entities. Since a lightpath is not allowed to change its wavelength, for each wavelength we can use exactly one optical non-blocking switch. Such structure allows switching of a lightpath between any combination of input and output fibres.

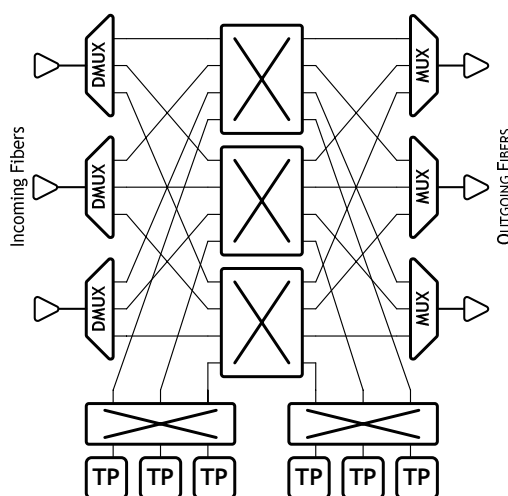


Figure 5.4 - Generic node structure for ASWN without wavelength conversion

The only difference with respect to PWN without protection is in the capability to automatically provide non-protected lightpaths.

#### ASWN with 1+1 Protection

The generic structure of a node for this type of network is the same as in the case of ASWN w/o protection. The main difference is in management software, or more precisely, part of the management software that manages allocation of resources. For each lightpath an additional, spare lightpath should be reserved. Once again, it is clear that in order to have a network capable of surviving any single failure, primary and spare lightpaths should be assigned onto two node independent paths.

As in the case of PWN with 1+1 protection, it is the responsibility of transponders to split signals at the transmitting side, and to pick the best signals out of the two in the receiving direction.

The motivation for this scheme is the fastest protection switching time, since for switching to a spare path, the use of complex network management resources is not needed.

#### ASWN with 1:1 Path Protection

Unlike in 1+1 protection, in 1:1 path protection scenario the spare path is used exclusively in the case when there is a failure on the primary path. Hence, in cases without such failures, a spare path can be used to carry the lower priority lightpaths. In a special case, the network elements allocated to spare lightpaths can be shared among a number of spare lightpaths. Clearly, if primary lightpaths sharing two spare lightpaths are element-independent, no single failure will ever affect both of the primary lightpaths. Consequently, if both spare lightpaths are assigned to the same network elements, a network will still be able to survive any single element failure.

In a real network, such scheme can be implemented by pre-calculation of primary and spare lightpaths for each demand.

#### ASWN with Path Restoration

Restoration is the most complex P&R mechanism that can be implemented in optical networks. Among P&R techniques, restoration is usually viewed as the most efficient one in terms of network resources use, but the obvious disadvantage is in the restoration time that can be of unacceptable level. Restoration time can be kept at a reasonable level by pre-computation of spare lightpaths. For example, whenever a demand is routed through the network, a corresponding spare lightpath is computed by the network management. Thereafter, a new spare lightpath is computed only if the primary or spare lightpaths are affected by a failure. If a failure affects the primary lightpath, the switching can be achieved in a fast manner.

#### **5.2.2. Availability Model**

Availability ( $A$ ) is the probability that the given entity is functional at the given instance of time. Unavailability ( $U$ ) is the complementary measure used to present availability figures that are typically very close to 1, in a format easier to read.

Availability model makes sense only for repairable systems. In the paper we present availability results for transport entities that are calculated from the availabilities of network components. Therefore, availability of transport entities can be treated as probability that corresponding logical expression, defining logical relationships between the considered transport entity and network components, is true. Input to the availability analysis, in general, are the probability density functions (PDF) for both failure and repair times of network components. In this report we assume exponential distribution of failure and repair times, thus, for each component the mean time to failure ( $MTTF$ ) and to

repair (*MTTR*) should be given. For the exponential failure and repair times, availability function tends to steady-state (asymptotic) availability  $A_s$  if  $t$  tends to  $\infty$ .

Instead of *MTTF*, we are using the notion of failure rate  $\lambda$  which represents the number of failures per time interval. This measure is expressed in FIT (failures in time) units; 1 FIT corresponds to 1 failure per  $10^9$  hours. Similarly, *MTTR* could be expressed as repair intensity  $\mu$ , the number of repair actions per time interval. The value of  $\mu$  is the reciprocal value of the mean time to repair;  $\mu=1/MTTR$ .

If a non-redundant and repairable system or component is considered, the steady-state (un)availability can be derived by using the Markov availability model. For the case if  $\lambda \ll \mu$ , the unavailability formula could be written as

$$U \approx \frac{\lambda}{\mu} = \lambda \cdot MTTR .$$

The first transport entity, starting from the bottom, is the *wavelength channel* (*WCh*). As described before, a wavelength channel corresponds to a single wavelength on any fibre in the network. A wavelength channel is available only if a corresponding fibre, optical line amplifiers, booster optical amplifiers, optical pre-amplifiers, multiplexers and de-multiplexers are available.

In most cases, however, a failure of a fibre is caused by the cut of the corresponding cable which affects all fibres within the cable. We assume that each link contains exactly one optical cable.

The next transport entity is the *lightpath* (also *wavelength path*, hence *Wp*). A lightpath can be viewed as the concatenation of wavelength channels on a given path. Note that a lightpath is determined by the source and destination nodes, the path on which it is routed, and the wavelength that is identical for each wavelength channel in a lightpath. A combination of paths, fibres and wavelength is usually determined during the design process.

A lightpath is a unidirectional transport entity of a standard bit-rate. The corresponding bidirectional transport entity is the logical channel (*LCh*), or just channel. In the simplest case, for the no protection case, a channel is composed of only two lightpaths, one per each direction.

For the 1+1 protection, each direction uses two lightpaths, routed over two independent paths (node or link), the primary and the spare. A channel will be available if for each direction at least one lightpath is available.

At the end, a logical connection (*LC*), or just connection, is composed of several channels. The number of channels depends on the capacity of each channel (full wavelength capacity, *C*), and the capacity of the logical connection. Two definitions for the availability of logical connection can be used. The first one assumes that connection is available only if all of the channels are available (series or non-redundant structure). The second one assumes protection (or restoration) in a way that connection is available if at least one of its channels is available (parallel or redundant structure).

In the report we calculate the availability for each logical connection in the network. This is calculated as the probability that corresponding logical expression is true. In our case, each network element is treated as a logical variable; an element can be available (1) or not (0). Analytical methods [4] have been used in order to obtain the availability performance figures.

### **5.2.3. Network Modelling Procedure**

Network was modelled using the above described PWN and ASWN architectures with no protection, 1+1 protection and 1:1 restoration. The modelling procedure comprises the following steps:

1. Creating topology based on the node list and connectivity matrix – each node represents a WDM Point-of-Presence node, which serves as ingress/intermediate/egress node of the WDM core network. Node structure depends on the chosen network architecture and P&R mechanism.
2. Calculating traffic demands between nodes. A traffic model presented in [1] has been used.
3. Creating transport entities. The number of transport entities depend on the traffic demands as the SDH channel capacity granularity is imposed on the total traffic demands between the end nodes (traffic demands must be supported by SDH interfaced line cards).
4. Using graph colouring in the architecture without wavelength converters to assign the list of wavelengths and fibres to each lightpath. This step depends on the chosen P&R architecture as protection transport entities are also being defined.
5. Creating WDM equipment and dimensioning the access and interconnection cables.

Described steps produce a WDM network capable of supporting the chosen P&R mechanism using chosen network architecture. This is followed by a network analysis which can include:

- Equipment statistics,
- Cost calculation,
- Availability calculation:
  - i. Analytical method,
  - ii. Monte Carlo simulation.

#### 5.2.4. Results

In this section results for the COST 266 case study topologies and different P&R architectures are presented. All calculations assume the availability data for photonic components as presented in Table 5.1. In addition, all networks have been designed with the assumption of 16 wavelengths per fibre. Traffic matrices are estimated for the year 2002. No wavelength conversion has been assumed. The distance between amplifiers is 80 km for all presented results.

**Table 5.1. - Availability data**

Component Type		Failure Rate [fit]	MTTR [hours]
Transponder (source + receiver)	T	2000	6
Wavelength Converter	WC	200	6
Multiplexer	M	200	6
Demultiplexer	D	200	6
Cross Connect	XC	1000	6
Cable [1/km]	C	100	21
Amplifier (EDFA)	A	650	6
Splitter		50	6

The following tables (Table 5.2, 5.3, 5.4) show availability results for five COST 266 case study topologies: The base topology (BT), core topology (CT), large topology (LT), ring topology (RT) and the triangular topology (TT). The results are for year 2002 traffic matrix estimation.

**Table 5.2 – Worst (W) and mean (M) unavailabilities / No protection**

Topology		BT	CT	LT	RT	TT
PWN (W) [10 <sup>-4</sup> ]	<i>A</i>	114.70	62.50	131.90	115.20	106.20
	<i>S</i>	126.50	61.20	135.80	116.50	113.00
ASWN (W) [10 <sup>-4</sup> ]	<i>A</i>	115.30	63.00	132.50	115.80	106.60
	<i>S</i>	115.00	63.50	124.70	124.00	116.40
PWN (M) [10 <sup>-4</sup> ]	<i>A</i>	47.62	29.92	52.35	51.51	43.12
	<i>S</i>	48.18	30.53	52.40	51.06	43.72
ASWN (M) [10 <sup>-4</sup> ]	<i>A</i>	47.90	30.16	52.65	51.81	43.33
	<i>S</i>	47.93	30.86	51.50	51.74	44.34

Table 5.3 – Worst (W) and mean (M) unavailabilities / 1+1 protection

Topology		BT	CT	LT	RT	TT
PWN (W) [10 <sup>-4</sup> ]	A	6.074	5.552	-	7.001	5.840
	S	9.283	5.600	9.745	9.034	9.241
ASWN (W) [10 <sup>-4</sup> ]	A	6.180	5.681		7.206	5.851
	S	7.113	5.711	9.103	8.609	8.761
PWN (M) [10 <sup>-4</sup> ]	A	5.040	4.853	-	5.210	4.924
	S	5.054	4.854	5.086	5.214	4.959
ASWN (M) [10 <sup>-4</sup> ]	A	5.056	4.860	-	5.228	4.936
	S	5.013	4.847	5.089	5.189	4.933

Table 5.4 - Worst (W) and mean (M) unavailabilities / Path restoration

Topology		BT	CT	LT	RT	TT
ASWN (W) [10 <sup>-4</sup> ]	A'	6.259	5.799	-	7.206	5.968
	S	11.964	7.161	11.041	13.811	9.160
ASWN (M) [10 <sup>-4</sup> ]	A'	5.126	4.989	-	5.335	5.047
	S	5.418	5.107	5.394	5.522	5.101

For the large topology, in the case of 1+1 protection and Path restoration, the analytical procedure was not able to produce results in an acceptable amount of time due to high complexity in terms of both space and time. For the rest of the cases, good agreement between the results obtained analytically and those obtained by simulation was achieved.

Obviously, the CT, being the smallest topology in terms of nodes and links, performs best in terms of unavailability. The difference is most obvious in the case of network without protection.

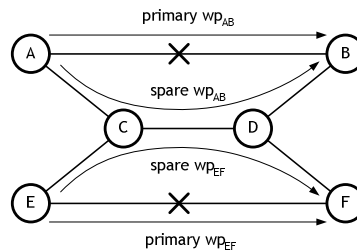


Figure 5.5.

In the case of Path restoration, analytical calculation also provides somewhat erroneous results (Table 5.4, A'). Due to the nature of the analytical model, it was not possible to take into account the cases of already reserved spare resources, and thus unavailable for restoration. For example, let us consider two communication links, one between nodes A and B, and the other between

nodes E and F (Figure 5.5). The spare wavelength paths for these links share the resource. It is clear that if both wavelength paths are affected by multiple failures in the network (i.e. failure of links A-B and E-F), it will not be possible to restore both links.

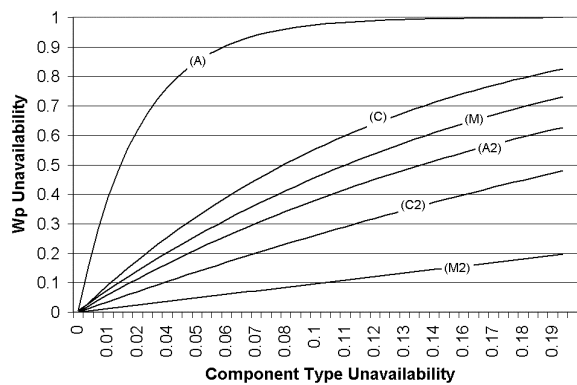
Clearly, a margin of such an error in the analytical model will be lower as the components availabilities get higher. Consequently, a probability of multiple failures will be lower.

The rest of this section deals with the influence of component availabilities on the availability of different transport entities. The goal of the study was to determine criticality of some components with respect to the unavailability of the transport entities, and to draw conclusions on what should be important in the design process. For this reason, two links were extracted from the CT, the shortest one in the topology connecting Strasbourg and Zurich and the longest one connecting London and Zagreb.

**Table 5.5. – Considered wavelength paths**

Wp	Description	Length	Initial unavailability [ $10^{-4}$ ]
1	LondonIP - ZagrebIP	2620	60.574
2	StrasbourgIP - ZurichIP	238	9.200

Figure 5.6 shows the influence of the component type unavailabilities on unavailability of the wavelength path. The components include amplifiers (A), cables (C) and multiplexers (M). Curves (A), (C) and (M) are given for the longest wavelength path (London-Zagreb), while curves (A2), (C2) and (M2) correspond to the shortest wavelength path (Strasbourg-Zurich).



**Figure 5.6. - Wp unavailability vs. component type unavailabilities**

From the Figure 5.6 it is clear that the amplifiers play a major role in determining unavailability of the photonic network. The influence on the unavailability becomes higher as the length of the communication links increases.

Another confirmation of the statement presented above comes from the Figure 5.7. The figure presents the number of components in terms of percentage for both considered wavelength paths. Note that the cable is counted in terms of quantity, not kilometers.

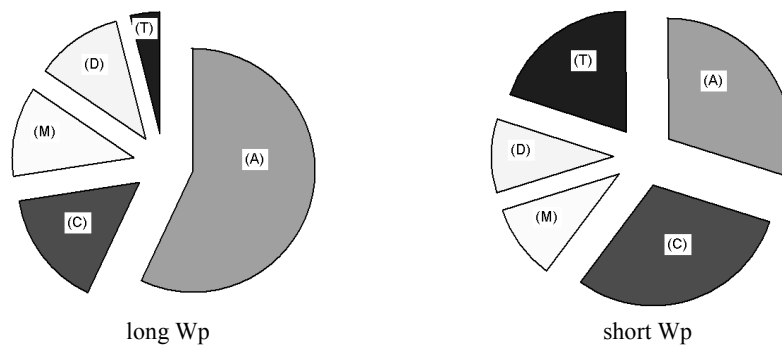


Figure 5.7. - Component count percentages

Figure 5.8 shows the unavailability of a logical channel as a function of amplifier (A), cable (C) and multiplexer (M) unavailabilities. Two cases are presented. On the left side of the figure the no protection case is presented. In this case the logical channel is composed of two wavelength paths, one per each direction. On the right side of the figure, the curve for the 1+1 protection is presented. In this case the logical channel is composed of 4 wavelength paths, a primary and a spare for each direction.

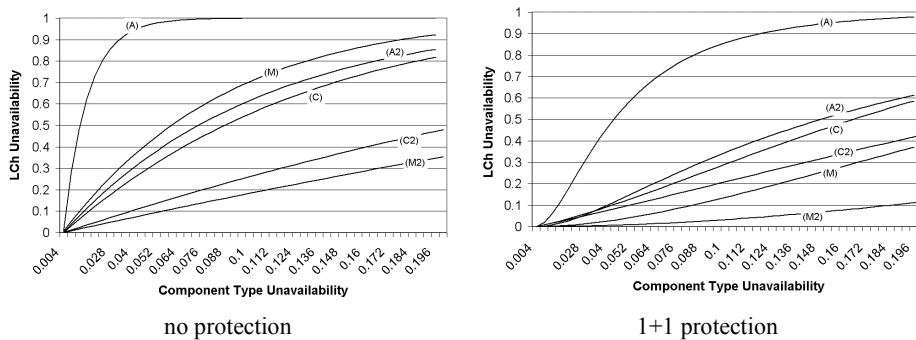
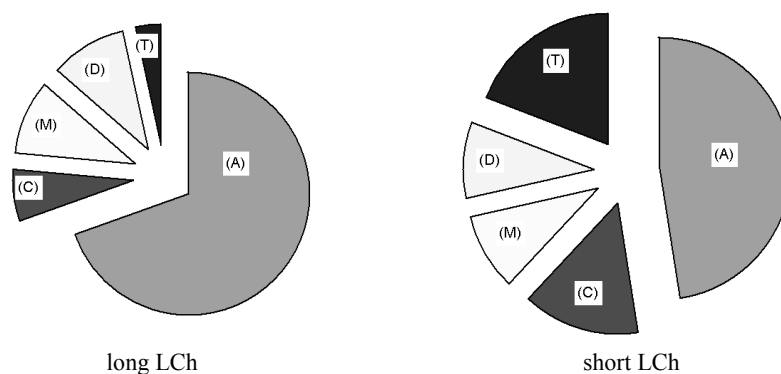


Figure 5.8. - LCh unavailability vs. component type unavailabilities



The figures show that the influence of the amplifiers on the availability of the logical channel can be reduced with the application of protection. It also shows that for the no protection case, unavailability of multiplexer influences highly unavailability of logical channel.

In the figure below (Figure 5.9), component count percentages for unprotected long and short logical channels are given. As expected, the amplifier is still the component with the largest count with (de)multiplexers at the second place. Clearly, the influence of amplifiers on unavailability of transport entity becomes even more evident. The number of cables is reduced for the no protection case since wavelength paths use the same cable for both directions. The logical channel with 1+1 protection keeps similar component count percentages as the unprotected logical channel with a slight increase in the number of amplifiers.



**Figure 5.9. - Component count percentages**

A logical connection (*LC*) can be considered as a series of logical channels, i.e., in the case where all of the logical channels should be functional in order for *LC* to be functional, or as a parallel connection of logical channels. In the second case, *LC* will be functional if at least one of the logical channels is functional. The second case can also be viewed as a protection (or restoration for that matter) applied at the level of logical connection.

We have chosen these two definitions of the *LC* to show the impact of the component sharing among several transport entities (*Wps*). Two *LC* examples have been analyzed in each definition. In the case of no protection the *LC* with 6 *LCh*, and the *LC* with only one *LCh* (denoted with the index (2) in graphs) have been included. In the case of *LC* with one *LCh* the results are the same in the case of a parallel and series definition. This *LC* corresponds to the longest *LCh* from the previous cases (Zagreb-London). The component count is in the

second case larger than in the first case of a *LC* with 6 *LCh* due to short *LCh* lengths. In the case of 1+1 protection the same *LC* with one *LCh* have been kept, but the *LC* with only two *LCh* have been analyzed to reduce the computation time.

All components will have the same impact on the *LC* availability in the case of *LC* defined as an *LCh* series. In the case of *LC* as an *LCh* parallel some components affect only one *LCh*, but some affect all *LChs* in the *LC* parallel thus having larger influence on the *LC* availability. The level of influence depends on the component sharing. The term sharing denotes usage of the same component by several transport entities at the same time and is not to be associated with time disjoint sharing in the 1:*n* restoration scheme.

Cables along the *LC* path are used by all the *LChs* in the no protection case. They will have the largest influence on the unavailability (Figure 5.10). This is visible when analyzing the availability change of one (any) cable. The cable would have the largest influence among all component types. In the case of component type unavailability variation the amplifiers would still be the component with the largest influence due to their large component count, but the cable will be on the second place because of being shared by multiple *LChs*.

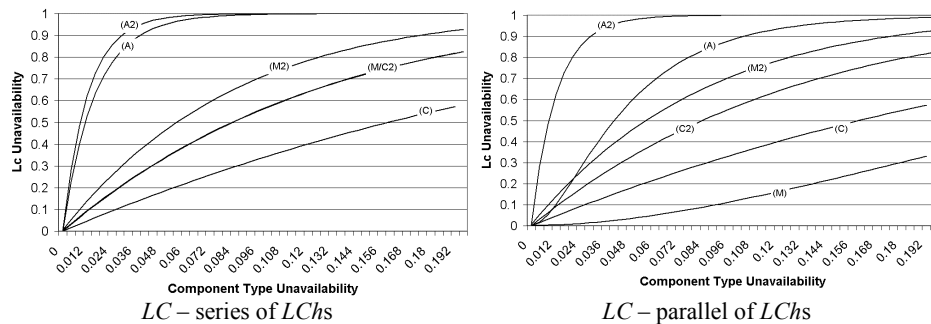


Figure 5.10. - LC unavailability vs. component type unavailability (no protection)

The strong cable influence on the *LC* availability will be also visible in the case of 1+1 protection. The access cable will be shared among all the *LChs*, and include their primary and protection *Wps*. This is the consequence of applying protection scheme only in the transmission network but not in the access network. The access cable is thus unprotected. The transmission network cables in the 1+1 protection scheme will carry only primary or protection *Wps* of the same *LC* and will thus have smaller influence on the *LC* availability than the access cables.

The component share increases the influence of the component on the *LC* unavailability, more than the component count implies. Amplifiers and (de)multiplexers could also be shared among several *LCh*, but this sharing depends on the wavelengths associated to each *Wp* in the *LCh* belonging to the same *LC*. This is a consequence of the RWA, and thus cannot be generalized like in the cable case.

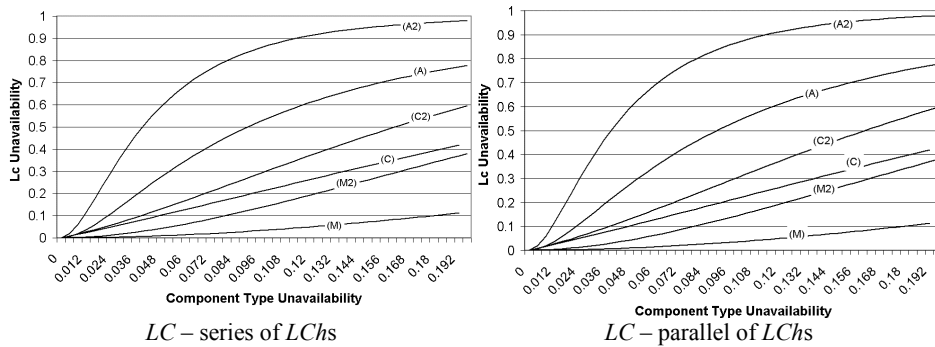


Figure 5.11. - *LC* unavailability vs. component type unavailability (1+1 protection)

Finally, Figure 5.12 shows the network unavailability as the function of component type unavailability.

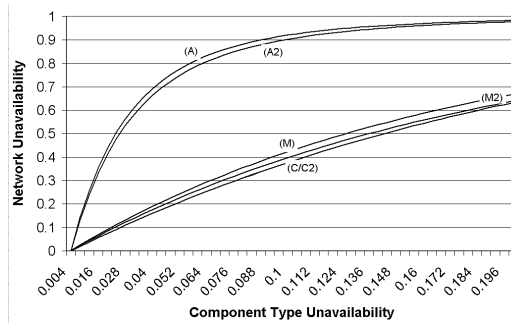


Figure 5.12. - Network unavailability vs. component type unavailability

### 5.3 Comparison of Pan-European Optical Backbone Networks in Terms of Expected Loss of Traffic

#### 5.3.1. Introduction

Where voice was still the dominant traffic source some time ago, data traffic (mainly IP traffic) has recently overtaken voice-generated traffic in volume.

Two things have to be noted here. First of all, the changing relative importance of voice and data traffic (including both transaction data and IP traffic) alters the traffic pattern. While voice traffic is mainly exchanged between locations that are geographically quite close, the exchange of IP traffic is less or not at all related with the distance between source and destination. Secondly, although data traffic has overtaken voice traffic in volume, the main revenue for operators still comes from voice.

The availability of the connections for voice as well as transaction data traffic is very important, both in our personal lives and for business. IP traffic, on the other hand, is generally accepted to transport best effort services. There is thus an opportunity for network operators to save costs by applying different recovery schemes for voice and transaction data on one hand and IP traffic on the other hand. In this section, the comparison is made between (a) the use of 1+1 dedicated optical protection for all traffic and (b) 1+1 protection for voice and transaction data traffic while IP data is routed unprotected. More details can be found in [1] and [2].

### 5.3.2. Case Study

The networks used in the case study presented in this section are the ones introduced in section 1.3. Also the traffic forecast explained there is used for this case study. The Basic reference Topology, the Ring Topology and the Triangular Topology are compared in terms of the integrity of the network design. The Average Expected Loss of Traffic (AELT, average amount of traffic that is expected to be lost each year due to failures) is used to express the availability of the network.

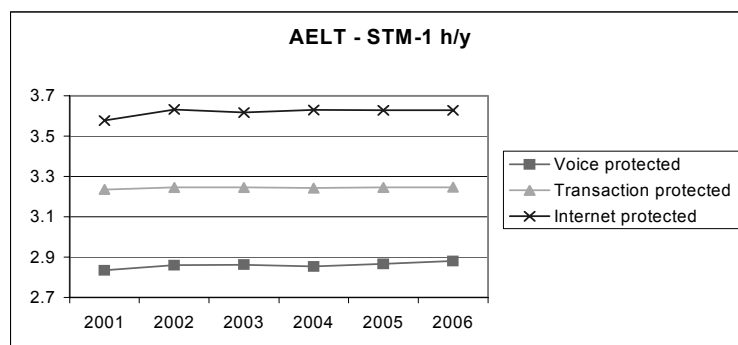
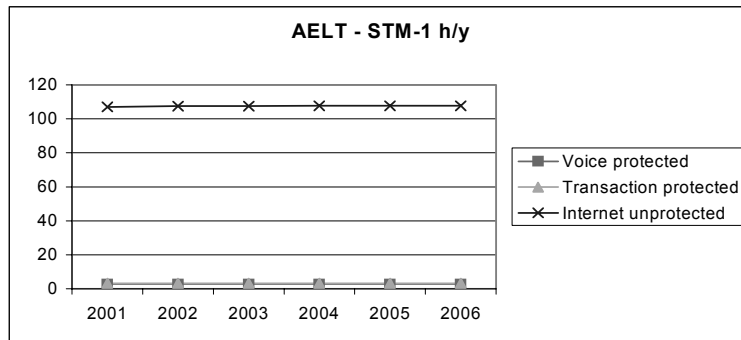


Figure 5.13.a. - AELT of all traffic types for 1+1 protection for all traffic

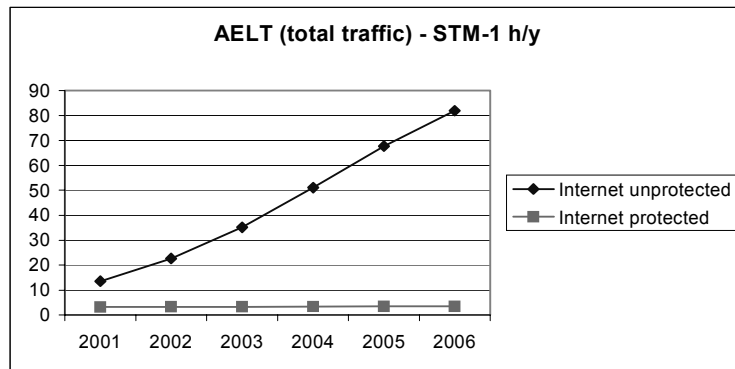
Figure 13 (a) illustrates the relationship between distance and traffic type. Most voice connections are between cities that are geographically close and

thus follow a shorter, more available path, as there is less chance to encounter failing equipment. Transaction data is still distance-dependant, but less than voice traffic. With Internet traffic there is no longer an influence of the geographical distance. These connections will thus on average follow a longer path, which is more prone to failures. Figure 13 (b) shows that the recovery scheme applied has a big influence on the AELT.



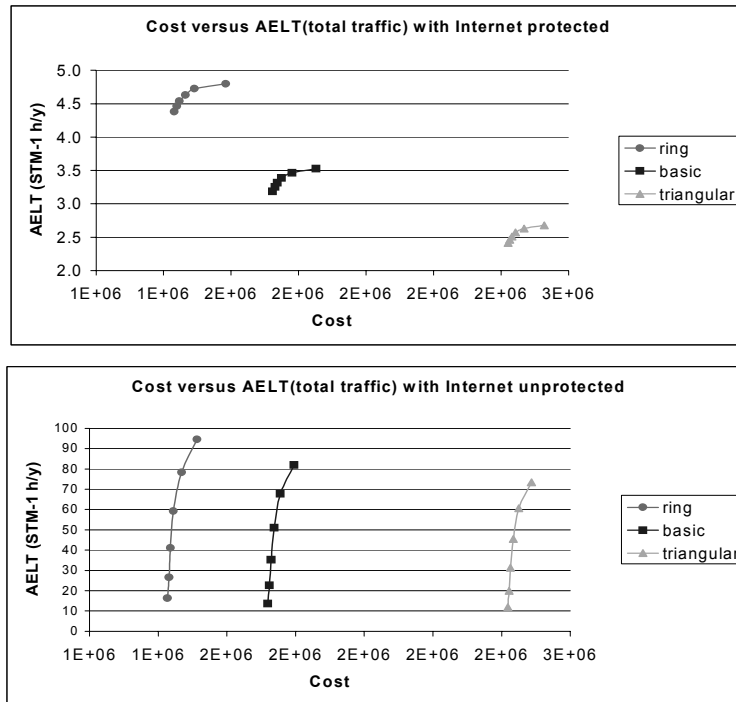
**Figure 5.13.b - AELT of all traffic types for 1+1 protection for voice and transaction data traffic, no protection for IP traffic**

The AELT over all traffic is shown in Figure 5.14. The trend of this graph can again be explained by the distance – traffic type dependency and the fact that Internet traffic grows very fast over time.



**Figure 5.14. - AELT(total traffic) for (a) 1+1 protection for all traffic and (b) 1+1 protection for voice and transaction data traffic, no protection for IP traffic**

Figure 5.15 illustrates the trade-off between cost and AELT for the three topologies over time.



**Figure 5.15. - Cost versus AELT (total traffic) of the basic, ring and triangular network designs**

When looking at the topologies of RT and TT, it is obvious that in the TT, the lightpaths are routed along a path with a limited number of hops. In the RT on the other hand, the lightpaths make a (sometimes quite long) detour. In the latter topology, more fibres will thus be used on each link than in the TT.

The triangular design is the most expensive one. It is much more expensive to dig all ducts required for the TT than for the BT or the RT. As the digging cost is the dominant cost component it defines the order of cost-efficiency of the three considered topologies

The availability of the connections is best in the triangular network and worst in the ring topology independent of the considered traffic type (voice, transaction data or Internet traffic). This can easily be explained. In the ring network, the connections follow on average a longer route and pass more OXCs than in the basic or triangular network. This means that the probability for a failure along the path of the connection (line or node failure) is higher, and thus the availability of this path is lower. In the triangular network, the routes between origin and destination node are the shortest, and these routes have thus the

highest availability. Note that the influence of the topology is even stronger than the influence of the changing traffic pattern over time (different relationship between distance and traffic volume for the three considered traffic types and changing dominant traffic type).

The RT design is the cheapest one, but also results in the highest values for the AELT. The TT is by the way the most expensive design, but gives better values for the AELT. Here a (strategic) decision has to be made whether the higher availability of the connections compensates the extra investment required to reach a lower AELT.

When the IP traffic is transported as unprotected traffic over the network, the difference in AELT between the three studied network topologies is quite small. Although the AELT values obtained with the expensive triangular topology are, for a specific year  $X$ , lower than with the two other considered topologies, one or two years later (year  $X+1$  or  $X+2$ ), they reach the same level as in year  $X$  with the least expensive but “worst availability” ring topology. When all traffic is recovered using 1+1 protection, the AELT tends to a constant value that is considerably lower with the triangular design than with the basic or ring designs. In this situation the extra cost for installing a more densely meshed network may be justified

### **5.3.3. Conclusions**

The changing traffic pattern, characterized by the growing importance of Internet traffic, has a serious influence on the AELT. The recovery scheme used for the IP traffic has a quite big impact. Transporting IP traffic as best-effort service is much cheaper than protecting this type of traffic, but results in a much lower availability of this service than when this traffic would be 1+1 protected. The RT design is the most cost-efficient one. The TT results in the best values for the AELT. The strategic trade-off between network installation cost and availability of the connections has to be made.

## **5.4 Instead of Conclusion – Further Work**

Generic all-optical network architectures that should be dealt with in the future can be classified according to the two criteria. From the topological point of view we distinguish ring, mesh and p-cycle based networks. Furthermore, on top of each of these network architectures, three different switching schemes can be laid - circuit, packet and burst switching. This report focuses on mesh based circuit switched networks, and perhaps future work should investigate

comparative advantages and disadvantages of network architectures by paying more attention to the ring based and p-cycle architectures.

An all-optical network can be viewed as a server layer (or server network) to higher layers (for example IP). Thus, in addition to providing only the raw capacity with or without QoS guarantees, research could also be focused on possible ways for an all-optical network to provide survivability guarantees to higher layers. This field is sometimes called resilience differentiated QoS.

## 5.5 References

- [1] S. De Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz, "Pan-European Optical Transport Networks: an Availability-based Comparison," *Photonic Network Communications*, vol. 5, no. 3, May 2003, pp. 203-225.
- [2] S. De Maesschalck, D. Colle, M. Pickavet, and P. Demeester, "Availability Comparison of Pan-European Optical Backbone Networks," *Proc. on COST 266 Workshop on All-Optical Routing (WAOR) in conjunction with ICTON 2002*, April 2002, Warsaw, Poland.
- [3] R. Inkret, M. Lacković, and B. Mikac, "WDM Network Availability Performance Analysis for the COST 266 Case Study Topologies," *Proc. Optical Network Design & Modelling, ONDM 2003*, vol II., pp. 1201-1220, Feb. 3-5, 2003, Budapest, Hungary.
- [4] M. Lacković and B. Mikac, "Analytical vs. Simulation Approach to Availability Calculation of Circuit Switched Optical Transmission Network," *Proc. of the 7th Int. Conference on Telecommunications, ConTEL 2003*, vol. 2, pp. 743-750, June 2003, Zagreb, Croatia.



## Chapter 6

# Tools

*Marko Lacković, Cristian Bungarzeanu (editors),  
Steinar Bjørnstad, Didier Colle, and Sophie De Maesschalck*

### 6.1 Introduction

Network evolution steps increased the complexity of problem imposed to telecomm engineers. More sophisticated design analysis, and simulation aids were necessary. Today, there is large number of simulators from different sources (academic and commercial) and for various purposes (narrowly-oriented or all-purpose). They are comparable in two key features – capabilities and price - as they serve for different purposes, from education and presentation, to commercial projects.

This chapter describes tools that have been created and used to help in the design and analysis stage of telecommunication network and component construction during the COST 266 project.

### 6.2 The *Cosmos* Tool

#### *6.2.1. Motivation and Idea*

Basic idea was to make a uniform network description, which would allow creation of the network and optimization algorithm libraries. The tool had to

be simple, stable and all-purpose oriented. The last demand was set in order to be able to solve a wide number of problems thus avoiding dealing with several tools. This would attract larger number of users and programmers helping to develop the tool, and upgrade it with their applications. The tool offered the platform with basic functionality being the keystone for future development. Therefore, the first part of the tool was the framework for network description.

### *Purpose*

*Cosmos* is the abbreviation for *COmplex System Modelling, Optimization and Simulation*. It can be used in a variety of telecommunication and engineering design and optimization processes as it can be used in modelling, simulating and optimization processes describing or dealing with diverse systems, from electronic devices to physical effects in the optical fibre. However, the main purpose of *Cosmos* is telecommunication system design and it is primarily focused on optical networks.

Wide and diverse demands were set with respect to various needs of telecommunication. Such general specification makes *Cosmos* adaptable to majority of problems. Uniform model and topology description implies that only one tool has to be learned. This accelerates engineering work and increases its efficiency by eliminating the learning and adaptation time. This is of primary concern as duration and quality of the result are among major concerns of every project.

### *Features*

The *Cosmos* tool has the following characteristics [1]:

- *All-purpose orientation* – tool is not oriented towards any specific issue. The diversity is based on flexibility of topology description and the use of simulation domains. These allow description and simulation of various types of systems,
- *Efficiency* – the tool is efficient in resolving simple and complex problems. Its scalability was one of key requirements during its specification and development, because many existing tools fail to give suitable solution concerning this problem. Efficiency is based on the uniform approach to network design using three structural elements - modules, layers and system,
- *Simplicity of use* – rapid learning and fast accommodation to the tool. Simple usage provides fast application development for problems of minor complexity. Simple problem transfer from mother tongue and/or from some formal specification to *Cosmos* representation is the basis of wide

tool usage. Simplicity is accomplished mainly by encapsulated simulation mechanisms (domains), and network and optimization algorithms in easily-to-use libraries. These libraries can be integrated with the user developed application eliminating the need of simulation mechanism coding. They, however, still have to remain susceptible to user modifications, which can be done by redefining the functions of the simulation mechanism,

- *Code portability* – user code has to be portable between various computer platforms (primarily Windows, Sun Solaris and Linux based systems), as well as different processor architectures (IBM compatible personal computers, SUN workstations). This is accomplished by using freely obtainable C++ compilers for different platforms and keeping the code structure within ANSI C++ specifications. This portability includes just the console part of the tool, restricting the GUI to Windows platform.

#### *C++ and the Object-orientated Structure*

The tool is fully programmed using C++ programming language on Windows based platform. The C++ object-orientation provides enough flexibility for the discussed requirements, and its all-purpose orientation allows all parts of the tool, from simulation mechanisms to graphic interface, to be coded in the same language. C++ widespread usage implies easily obtainable compilers and large programmers community, being essential requirements for wide tool usage.

Inherent similarity between the models implies that adding reusable features or functional parts is more appropriate than implementing the whole solution in one large monolithic model at once. So presented model is easier to develop and verify, because the feature and functionality description can be implemented and tested stepwise [2]. In addition, reusability of the model code shortens similar model development process. These demands can be easily carried out using object-oriented paradigm (OOP). The C++ is not only used for the tool development, but also for model and behaviour definition implying object-oriented approach in modelling and simulation.

Existing systems can be described by their features (e.g. length of the optical fiber) and functionality (e.g. transmission capability of the optical fibre). This property is implemented in OOP encapsulation feature. OOP classes (general templates) and objects (classes' instances) consist of the attributes and functionality (methods), thus giving appropriate basis for model development.

Adding functionality to OOP terms means developing a class hierarchy, starting from more abstract classes (like those representing failure models common to all systems) and then deriving more specialized ones that represent some

system in more details. This mechanism is called inheritance. Each derived (inherited) class contains all functionality and features of the basic class.

Inheritance is the basis of the OOP code reusability because more abstract classes are expected to be used in different model development processes sharing the same basic model (class). Inheritance is accompanied by polymorphism which enables application of the functions developed for base classes (=developed earlier) to all classes inherited from the base class (=developed later).

### **6.2.2. Main Capabilities**

- *Network structure description (topology description)* – supports the user network structure description and model development in terms of topological design. Network description does not imply only telecommunication network description, but a description of any system, as the consequence of all-purpose orientation,
- *Behaviour description* – specifies a simulation model. The user provides a description of network behaviour (function), or network element behaviour (function),
- *Simulation* based on network structure and its model,
- *Optimization* based on optimization and network algorithms and/or simulation results. It can be used in junction with simulation, or analytic procedures.

A more detailed description of these features is given in further sections.

### **6.2.3. Cosmos Tool Structure**

The *Cosmos* tool comprises following parts (structure given in Figure 6.1):

- *Cosmos kernel* is the basic part that provides minimal functionality, necessary for network and model description and simulation.
- *User libraries* contain user provided descriptions of simulation models (like the *On-Off* model used in Monte Carlo availability simulation or parts of WDM network). User libraries make design process easier as the user does not have to provide model description. Code writing, as the most difficult part of design, especially for nonprogrammers, becomes unnecessary.
- *Analytical algorithms* are used for network analysis based on the analytical procedures. The precondition for gathering large number of algorithm and

model libraries is to attract enough users. Many tools fail to accomplish this goal,

- *Nyx tool* is used for defining general purpose heuristic methods, which are applied to defined optimization problems,
- *Cosmos GUI* is essential to achieve user-friendly tool and simplicity of its use. Graphic interface makes topology design easier, but, on the other hand, limits its scope. A unique graphic interface has been developed. Analysis tools make processing of simulation results easier and allow visualization. This could be essential for understanding and making conclusions.

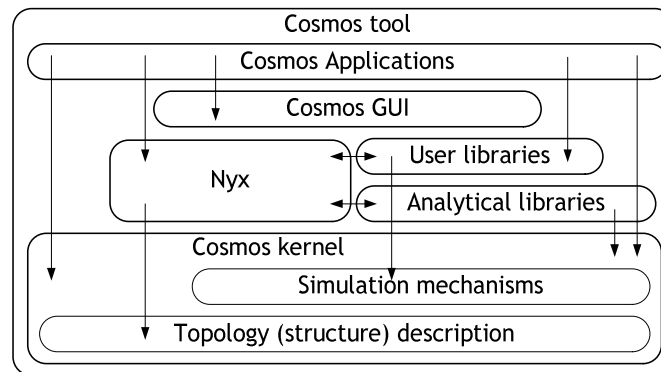


Figure 6.1. - Structure of the *Cosmos* tool

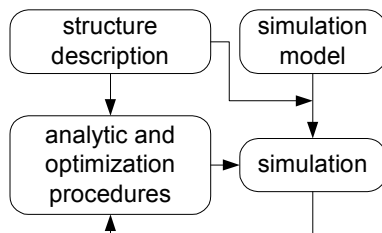
#### 6.2.4. *Cosmos Kernel Structure*

*Cosmos* kernel has two main parts:

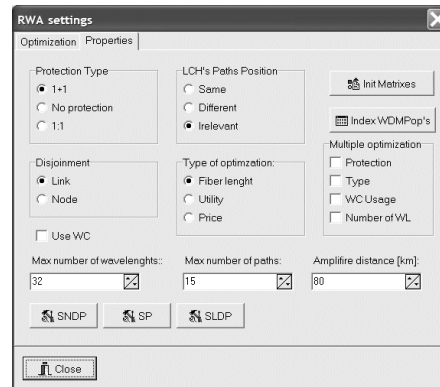
1. Topology (structure), and
2. Behaviour description.

The kernel allows user network description. It provides basic elements of network topology (e. g. nodes in telecommunication network) and the framework for behaviour description.

Topology description defines a passive network structure, while behaviour description defines a simulation model. Topology itself is sufficient for some analysis and optimization processes, but simulation requires both steps. The results of simulation can serve as input data for the analysis and optimization algorithms, (Figure 6.2).



**Figure 6.2. - Steps in network representation in *Cosmos* tool**



**Figure 6.3. - Analytical algorithm parameterization in *Cosmos* GUI**

Analytical and optimization algorithm used in network structuring can be directly parameterized from the *Cosmos* GUI. Figure 6.3 depicts a parameterization form for the RWA algorithm.

Following section shortly describes kernel structures used for topology description. Kernel structures are coded as C++ abstract classes, which are the basis for further development by means of the inheritance mechanism. The inheritance allows the addition of the attributes and methods to the basic set of classes provided by the kernel. Kernel structures are, therefore, the most abstract view brought closer to concrete network elements.

### 6.2.5. *Nyx* Structure

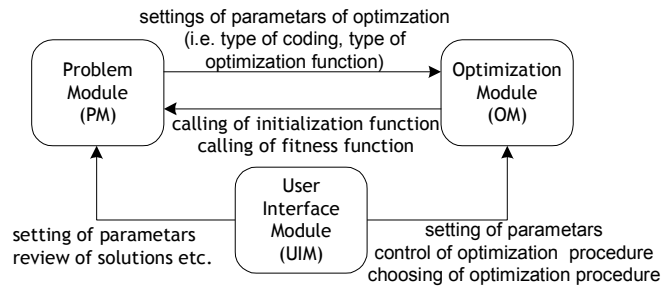
The main motivation for the *Nyx* development was the need for a monolithic software solution that could produce description and solution to a broad range of problems without a large programming effort. We have developed a tool called *Nyx* with such assumptions on the department of telecommunications. Its main features are short developing time and diversity of problems it can be applied to.

*Nyx* optimization kernel comprises three main modules:

- Problem module (PM),
- Optimization module (OM), and
- User Interface Module (UIM).

Figure 6.4 depicts the interaction of the modules. The OM calls fitness function with the possible coded problem solution. The PM decodes solution given

from OM and calculates fitness function implemented in that PM. As a response, PM returns fitness function value of that particular solution.

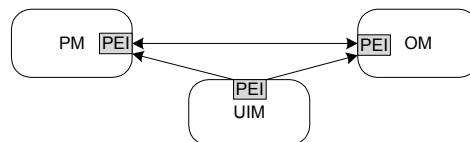


**Figure 6.4. - Communication and basic function of NYX modules**

The UIM contains implementations of user interface which provide function for controlling optimization procedure.

*Communication between Nyx modules*

Optimization procedure depends of many parameters. Nyx allows static and dynamic parameters. Most of the parameters are encapsulated inside module. A parameter is defined by name, type and permissions and must be registered. Each module has an interface – *Property Exchange Interface (PEI)* in order to make the parameter settings possible (Figure 6.5).



**Figure 6.5. - Communication between modules is accomplished with PEI**

Interface PEI contains dynamic list of parameters and methods for setting the parameters.

The UIM module enables setting and reading parameters of the PM and OM. When accessing some parameter, UIM gets a dynamic list of all parameters using a PEI and finds the required parameter in this list. Then it calls appropriate method in PEI to access that parameter. UIM is implemented in the command line shell.

A problem module is the user defined module containing the knowledge about problem being optimized. The main function of this module is to implement fitness function for a specific problem. OM communicates with the PM by calling the fitness function, which has a coded possible solution as the only

input parameter. Each solution has to be encoded, and represented as an array of objects where only PM knows meaning of code. The coded solution has no meaning for the OM. It just connects coded solution with its fitness function and generates next possible solution using that information. A user has to define the type of coding solution. He can use some of included codings like binary, stream of chars and stream of number, or he can build his own type of coding suitable for his problem.

Nyx optimization kernel was implemented as a set of optimization modules and user defined problems, and includes a catalogue of implemented problems. Each new problem has to be register in this catalogue.

Optimization module (OM) contains implementation of various general optimization methods independent of the problem type. Each method is correlated to some type of coding (a solution representation) and uses particular methods. User registers his PM after creating a problem module and choosing type of coding. He is then offered an appropriate set of optimization methods gets a set of optimization methods which can solve the created problem. These optimizations methods are also registered in the catalogue using the type of coding as a key. That concept is shown on the Figure 6.6.

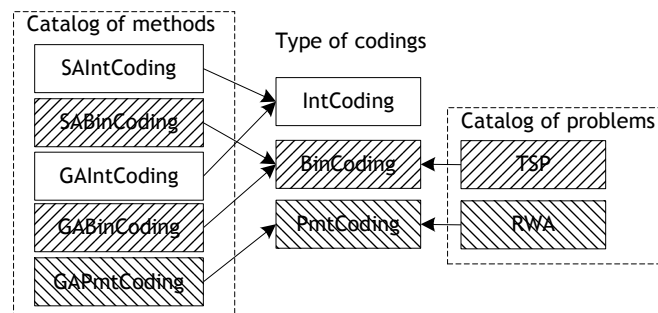


Figure 6.6. - Relationship between problems and methods catalogue

All methods inside optimization kernel must be registered in the catalogue of methods.

Type of coding defines a presentation of some optimization problem suitable for various optimization methods. It is important to choose appropriate type of coding which will adequately describe a solution of stated problem. After choosing the coding type and creating a problem module, the user can choose various optimization methods. Nyx optimization kernel defines three coding types - binary, integer and permutation integer coding.



### 6.2.6. Topology Description

Network topology or system structure consists of list of network elements, and a list of connection between network elements.

Topology is a passive structure (in terms of simulation modelling), not necessarily linked to simulation. It is a part of input data to simulation mechanism (Figure 6.2). *Cosmos* allows the user to apply, besides simulation mechanisms, some analytical procedures, such as topology optimization. Optimization and analysis can be based on simulation result too, but this implies model description in addition to topology description.

Network elements, as the basic building block of network topology, comprise:

- *Attributes* – the features of real-world elements they represent (like amplification of an optical amplifier);
- *Ports* - allow connections and communication between the elements, and
- *Behavior* – model description. It is necessary for simulation only.

#### Modules

Modules are the bases of network topology description. They provide basic functionality that has to be supplemented by user code (attributes, methods) in order to get templates (C++ classes) for element creation (C++ objects). Modules are network elements described in *Cosmos*.

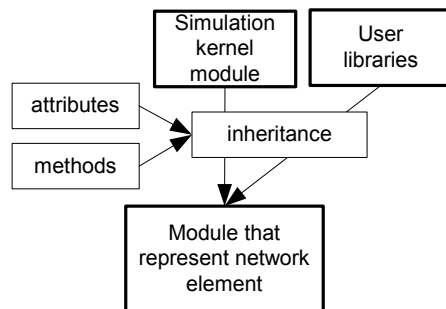


Figure 6.7. - Creation of element representation

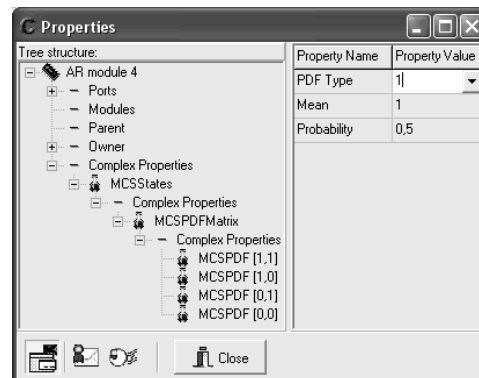
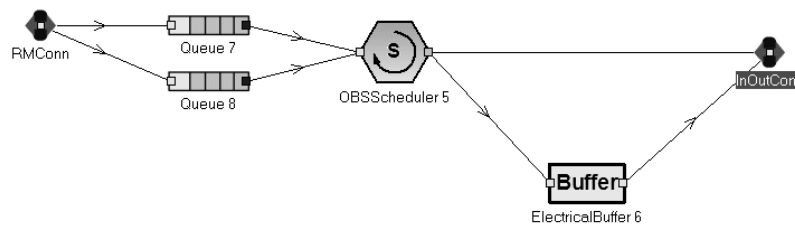


Figure 6.8. - Module attributes in *Cosmos* GUI

There are two ways of deriving the modules that represent a specific network element. The first one uses standard C++ mechanism of inheritance and the other one uses the modules from user libraries (Figure 6.7). Inheritance can be applied to the latter.

Each module contains a number of attributes, which can be included and accessed from the *Cosmos* GUI (Figure 6.8).

Modules can be nested in arbitrary depth. A module can contain arbitrary number of modules (Figure 6.9).



**Figure 6.9. - An internal module structure (burst switching application)**

#### *Ports and logical connections*

Connections between modules are created between ports. These connections are called logical as they don't implement any physical feature. Ports can be unidirectional (in, out) or bi-directional (in-out). They cannot have behaviour description and serve just as the support in topology creation.

Logical connections link two ports and edge port directions determine their direction. Several connections can be linked to one port. Connections along with the list of elements make topology description and serve as input data for network and optimization algorithms. They are essential for message exchange mechanism, which is a key for telecommunication network simulation.

Logical connections do not implement any physical features and are not network elements either. They exist just in the list of connected elements contained in each module comprising at least one port.

Physical connections have to be created as new modules. These modules are connected to edge modules communicating through the connection with physical features. Figure 6.10 depicts the general idea of implementing the physical connection features inside a new module.

Logical connections can exist only between modules that belong to the same module, or the modules that are not a part of any module (that are directly added to the system). A module aggregating other modules (the use of the module nesting mechanism) contains ports that provide an interface to the other world. This interface allows nested modules to be connected to modules on the outside of the parent module as shown in Figure 6.11. These ports are the same as any other port, but they allow the connections from the inside and from the outside of the module.

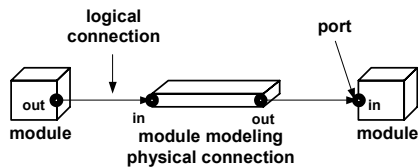


Figure 6.10. - Logical and physical connections

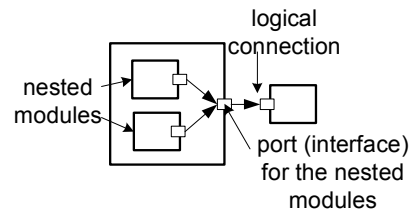


Figure 6.11. - Module nesting and ports

### Layers

Layers serve as logical aggregations for modules. They do not change the modules in any way, and do not add anything to their functionality. The modules that have something in common can be organized in one set or one layer. In telecommunication networks design the layers are mostly used to represent the same topology with regard to different OSI layers.

Network and optimization algorithms can be applied to the modules of one layer due to different topology of the same network (a mesh in TCP layer and rings in SDH layer).

One module can be the part of several layers. The layers can occur within the modules because they can aggregate the modules inside some other module. Layers cannot have behaviour description.

### System

A system is the highest part of topological hierarchy. After defining topology, and aggregating the modules in layers (this step is optional), the user adds the modules and layers to the system. The system itself does not change module features and in this aspect it is comparable with a layer. However, besides logical aggregation, the system implies simulation mechanisms that give the meaning to behaviour description. The system implements some simulation mechanisms, as described in section 6.9.

Figure 6.12 depicts system properties form in the *Cosmos* GUI (similar as in the case of module parameters), while Figure 6.13 depicts system parameters regarding simulation.

Modules, layers and systems are three main parts of topology description (Figure 6.14). The ports are used in connection creation. Logical connections between the modules are not implemented as objects.

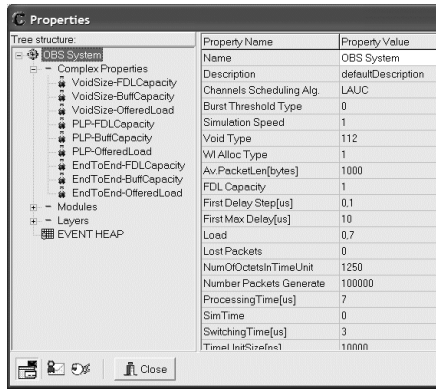


Figure 6.12. - System properties

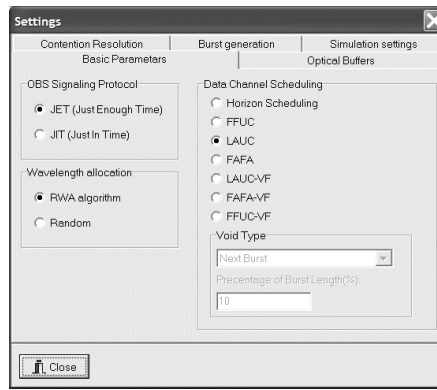
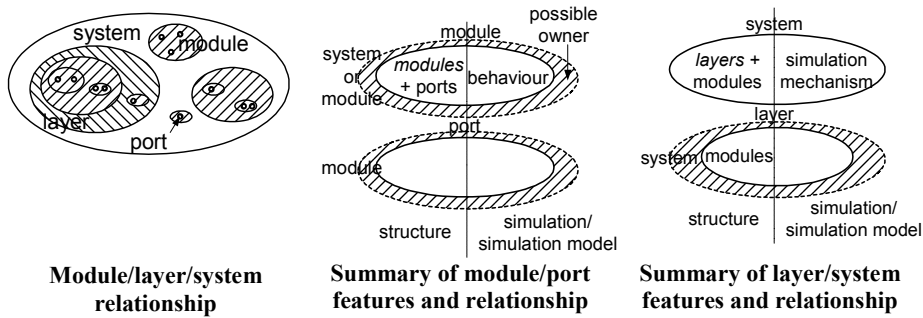


Figure 6.13. - System simulation properties



Module/layer/system relationship

Summary of module/port features and relationship

Summary of layer/system features and relationship

Figure 6.14. - Modules, layers and system

### 6.2.7. Behaviour Description

During the previous step the user defined network structure sufficient for some analytical calculations (Figure 6.2). Simulation requires a simulation model representing module behaviour, which is more or less simplified behaviour of a real-world element.

#### Module behaviour

C++ is essential to define a module behaviour. This makes this stage difficult for users unaccustomed to C++. However, they can still use the modules from module libraries to eliminate this phase, and create only topology description.

Module behaviour can be divided in two sets with respect to the execution time:

1. Initialization, and

## 2. Run.

Initialization code is called before the start of simulation. It sets modules (their attributes) to some initial values (in availability simulation this can be *On* or *Off* state). The system initialization method is called before the simulation run. The system initialization method runs initialization methods of all modules the system comprises, and each module calls the initialization method of all modules it contains. Initialization method of every module has to be called. The user has to provide code for initialization methods.

Simulation mechanism during simulation execution (determined by simulation domain) calls different run methods belonging to different modules. The selection and order of run calls depend either on a user provided mechanism, some condition, or an event expiration in the discrete event domain. The code within the run method depends on the chosen simulation domain, and on the module actions and interactions.

### **6.2.8. Communication Mechanisms**

Network behaviour is based on behaviours of the modules it contains, as well as on module communications. Communication process can be accomplished in three ways:

1. *Direct method calls* are the most explicit way of communication. One module directly calls a method (user defined action on a module) of some other module;
2. *Events* are the characteristics of discrete event simulation. This way of communication is limited to the given simulation domain, and
3. *Message exchange* is the most used communication mechanism in telecommunication network simulation.

Direct method calls are the simplest way of communicating where one module interacts with another by calling its method. This approach is not suitable for network simulation because of the distributed nature of network structure.

#### *Message Exchange*

The message exchange (ME) mechanism is implemented in the message exchange domain (implemented in a separate class), and is independent of chosen simulation domain (mechanism). The ME domain is usually associated with the discrete event simulation domain, as suits the telecommunication network simulation the best. In cases where discrete event domain doesn't describe the modelled system well, like in the case of modelling telecommunication network at the signal level, the ME domain can be used as a communi-

cation mechanism in association with some other simulation domain (like the continuous time simulation domain). This is possible because the ME domain doesn't use any part of the simulation mechanism. Modules are communicating with no correspondence of the simulation mechanism.

The message exchange domain is implemented in the same way as direct method calls, but its key features make it more suitable for network communication. It is limited just to two generic methods (`sendMessage` and `getMessage`). The user provides methods' code for each module.

Message exchange domain is the only one that has to be directly implemented in a module (get and send message capability). These features are contained in separate class (`meDomain`), and they are added to any module class (i.e. ancestor of the base module class `cModule`) using the inheritance mechanism (Figure 6.15).

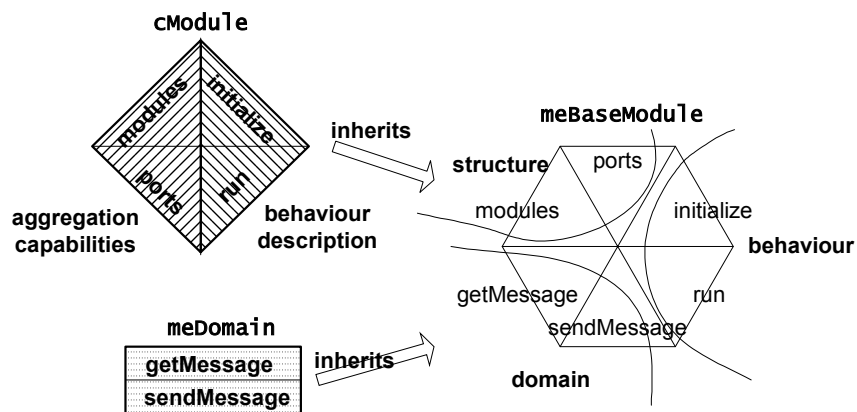


Figure 6.15. - Integration of the message exchange mechanism with topological and behaviour support

### 6.2.9. Simulation Mechanisms

Simulation mechanisms are the core of simulation. They run user provided behaviour description and give some meaning to the code.

Simulation domain implements a simulation mechanism. Different projects require different simulation domains that are most efficient in dealing with a particular problem. The user must choose the best simulation domain relevant to the purpose. Discrete event domain fits most simulation demands when it comes to telecommunication network simulation, but electronic devices are described best by differential equations.

### Event Handling

Event handler and event heap implement the discrete event simulation mechanism (Figure 6.16). During initialization modules produce events, which are put on the event heap by their expiration time. Scheduling mechanism takes the head of the event heap (an event with the shortest expiration time) and makes an appropriate action on the module that created the removed event (the event owner). This can be done directly, by calling the run method, or using the event handler, which calls the method associated with simulation domain. In the first case the run method calls another method or performs some action. In the message exchange domain in the latter case, the `sendMessage` method would be called.

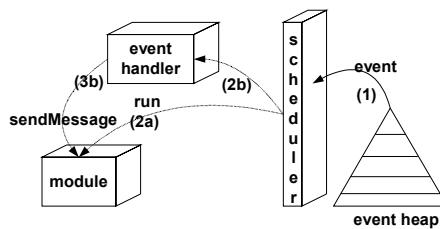


Figure 6.16. - Discrete event simulation principle

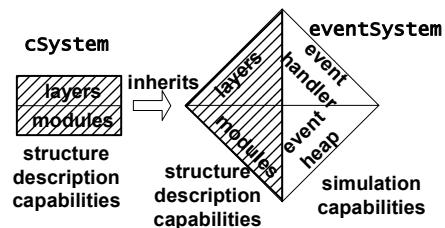


Figure 6.17. - Incorporating a simulation mechanism with the topological and behaviour description

It is important to note that the event should carry information about simulation domain, because simulation mechanism does not imply simulation domain, and vice versa (except in the case of communication via events connected to the event driven simulation). An event can also contain an object used as the input parameter for the function called after event expiration (e.g. object of the `cModule` class that will be passed to the `sendMessage` function when using the event handler in the message exchange domain).

Figure 6.17 depicts a way of integrating the topological and behaviour description with the simulation mechanism. An inheritance has been used and the idea is the same as with integrating the module topological and behaviour description with the communication domain (Figure 6.15). This approach produces great flexibility as the simulation mechanism (like the communication domain) can be integrated with a user developed system code.

Figure 6.18 shows a discrete event simulation properties including heap size, animation settings (for the message exchange domain), and simulation controls (run, pause, stop, ...). Figure 6.19 shows the support for graph and histogram analysis after simulation (or analytical algorithm) execution.

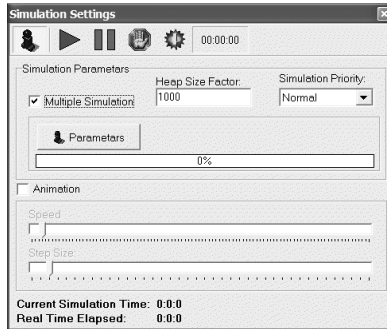


Figure 6.18. - Discrete event simulation properties

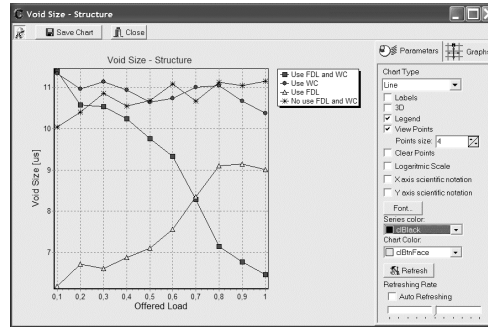


Figure 6.19. - Simulation result analysis

### 6.2.10. Examples

#### *Availability and reliability simulation (Graphical)*

Monte Carlo simulations based on discrete event domain. Dijkstra's shortest path algorithm is used to determine network availability and reliability between two chosen nodes. Availability and reliability modules implemented as Markov process.

#### *Modelling, optimization and analysis of the WDM all-optical network based on configurable cross-connects and circuit switching (Console)*

Input data comprise WDM network topology, traffic demands between nodes (PD model) and P&R scheme [3]. The goal is to create a WDM network (node structuring + link dimensioning) that supports traffic demands and chosen P&R mechanism. Design includes or excludes wavelength converters. BER analysis based on power levels, ASE and crosstalk noise addition.

#### *Protection and restoration schemes in the automatic circuit switched optical network (Console) [4]*

Supported P&R mechanisms: 1+1, 1:1, Span and SPAF. Monte Carlo based network availability calculation.

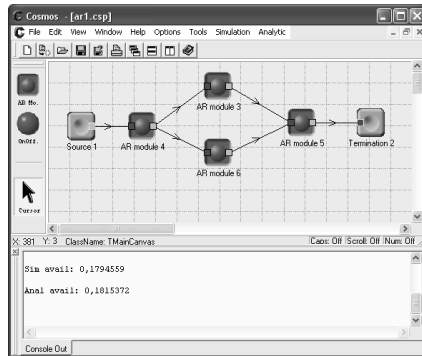
#### *Gigabit Ethernet network modelling, simulation and analysis (Graphical)*

GbE mechanism implementation, with performance analysis (throughput as a function of input traffic). Two network components – Ethernet stations and GbE hub.

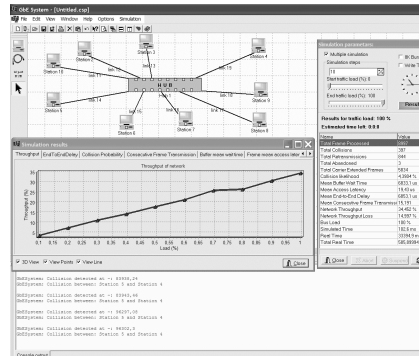


*Broadband IP-router network modelling and analysis (Console)*

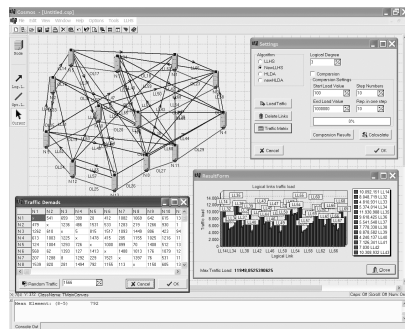
Implementation of several router architectures regarding switching fabric type (bus, crossbar). Throughput and PLR analysis.



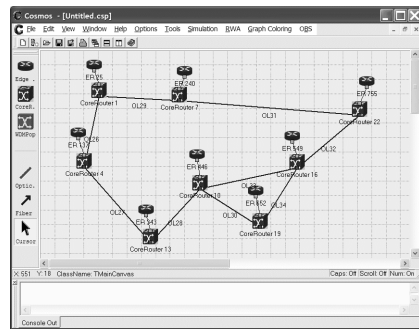
**Reliability and availability**



**Gigabit Ethernet**



**Logical topologies**



**Burst switching**

**Figure 6.20. - Examples of applications in Cosmos**

*Traffic modelling (Console)*

Tradition traffic types, self-similar traffic (Pareto aggregation and FGN).

*TCP (Graphical)*

TCP layer modelling, educational purpose.

*Logical topology optimization (Graphical)*

Creating a logical topology using the physical topology and traffic demands as input data. Application of optimization algorithms on the logical topology. Link dimensioning.

### *Burst Switching (Graphical)*

Modelling, optimization and simulation of the optical burst switched network. Figure 6.20 depicts graphical interface of several applications in *Cosmos*.

## **6.3 Computer Aided Network Planning Cockpit CANPC**

### **6.3.1. Introduction**

Network planning tools are essential for designing multi-service networks that can meet today's growth in user applications and traffic demands. These networks have to provide guaranteed service quality and must be optimized according to various cost constraints. Since the complexity is also growing rapidly, planning tools have to be powerful and enough flexible to handle a large variety of design issues. Enhanced extension facilities are also required to minimize the time necessary for the integration of new algorithms and methods. The first version of this planning framework was using a sophisticated architecture, but the time and effort required to integrate new applications was a serious obstacle for further evolution. For that reason, a new version has been developed, keeping in mind both requirements of power and flexibility. The main goal was to deliver a user friendly framework that allows the development and integration of new applications in a simple and fast way, to keep pace with technology evolution.

### **6.3.2. The CANPC Framework**

The CANPC framework [5][6] provides the main user interface and acts as a nexus for the network editor, the applications and the extension manager. It also provides the control of all inputs and outputs to the main program (opening and saving network models, descriptions of new items, extensions, etc.). The programming employed in CANPC is Python [7], which being a high level interpreted language allows a fast development of algorithms and applications.

#### *The network editor*

Within the CANPC network editor, nodes, links and traffic demands can be placed graphically. For larger networks, scripts can be used to automate the placement of nodes and links, the set-up of traffic demands and the configuration of all network elements. The network may be structured at different levels, each node representing a sub-network. Figure 6.21 shows the graphical interface of CANPC network editor.

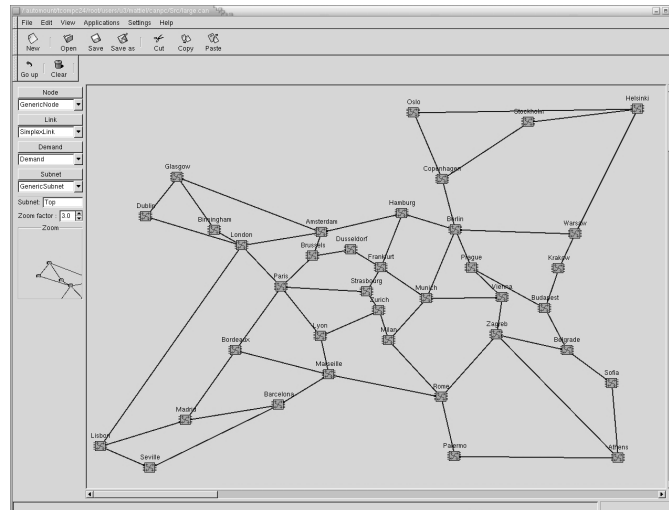


Figure 6.21. - CANPC network editor

### *The network model*

The network model contains all input parameters that are taken into account by applications and algorithms to perform evaluations. To construct the network model, the user has to follow several steps within the network editor. The main steps of editing the network model are:

- *Placing and interconnecting the network objects.* It can be done graphically by selecting the object types and placing them on the canvas. When the number of objects becomes too large, the scripting application offers a way to simplify the editing procedure. Thus, the user can write a simple script to specify node the locations and the way to interconnect them. The traffic demands between the nodes can also be defined in the script. The script description can define easily complex networks and offers a efficient way to exchange network descriptions between various planning tools (e.g. *CANPC* and *Cosmos*).
- *Specifying the object properties.* Initially, network objects contain only few data elements, like the assigned icon, coordinates, name and type. To attach the information required by applications sets of specific attributes called protocols have to be created. A protocol can be viewed as a variable container that is inserted into the network objects to increase the amount of configuration parameters. A protocol editor allows the user to select the protocols, edit the parameter values and assign them to the nodes, links, traffic demands or subnets. It is important not to associate the concept of

protocol object to communications protocols (IP, RSVP, etc.), since the first one is just a simple attribute container. Once the network defined and configured, parameter values will be extracted by exploring directly the appropriate protocols which were inserted into the network objects. This provides the needed flexibility, as, for example, a link can be considered at the physical level, containing a parameter such as the fibre chromatic dispersion, or to the network level with a traffic load associated to it. In this case, two different applications may consider the same link, but the relevant parameters for each application are stored into two different protocols. Therefore, the evaluation can be performed at different levels without any interference. Since the results issued by one application may be used as input data for another one, the system will put these results into containers which are inserted into network objects. Thus, other applications can extract and use them later. This is the case, for example, for the routing application which provides a path for each pair of traffic demands. The results of this application consist in a set of “route” objects, which are inserted in all traffic demand objects. If another application requires the routing information, it can retrieve it by the exploration of “route” objects. The structure of a network object is represented in Figure 6.22.

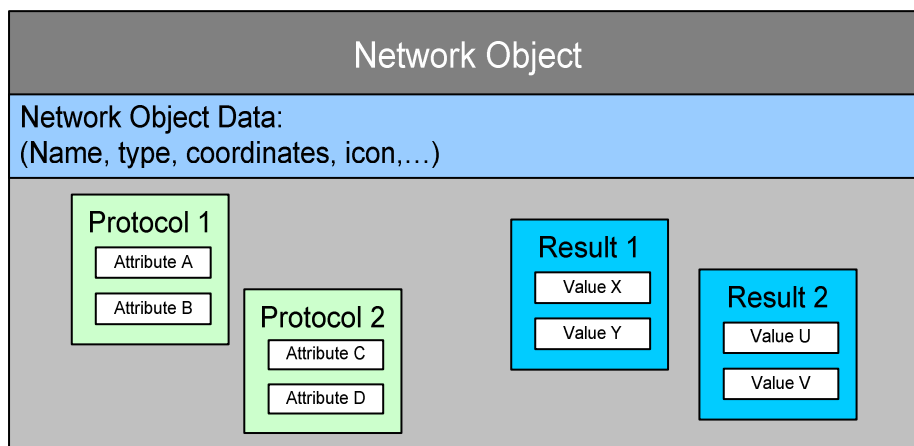


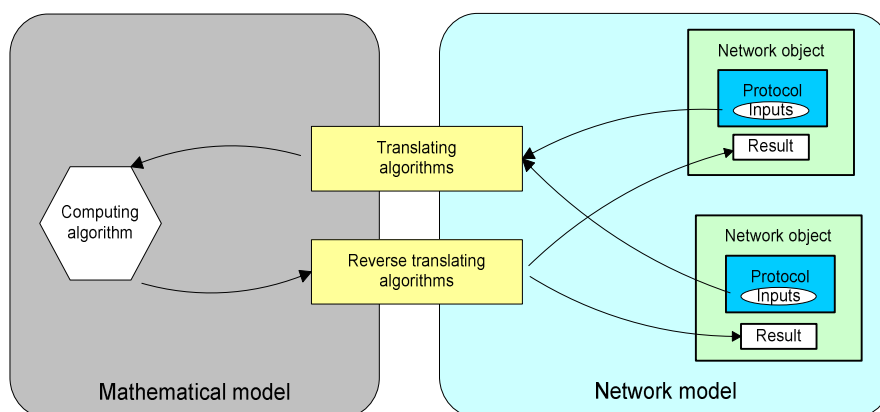
Figure 6.22. - Internal structure of a network object

### 6.3.3. Applications and Algorithms

Once the network defined and specific parameter values assigned, the evaluation may start. To perform an evaluation, the appropriate application is selected from the menu “applications”. An interface window containing all the control options for the chosen application will appear. Some basic applications

are already implemented into the framework. However, the main goal was to make it very accessible to developers, offering them a maximum of facilities to build and integrate new user applications. To structure the process of development and make possible the reutilization of common routines by different applications, computations are split into algorithms that are called by applications to perform a well-defined single task. There are two different kinds of algorithms, depending on the task performed. Thus, the translating algorithms collect all the input parameters from the protocols of network objects (directly from the network model) and transform this data into mathematical objects (matrices, graphs, etc.) that can be used by computing algorithms. The computing algorithms perform the appropriate computations and their inputs and outputs are only mathematical objects.

An example of computation algorithm could be the Dijkstra's routing algorithm. It calculates the shortest path between two nodes in a network, given the incidence matrix which specifies how the nodes are connected. To reduce the computation time, these algorithms can be implemented using numerical libraries or C compiled libraries, which are considerably faster than the interpreted Python. After computation, the inverse translating algorithms will take the results and insert them into the corresponding network objects, to be displayed by the network editor and made accessible to the user. Figure 6.23 shows the different algorithm types in the application structure.



**Figure 6.23. - Application structure**

This architecture proved to be very efficient in terms of reutilization and applications can be developed successfully in relatively short times. Thus, the application library grows permanently by adding new applications in different fields. Below are briefly described the main applications that are available today in CANPC.

### *Network Editing Applications*

- *Network generator*: generates hierarchical Internet networks automatically from some user-defined parameters (number of LANs, number of nodes per sub-network, etc.).
- *Scripting application*: very useful when the amount of network elements is considerable. It consists in a script editor and the functionality to execute the scripts. Allows the automatic placement of network elements, traffic description, specification of various parameters and other useful features.
- *Link editor*: allows the interactive placement of links by means of a user-friendly graphical interface. Facilitates the specification of protocols and the input of parameters values.
- *Traffic editor*: very similar to the Link editor, establishes the traffic demands between the nodes. Protocols and input parameters of these demands can also be specified.

### *Routing applications*

- *Shortest path routing*: paths can be computed automatically (shortest paths) or be defined manually by the user. The objective function to be minimized can be the number of hops or a cost parameter (length) defined by the user.
- *Hierarchical routing*: provides the shortest paths considering sub-networks at different levels (LANs, MANs and WANs).
- *Routing with load balancing*: in this case, the routing is performed not only in function of the length, but taking also into account the total load of each route.

### *Dimensioning of WDM photonic networks*

Most of the dimensioning applications integrated in CANPC deal with the WDM photonic networks. The most relevant are:

- *Wavelength assignment*. Several routing and wavelength assignment approaches have been programmed and integrated in CANPC. They concern the WDM with optical cross-connects and without wavelength conversions [8][9]. Several assignment algorithms have been tested on various topologies proposed within COST 266. A comparison of different wavelength assignment algorithms is given in [10].
- *Dimensioning of optical packet switched networks*. An analytic dimensioning methodology based on probabilistic computations [11] has been devel-

oped and tested to medium and large topologies. It allows the evaluation of the link capacities necessary to guarantee a ceiling packet loss ratio [12].

- *Dimensioning optical transport network with optical add-drop multiplexers.* For a given topology and traffic matrix, it provides an optimum choice on type and location of ONDMs.
- *Planning a logical topology for IP over WDM networks.* Provides an optimum configuration that minimizes the overall cost. Figure 6.24 shows the user interface of this application.

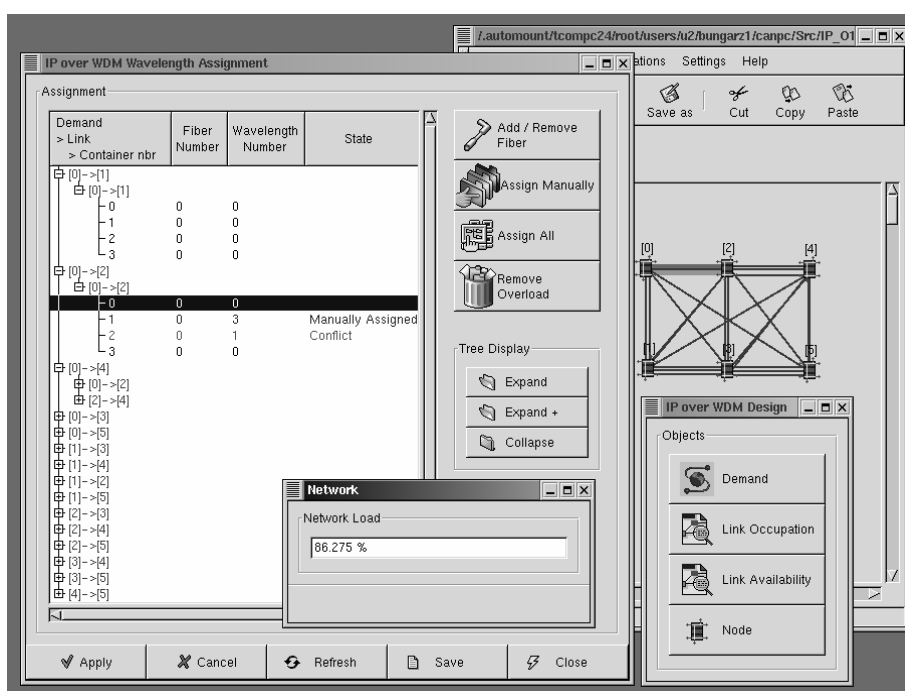


Figure 6.24. - User interface of IP over WDM application

### 6.3.4. The Extensions Manager

The integration of new applications, protocols and algorithms in the CANPC framework is controlled by the extensions manager. It allows the user to extend the set of active applications without any programming. The main procedures related to the extensions manager are described below.

#### *Registering a new application*

All the installed applications are visible in the menu “Applications” of the main window. To register a new application, only the location of the applica-

tion file is needed. This file should contain the main code of the application so that the framework can automatically perform the call with the appropriate parameters. In this way, the links to the network editor and the network model are automatically provided by the framework and the user must only follow a set of simple rules to access the objects.

#### *New protocols registration*

Only a file all the protocol descriptions must be written and registered. Thus, the protocol will be available in the framework to be inserted in the network objects.

#### *Registration of new user network objects*

The user may define new network objects without any programming, as a combination of existing network objects and protocols. To simplify and accelerate the design tasks, a new network object can be interactively created. During the creation process, the protocols to be inserted are selected and the default values for the specific parameters are assigned. It is possible to configure even the appearance of a new network object (icon, color, etc.) can be configured. Once the registration done, the new object is immediately available in the network editor, to be selected and placed.

#### **6.3.5. Conclusion**

CANPC proved to be a versatile platform for network planning and performance analysis applications. The short prototyping time and the ease of integration of new applications allow the users to test several methods and algorithms to find the most efficient way to address a specific problem. Student projects showed that the time to learn Python programming language is acceptable. However, people familiar with object oriented programming (like C++) can learn Python in a couple of days. Thus, the library of applications associated to CANPC is growing at fast pace.

### **6.4 Other Tools**

The other example of discrete event tools beside *Cosmos* is a general simulation library for the programming language Simula, called “DEMOS” [13]. The simulation program itself is made especially for simulating OPS-switches with a strictly non-blocking switching matrix and a limited number of buffer inputs.



### 6.4.1. WDM Guru

OPNET WDM Guru [14] (formerly known as WDMNetDesign) is an advanced multi-layer network planning solution that enables to design resilient, cost-effective optical networks.

#### *Introduction*

WDM Guru enables network designers to create robust and cost-effective Wavelength Division Multiplexing optical networks. Guru's extensive, built-in network expertise provides powerful capabilities for routing, grooming and dimensioning networks to meet current and future traffic demands. Users can quickly create and test different "what-if" scenarios with varying topologies, traffic matrices and configurations. Network managers can use Guru's extensive reporting features to quickly compare the results of differing scenarios, and thereby determine the most effective (and least costly) network designs to meet future demands. This section provides a brief introduction to many of these features and an overview of the design process.

#### *Product Features*

Key features of WDM Guru include:

- **Layered network models** – Guru creates four-layered network models compliant with ITU and ANSI recommendations. This allows the user to explicitly model his network at four different layers: physical (buildings and cables), multiplexing (equipped fibres), optical-wavelength channelling and routing, and digital client (SONET/SDH traffic and routing).
- **Optical network architectures** – WDM Guru includes two different modes for modelling different architectures:
  - Opaque mode, for optical networks in which the OXC is surrounded by long reach transponders interconnected to the WDM line systems. In these configurations, regeneration occurs at every intermediate node.
  - Transparent mode, for optical networks with a transparent OXC. The OXC can be interconnected to a bank of transponders to overcome transmission limitations for wavelengths transiting through the node.
- **Equipment models** – Guru allows the user to generically model different types of WDM line systems and node equipment at the different layers. Different cost models can be applied to line systems or node equipment so the optimization algorithms select the most appropriate network equipment.
- **Topology design and traffic modelling** – With Guru's graphical interface the user can easily create and modify his network designs. Guru also calculates

topology metrics such as network connectivity and node connectivity. The Traffic Matrix Editor makes it easy to create varying traffic matrices for testing different protection strategies.

- **Multiple scenarios per project** – A single project can contain multiple scenarios. Each scenario contains a network that is independent of the other scenarios in the project. The user can extract and compare network metrics (like network cost and capacity) across multiple scenarios.
- **SONET ring design** – WDM Guru supports SONET/SDH ring architectures. As part of the DCL infrastructure the user can specify SONET rings on top of the physical topology. The user can route DCL traffic on these rings, upgrade the rings by adding stacked rings on top of them, or let the ring design operation determine the stacked rings to deploy.
- **Protection and restoration** – Guru can consider various protection mechanisms at the optical and electrical layers: dedicated path (1+1), shared path ( $N:M$ ), end-to-end path restoration, and link-based restoration. All these mechanisms reserve sufficient spare capacity on the links to protect against link failures.
- **Physical link design** – Guru performs the link design (it places amplification and regeneration huts along a link) based on the transmission characteristics of the different WDM line system types. As a result, the cost of amplifiers and regenerators can be taken into account correctly during network dimensioning.
- **Network dimensioning** – Given a particular traffic demand, WDM Guru dimensions the nodes and links to accommodate the demand. Custom algorithms dimension the network to maximize throughput at the lowest possible cost.
- **Routing** – Guru can perform routing at both the optical-channel and the digital-client layers. The user can apply both unprotected and diverse-path protected routing algorithms. Guru takes capacity limitations resulting from the equipment specified into account during routing.
- **Manual routing** – In addition to routing an entire traffic matrix, the user can also establish connections in the network on a connection-per-connection basis. The user can inspect the candidate routes and metrics and then select the route to use for a particular connection.
- **Transparent routing** – With WDM Guru, the user can design and evaluate networks that have limited transparency reach and selective regeneration. This allows the user to balance the amount of required regeneration and the maximum length of a transparency connection in his network.

- **Grooming** – Guru can dimension a network by performing optimized translation of SONET/SDH DCL traffic at the digital-client layer into optical wavelengths. Thus the user can obtain a cost-effective balance of routing and switching cost over the optical and electrical layers.
- **Availability calculations** – Based on the failure rates specified for the different network elements, and the recovery mechanism applied to a service, Guru can calculate the availability of this service and calculate metrics such as expected loss of traffic.
- **Failure analysis** – The user can also define his own failure scenarios within WDM Guru. First the user selects the network elements that fail (such as cable cuts or OXC failures); then Guru evaluates how these failures affect the traffic routed in the network. Then the user can investigate which connections are affected, lost and recovered.
- **Browsers** – Browsers provide intuitive access to relevant node, link and connection information at the different layers. The tree structure allows the user to expand or collapse branches, showing or hiding detail.
- **Bill of Material (BOM)** – The user can easily generate and view a Bill of Material which includes a comprehensive report on line and node equipment at different areas in the network: tributary cards per bit rate, digital and optical cross-connects (DXCs and OXCs), OADMs, long- and short-reach transponders, in-line OAs and regenerator equipment.

#### *Use in COST action 266*

The availability studies presented in Chapter 5, section 1.3 [15][16] were obtained using this network design and evaluation software tool .

The availability studies from the Ghent group were based on the WDMGuru tool of OPNET. This is also the new name of the WDMNetDesign tool (part of the Comsof tool) [15][16].

## **6.5 Conclusion**

Simulation tools are widely used in the design and analysis of telecommunication system. They are necessary in majority of the projects involving telecommunication networks because of the network complexity and price. Network complexity makes exact analytical methods impossible because of computer limitations and considerably short time required for the projects to give the results. On the other hand, network complexity and size are very cost-

demanding, making faulty design implementation and/or prototypes unaffordable.

This chapter included the overview of tools Cosmos and CANPC developed during the COST 266 project. They are both characterized by the general purpose framework which allows a diversity of problems to be described and solved. However, during the project just the applications concerning the optical networks have been developed. Applications in the Cosmos tool were oriented to the circuit switched network design and optimization and protection and restoration analysis. The applications developed in the Canpc tool were oriented to the routing and wavelength assignment in the circuit switched optical network, as well as planning of the packet switched optical network.

A short overview of two other network tools, the tool Demos, and the commercial WDM Guru tool, has been given as well.

## 6.6 References

- [1] M. Lackovic and R. Inkret, "Network Design, Optimization and Simulation Tool Cosmos," *Proc. of WAON 2001*, June 13-14, 2001, Zagreb, Croatia, pp. 37-44.
- [2] M. Lackovic, R. Inkret, and B. Mikac, "An Object-oriented Approach to Telecommunication Network Modeling," *Proceedings of ESM*, June 3-5, 2002, Darmstadt, Germany.
- [3] M. Lackovic, R. Inkret, and B. Mikac, "A Modeling Procedure for IP over WDM Transmission Network," *Proc. of MIPRO*, May 20-24, 2002, Opatija, Croatia.
- [4] M. Lackovic, R. Inkret, and B. Mikac, "An Approach to an IP over WDM Transmission Network Modelling," *Proc. of ICTON*, April 21-25, 2002, Warsaw, Poland.
- [5] D. Tarongi, J. Ehrensberger, S. Kessler, and C. Bungarzeanu, "Computer Aided Network Planning Cockpit CANPC," *Proc. of WAON 2001*, 13-14 June 2001, Zagreb, Croatia.
- [6] J. Ehrensberger, C. Lefeuvre, D. Tarongi, and C. Bungarzeanu, "Computer aided planing tool for all-optical WDM transport networks," in *the ECOC 2000 Workshop 'Modelling and Design of Optical Networks and Systems'*, Munich, Germany, September 3, 2000.
- [7] [www.python.org](http://www.python.org)

- [8] C. Bungarzeanu, L. Besuchet, and D. Rodellar, "Efficient planning tool for WDM transport networks," in *the 4th Working Conference on Optical Network Design and Modelling (ONDM 2000)*, Athens, Greece, February 7-8, 2000.
- [9] C. Bungarzeanu, J. Ehrensberger, and D. Tarongi, "Planning Tool for Wavelength Assignment in Static WDM Transport Networks," *IEEE ICT 2001*, 4-7 June 2001, Bucharest, Romania.
- [10] M. Matiello and C. Bungarzeanu, "Static Wavelength Assignment in WDM Networks with Optical Crossconnects," *Proc. of CONTEL 2003*, June 11-13, 2003, Zagreb, Croatia.
- [11] P. Strahm and C. Bungarzeanu, "Planning Tool for WDM Networks with All-Optical Packet Switching," *Proc. of WAON 2001*, 13-14 June 2001, Zagreb, Croatia.
- [12] M. Lackovic and C. Bungarzeanu, "An Approach to Planning and Performance Evaluation of Networks Based on Optical Packet Switching," *Proc. of ONDM 2003*, 3-6 February 2003, Budapest, Hungary.
- [13] DEMOS: a system for discrete event modelling on Simula / G.M. Birtwistle. London : Macmillian, 1979, ISBN: 0-333-23881-8
- [14] OPNET WDM Guru web site,  
<http://www.opnet.com/products/wdmguru/home.html>
- [15] S. De Maesschalck, D. Colle, M. Pickavet, and P. Demeester, "Availability Comparison of Pan-European Optical Backbone Networks," *Proc. of COST 266 Workshop on All-Optical Routing (WAOR) in conjunction with ICTON2002*, April 2002, Warsaw, Poland.
- [16] S. De Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz, "Pan-European Optical Transport Networks: an Availability-based Comparison," *Photonic Network Communications*, vol. 5, no. 3, May 2003, pp. 203-225.

THIS PAGE HAS BEEN LEFT BLANK INTENTIONALLY

## Chapter 7

# Evolution of Next-Generation Optical Networks

*Didier Colle, Erik Van Breusegem, Sofie Verbrugge, Jan Cheyns,  
Chris Develder, Sophie De Maesschalck, Mario Pickavet, and Piet Demeester*

### 7.1 Introduction

Chapters 3 and 4 of the final report deal with, respectively, optical circuit-switched and burst/packet-switched networks. The goal of this chapter is to discuss where these network technologies can fit in the evolution of next-generation networks.

The goal of section 7.2 is to discuss from a technology viewpoint one possible migration scenario: this results in an evolution from circuit- to packet-switched optical networks.

On the contrary, section 7.3 favours an evolution from packet- to circuit-switched networks, as circuit-switched networks become more cost-effective for larger traffic volumes. Finally, section 7.4 presents a novel node architecture that aims at combining the benefits of both circuit- and packet-switched networks.

Conclusions and opportunities for future research are highlighted in section 7.5.

## 7.2 A Technology Perspective

### 7.2.1. Introduction

It is well-known that optical burst- and packet-switching technologies are still in their infancy, while Wavelength Division Multiplexing (WDM) line-systems and Optical Cross-Connects (OXC) are mature and start to be deployed in operational networks. Therefore, from a technology perspective, an evolution from circuit-switched (see Section 7.2.2) to burst/packet-switched (see Section 7.2.3) optical networks is an obvious vision. The evolutionary scenario presented here consists of three major steps beyond the initial phase.

### 7.2.2. Evolution of Circuit-Switched Networks

#### *Initial Phase: static optical networking*

We start from a static optical network, containing Optical Cross-Connects (OXC) interconnected by WDM point-to-point systems. In such a network (see Figure 7.1) it is (conceptually) possible to set up lightpaths from each node to each other node in the network. The network (and the lightpaths it is carrying) are configured and managed by the network operator through the logically separated Telecommunication Management Network (TMN).

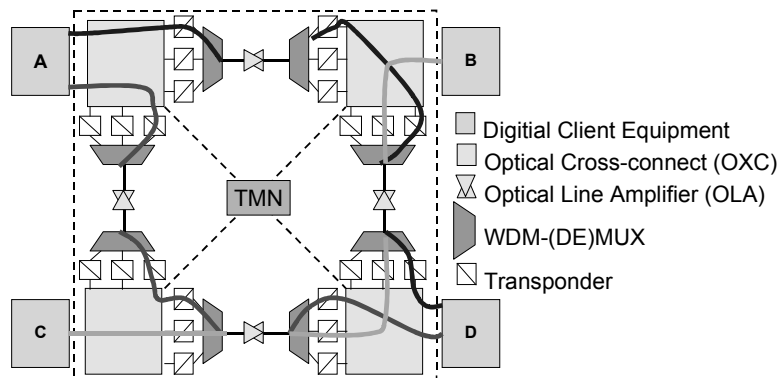


Figure 7.1. - Initial optical transport network

#### *Step 1: dynamic optical networks*

The initial network is managed by the network operator through the centralized management system. One of the main problems is that this does not allow performing fast lightpath provisioning. The introduction of a distributed control plane should result in an Automatic Switched Optical Network (ASON) [1], removing the need for any manual intervention of the network



operator in the lightpath provision process. The IETF is Generalizing the Multi-Protocol Label Switching (GMPLS) protocol suite, which should result in one candidate control plane solution [3]. A more detailed discussion on control plane solutions and architecture can be found in Chapter 3.

An example is illustrated in Figure 7.2. When a node is triggered to set up a lightpath to another node, then it sends a label (wavelength) request to the next hop, which forwards the request again to the next hop, until the request reaches the destination. The destination will answer the previous hop, by mapping (assigning) the lightpath to a particular wavelength, which means that it is ready to receive that particular wavelength or, in other words, that the previous node should send its signal on that wavelength. When the previous hop receives the label (wavelength) mapping from the next hop downstream, then it can send on its own a mapping to the previous upstream hop, while configuring its OXC.

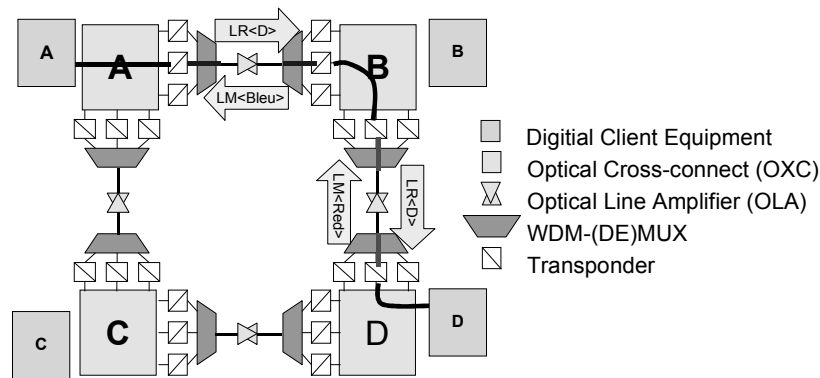


Figure 7.2. - Illustration of the label distribution process

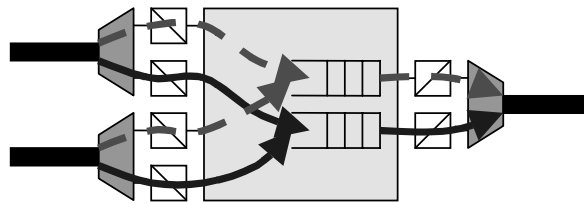
At first view, it seems that this evolution mainly impacts the (control of the) network level. Nowadays, there are still a bunch of open issues, with respect to these distributed control planes in optical networks, under investigations. A major issue to resolve is what functions the control plane will exactly take over from the management system and how the control and management plane will inter-work with each other.

### 7.2.3. Gradual Introduction of Optical Burst/Package Switching

#### *Step 2: introducing optical packets and buffers*

Up till now, we have been considering lightpaths, which have a fixed bandwidth due to SDH or DW framing, but which can be setup and torn-down in a very flexible and automated way (see also “Step 1: dynamic optical

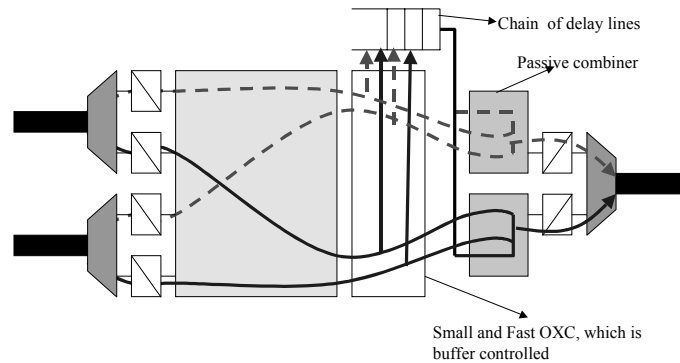
networks”). In this phase it was not possible to merge multiple incoming wavelengths into a single outgoing wavelength. This becomes possible when we evolve to optical bursts/packets ending up with a packet-switched network. At this stage, we assume that packets are switched based on their wavelength only. By merging multiple incoming wavelength channels into a single outgoing wavelength channel, contention may occur. Therefore, as illustrated in Figure 7.3, optical buffers might be required.



**Figure 7.3. - Introducing optical packets and buffers**

As discussed in Chapter 4, there are several technical issues (e.g., fixed versus variable length packets, synchronous versus asynchronous operation, which switching technology to adopt, etc) under investigation.

As shown in Figure 7.4, to facilitate the gradual introduction of burst/packet-switching, one might consider installing the wavelength merger mentioned and possibly the optical buffers at the output ports of an existing (large but slow) OXC. The figure shows that only a single output buffer – based on fiber delay line(s) – could be foreseen per outgoing link. A small and fast OXC (e.g., based on the SOA-technology) can then send each packet into the right fiber delay line or immediately to the outlet.



**Figure 7.4. - Output buffering**

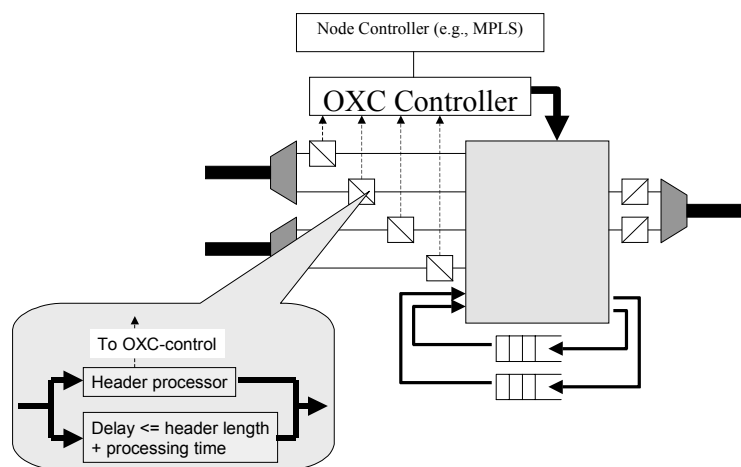
Another issue concerns the gradual introduction of optical packets and buffers. Has each node to be upgraded to optical packet switching at the same time? Is

it still possible to have circuit and optical packet switched wavelengths to transit the same equipment (lines and nodes)? Similarly, can data travel over a packet-switched wavelength, while being encapsulated/tunneled through a circuit-switched domain?

*Step 3: optical packet-header processing*

Up till now, all data entering the node on a particular line and wavelength combination leaves the node on a single line and wavelength combination. The goal of this stage is to upgrade the node in such a way that each optical packet entering the node on the same line and wavelength combination can leave the node on any line and wavelength combination independent of the routing of previous or later packets.

This implies that the core of the switch – in contrast to a wavelength merger at the output stage – needs to be capable of switching individual packets or bursts (examples are [5] or [6]). Nevertheless, the existing OXC could still be kept for cross-connecting/switching circuits through the network. As Figure 7.5 also highlights, processing (inclusive reading and rewriting) will be a major challenge in this phase of the evolutionary scenario.



**Figure 7.5. - Optical packet-header processing**

An important fact is here again - how can this stage coexist with previous stages? Can circuit-switched wavelengths transit the same equipment? Is it possible to send packets with a header through a domain, which is not optical packet-header processing capable? And how to control/manage such a hybrid network?

### 7.3 An Economical Perspective

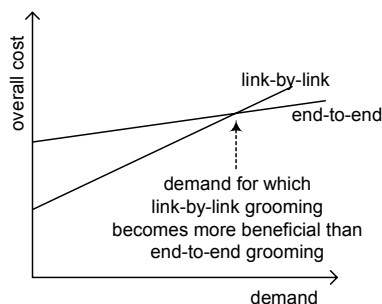
Although all-optical switching can significantly reduce the equipment cost compared to electronic switching techniques, switching individual packets will always remain more expensive – per transported bit – than switching complete circuits. However, as this section highlights, optical circuit switching only becomes attractive from an economical viewpoint, when the traffic volume over the network becomes sufficiently large.

#### *Some cost considerations*

In the packet switching or link-by-link grooming approach the initially required investment is limited, as no OXCs are needed. The incremental cost growing with the demand (router line cards), on the other hand, is significant. With the circuit switching or end-to-end grooming approach, the initial investment (OXCs) is important, but the incremental cost increase is limited. With a growing demand, the efficiency of the optical layer capacity in the end-to-end grooming approach increases, having a positive impact on cost per transported bit. Several scenarios are possible for the migration indicated here. Three of them are indicated below. In all cases initially link-by-link grooming is used throughout the whole network

#### *Network-wide migration in a single step*

In this scenario, at a certain point in time the network operator decides to alter the approach to end-to-end grooming and from that point onwards all traffic is routed directly from source to destination, without any special effort to fill the wavelength as well as possible. The migration point (see Figure 7.6) is determined by the overall cost of both approaches and the transition cost between them.



**Figure 7.6. - End-to-end (= circuit-switching) versus link-by-link (= packet-switching)**

*Node-by-node migration*

This scenario tries to find an answer to the question “In which nodes do we have to install an OXC?”. The nodes are transformed into little islands where we use end-to-end grooming. Gradually, more of these *end-to-end grooming islands* are introduced. An example is given in Figure 7.7. At the beginning link-by-link grooming is used throughout the whole network. If a certain node gets too heavily loaded, we install an OXC in that node and it becomes an end-to-end grooming island. As the time passes, more of these islands get introduced and islands can be merged to become a bigger island. Eventually (if the traffic keeps growing), the whole network can become end-to-end grooming.

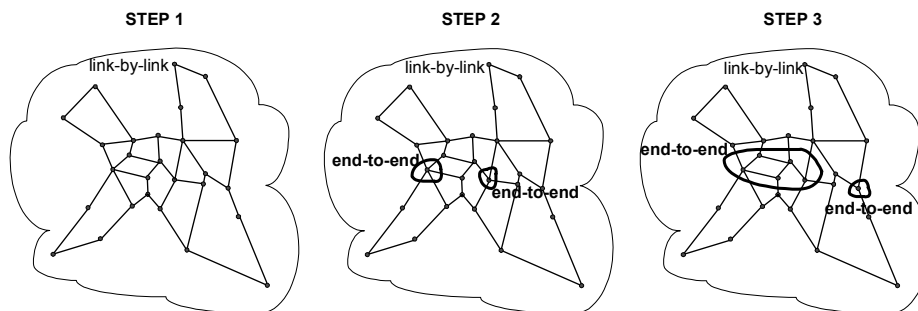


Figure 7.7. - Migration from end-to-end to link-by-link grooming

*Gradual migration, decision on a source-destination pair basis*

In both previous scenarios a network node always was completely link-by-link or completely end-to-end grooming. Nevertheless, also intermediate forms can occur. This happens, for example, if we observe the traffic on a source-destination pair basis. If there is enough high-speed traffic for a certain node pair, then this traffic is groomed on an end-to-end basis. In the meantime, traffic for other node pairs that have an insufficient amount of high-speed traffic is still groomed on a link-by-link basis. In this way there will be some nodes in the network where the traffic streams are completely unpacked in order to make it possible to perform link-by-link grooming, while other streams will be transmitted all-optically through those nodes.

## 7.4 Optional Step: ORION

As mentioned in the previous section, circuit-switching networks have the advantage of being rather cheap per transported bit, compared to packet-switched networks. However, packet-switched networks do not suffer from the

low bandwidth granularity imposed by the circuits. This section presents a node architecture that allows transporting most of the traffic volume over end-to-end circuits while switching only a minority of the traffic on a packet-per-packet basis. The goal is to combine in this way the cost-efficiency of circuit switching and the statistical multiplexing gain of packet-switching.

Sections 7.4.1 and 7.4.2 start by discussing the principle and possible node architecture. Afterwards, the case study in section 7.4.3 quantifies the potential benefits of the proposed concept. Finally, section 7.4.4 discusses the relevance of the proposed concept with respect to the evolution of next-generation networks. More details about ORION can be found in [4].

#### 7.4.1. Principle

ORION is a hybrid technology aimed at reducing the amount of transit traffic seen in the packet switches, while still allowing maximal statistical multiplexing. It accomplishes this by combining the wavelength switching concept with packet switched capability. Figure 7.8 shows a node that conceptually can support this. Fundamentally, it is based on the ASON concept, with a central OXC providing for a wavelength switched operation. Reconfigurability allows the option for medium to long-term adaptations to traffic variations.

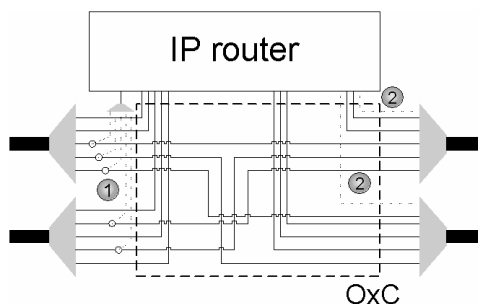


Figure 7.8. - A conceptual ORION node

In addition to this functionality, there are *taps* (1) added before the OXC, as well as *drains* (2) at the exit. The taps are capable of temporarily redirecting the established wavelength paths to the packet switched capable router. This means that, while packets on these wavelength paths normally bypass the packet switch, we add back the capability of local termination. To save interface cards however, the taps are multiplexed together. In this example, all are multiplexed together, effectively limiting redirection to the packet switch to only one packet at a time.

At the same time, at the exit, the drains are actually tunable lasers that can tune to any wavelength supported within the fiber.

What this kind of node concept allows us to do is, first of all, to operate the node standard as wavelength switched, with ASON capabilities. This means that data inserted in a wavelength follows a predestined path, as with standard wavelength switching. However, a situation can occur in which there is not enough capacity on a direct wavelength path. The taps and drains then allow for an operation of the node in a packet switched style mode. The difference is made by fitting the data with a signal, overriding the standard wavelength switched behaviour. When this signal is detected, the data is treated as if packet switched, and is directed through the taps to the packet switch. Note the fundamental difference with deflection routing here: deflection routing is not capable of interrupting the wavelength switched behaviour. Deflection can only use a concatenation of wavelength paths, possibly forcing data to take a longer path. The ORION concept, in contrast, offers full freedom. When data is treated in the special packet switched mode, it is said to be in “overspill mode”, as data “spills over” from its normal wavelength path to others that are not being used.

At the exit of the node a similar action can be taken. Indeed, with the wavelength switched approach, not all wavelengths are accessible at each node. With the addition of the tuneable lasers however, any wavelength can be chosen on which data can be sent for which not enough direct wavelength capacity is available. One last hurdle – from a network perspective - is that, when sending data on a normally inaccessible wavelength, we need to be careful not to garble up data which may possibly be already present on the now accessible wavelength path. Therefore, we need some form of idle detection. How this architecture can be implemented is further clarified in the next paragraph.

#### **7.4.2. Node Architecture**

To illustrate that we can achieve the required functionality for ORION without having to terminate each wavelength channel and to process every packet at intermediate nodes, we show an example switch design in a 4-wavelength network, Figure 7.9.

There is a central cross-connect (1), with four wavelengths coming in, of which two,  $\lambda_1$  and  $\lambda_3$ , are terminated towards the electrical edge node (2). In other words,  $\lambda_1$  and  $\lambda_3$  are the last section of a wavelength path.  $\lambda_2$  and  $\lambda_4$  pass through transparently (3). Data is intensity modulated (IM), while at the same time there exists a possibility to modulate orthogonally a label (at a lower bit

rate). We will use FSK as an example orthogonal signal, an approach pursued in the IST-STOLAS project.

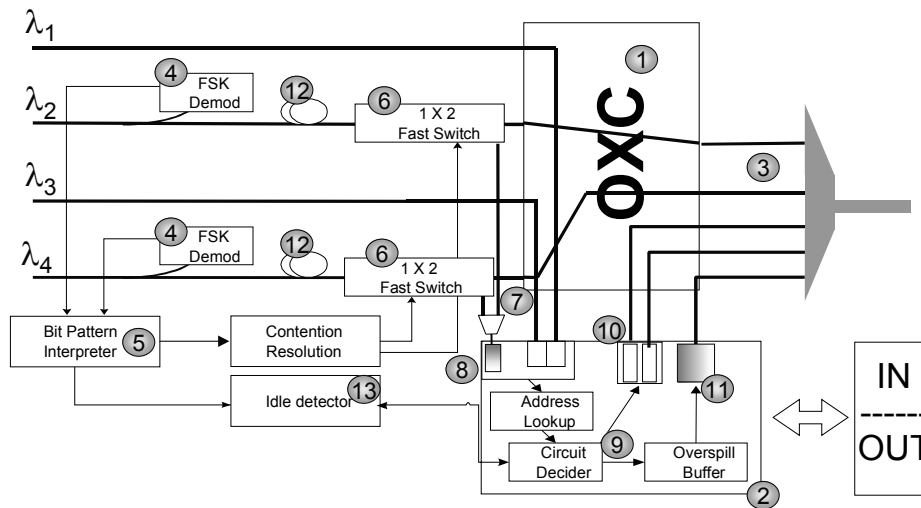


Figure 7.9. - An example: node architecture for the ORION network

Then, as a way of detecting whether or not a packet is in overspill mode, i.e., whether it is on a wavelength passing through but it does not belong there, the FSK label indicates this status. Before the packet enters the switch we split some of the power at (4) and read this label (5). If an overspill packet is detected, the  $1 \times 2$  fast optical switches (6) get set up to lead the packet towards the electrical edge node. Note that in the figure we combined (7)  $\lambda_2$  and  $\lambda_4$  (that were switched by the  $1 \times 2$  switches) into a single signal to save an interface card. The combined signal can (in this case) carry only a single wavelength since there is only one receiver (8) present. This implicitly means that we expect that packets in the overspill mode will be an exception rather than occur very frequently. Dimensioning the number of ORION enabled interfaces needed is a design parameter, depending on how many overspill packets there can be expected at the same time. Generally however, we reduce the number of wavelengths  $W$ , for which ORION packets can be handled simultaneously, to a smaller range  $N$ , without fixing beforehand on which wavelengths they are. How this is best done and the exact dimensioning of  $N$  currently remains open.

After the  $1 \times 2$  switch is set, the overspill packet is received by a wide-band receiver (8), after which the electrical edge node can either decide (9) to send



it via the normal regime with fixed (or slow tuneable) lasers (10), or again in overspill mode (11).

The reading and detection and the setting of the  $1 \times 2$  switches of course takes some time. Therefore, we provide some delay lines (12). All wavelengths can be equipped with delay lines, allowing a simpler node, since then we can multiplex all wavelengths in one physical fibre delay line.

The required delay lines also have a function for inserting packets in overspill mode on a pass-through wavelength. The same way we used an orthogonal FSK label to identify overspill packets, we use it for detection of idle periods on the wavelengths. Each packet on a circuit can be fitted with two labels, one at its end and one at its beginning, which indicate when the underlying intensity modulated packet occupies the line. As such they function as markers indicating the start and the beginning of a normal packet. The bit pattern interpreter now has a double function: identify overspill packets and detect idle periods, which can be used for overspilling.

Alternatively, power detection instead of FSK labels could be used to assess whether or not a packet is present on the pass-through wavelengths, or only a label at the beginning of the packet, which also indicates its length. Also, using dummy light for the idle periods would allow marking only those. This should also help to keep the optical power constant.

The information about availability is passed through to the idle detector (13) that tells the electrical edge router which wavelength to use for an overspill packet, should the provided circuit switched capacity be insufficient.

In order to be able to access all outgoing wavelengths without having to install a transmitter (and interface card) for each wavelength, a fast tuneable laser can be employed (11). The shown schematic implies that only one overspill packet can be sent at a time, which can of course be changed by simply adding more tuneable lasers.

### **7.4.3. Case Study**

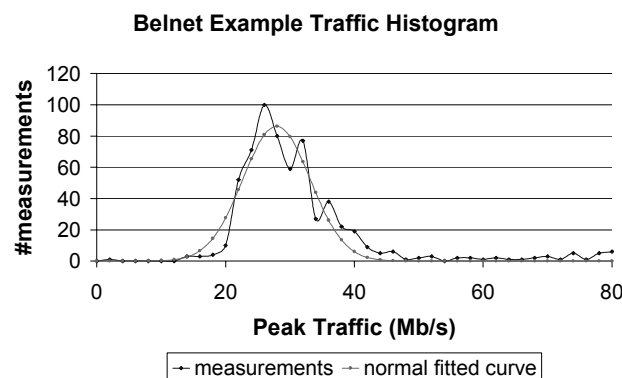
In order to quantify the advantage of the ORION concept, we ran a simple case study on the basic version of the COST 266 reference network, using the (revised) traffic profile forecast for 2008 (see Chapter 1). We compared three different network technologies types, calculated as follows:

- Fully packet switched (point to point WDM): Route all demands, and determine for each link how much wavelengths are needed to fulfil the demand it carries. Note that since full statistical multiplexing is possible, all demands can be combined, resulting in a minimal amount of required

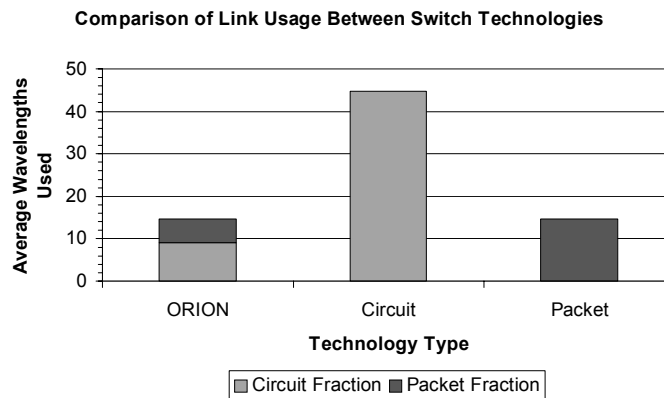
wavelengths while allowing a maximum amount of packets to be switched. In essence, all demands on a link are grouped and then collectively dimensioned.

- End-to-end wavelength paths only: Like in the previous case, but now we do not have any statistical multiplexing between demands (as explained earlier). Therefore, no demands can be combined, which result in the highest amount of wavelengths needed to route them. On the other hand, since there is no transit traffic anywhere, packet switches only have to route demands originating or terminating here. As a result, this network architecture has the minimum amount of traffic a node has to cope with. In essence, all demands on a link are dimensioned for separately.
- ORION: If idle periods in wavelengths can be completely re-used, ORION uses as much wavelengths as the packet switched technology since it can also employ statistical multiplexing. Furthermore, it only has to switch as much packets as in the combined case, since likewise wavelength paths are set up. As in the combined case, we calculated these figures to be optimized towards the number of wavelengths, as optimizing towards the number of packets handled in the node again would result in the end-to-end wavelength path case.

In order to make the comparison more realistic we also made a modification to the standard traffic matrix. Instead of fixed numbers we used distributions. In this way we can introduce statistical multiplexing effects when combining several streams, which should result in a more fair comparison (more favourable for packet switching). The conversion of the fixed matrix to a distribution form was based on measurements of peak moments from Belnet, the Belgian public transport network.



An example distribution is shown in Figure 7.10. It is fitted to the actual measurements of the Belnet network. From the figure it is already very clear that variance is significant, which will enlarge the benefits of a packet switched scheme. On the other hand, when the variance is not that large, probably most of the traffic can fit nicely in wavelength paths most of the time (minimal overspill traffic).



**Figure 7.11. - Dimensioning results**

Looking at the average amount of 10 Gbit/s wavelengths needed in the network in Figure 7.11, we see this confirmed. The end-to-end wavelength path case uses a lot more wavelengths than the two other alternatives. ORION uses the same amount of wavelengths as the packet switched case. Again we see an important feature of ORION: the most traffic gets sent through the direct wavelength paths, while only a small fraction uses the overspill mode, essentially packet switching. Yet, ORION still manages to stay within the same order of magnitude as circuit switching regarding the amount of packets that need to be switched. The reasons for the high amount of wavelengths needed in the network are the high bandwidth of one wavelength relative to the demand, and the fact that only end-to-end wavelength paths are set up. While not actually realistic, it does give a feel of the potential benefits of ORION.

#### **7.4.4. Migration Options**

Whereas section 7.2 favours – from a technology perspective – an evolution to optical burst/packet switched networks, section 7.3 favours an evolution to circuit-switched networks. As a hybrid architecture, ORION might fit as an intermediate step in one of these scenarios or one might consider a third possible evolutionary scenario, facing ORION as final step. The remainder of this

section elaborates on a seamless introduction of the ORION concept with respect to the evolutionary scenario described in section 7.2.

Note that a key requirement for ORION is the ability to recognize overspill packets. In the example this is based upon orthogonal FSK labelling. This means in fact that packets need to be put on the wire also in packet form, and that there in fact are idle times physically detectable. As a consequence current SDH framing technology poses a problem. In other words, ORION can be introduced at its earliest with the introduction of optical packets in the network in Step 2. From then on a seamless migration approach can be introduced. Instead of introducing optical buffers – an electrical packet switch is retained. No FSK/DPSK signal is employed, and the  $1 \times 2$  small, fast optical switches are not introduced yet. The result is still a dynamic optical network – with the same issues and functionality as mentioned in 7.2.2.

Next, delay lines are introduced. This increases the delay of the total network slightly (by a few ms), but should not be a big breach in the normal operations of the network.

Then, gradually, the orthogonal modulators/demodulators are installed on some wavelengths, together with the  $1 \times 2$  fast switches, wideband receivers and the tuneable lasers as mentioned in 7.4.2. Not all wavelengths need to be equipped with  $1 \times 2$  switches at once. Instead they can be added gradually as the network traffic profile or planning requires. The packet switch router remains electrical.

In the next phase the electrical packet switch can be replaced (in a drop-in fashion) by an all-optical one. Note the similarity between this stage and Step 3 in 7.2.3.

In the final phase then the OXC can be taken out – as well as the ORION specific components – ending up with a fully optical packet switched node.

## **7.5 Conclusions – Opportunities for Future Research**

Within this chapter several plausible evolutionary scenarios have been discussed. However, no consensus was obtained on what would be “the” evolutionary scenario. Future research is expected to find a consensus on this, based on an elaborated comparison between the different plausible scenarios. This does not exclude future research from coming up with more realistic scenarios (e.g., circuit-switched  $\rightarrow$  ORION  $\rightarrow$  packet-switched  $\rightarrow$  circuit).

It will be of utmost importance that this future research investigates different ways of seamless and gradual migration from one step to another one, inclusive the coexistence of two or more steps. When validating different possibilities it will also be important to verify on a system, component and even physical level that the proposed alternatives are technically feasible. Not only the hardware feasibility needs to be verified but also the viability of software architecture for the network control and management.

## 7.6 References

- [1] D. K. Hunter and I. Andonovic, "Approaches to Optical Internet Packet Switching," *IEEE Comm. Mag.*, vol. 38, no. 9, Sep. 2000, pp. 116-120.
- [2] "Architecture for the Automatic Switched Optical Network (ASON)", ITU-T draft G.ason, version 0.4, July 2001.
- [3] A. Banerjee *et al.*, "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements," *IEEE Communications Magazine*, vol. 39, no. 1, pp. 144-150, January 2001.
- [4] E. Van Breusegem *et al.*, "Overspill Routing in Optical Networks: a new architecture for future-proof IP over WDM networks," *Opticomm 2003*, October 13-17, Dallas, USA.
- [5] J. Cheyns, *et al.*, "Routing in an AWG-based optical packet switch", *Photonic Netw. Commun.*, vol. 5, no. 1, Jan. 2003.
- [6] C. Develder, J. Cheyns, E. Van Breusegem, E. Baert, A. Ackaert, M. Pickavet, and P. Demeester, "Node architectures for optical packet and burst switching," *PS 2002, the 2002 International Topical Meeting on Photonics in Switching*, July 21-25, 2002, Cheju Island, Korea.

THIS PAGE HAS BEEN LEFT BLANK INTENTIONALLY

## Chapter 8

# Summary of Results and Future Work

*Anton Kuchar, Ioannis Tomkos, Carmen Mas, and Anna Tzanakaki*

Technological progress is a never-ending process. Therefore, in this last chapter, after summarizing the results achieved, we propose the future work to be done in the framework of a follow-up COST 266 Action.

### 8.1 Summary of Results Achieved in COST 266 Action

The team of researchers that conceived COST 266 Action four years ago formulated a rather ambitious goal - to study all essential issues related to the next generation photonic networks. This included identification of the general telecommunication networks evolution trends from which requirements placed on the future photonic networks could be derived. It was soon concluded that these networks will be - apart from providing much higher transmission capacity spanning all-optically ever longer distances and at the same time advancing towards the end users - more flexible, agile, more reliable and therefore much more sophisticated. Another obvious trend has been the shift from the voice-centric to data centric networks. This implied putting more functionality and intelligence at the edge as well as within the networks, therefore, the overall architecture of photonic networks had to be overhauled.

With the aim of contributing to the development of the next generation photonic networks as outlined above, work of COST 266 research team has developed along three lines that corresponded to the three working groups formed within the Action.

Various aspects of physical limitations on transmission capacity and reach of photonic networks have been successfully studied in the framework of the Action. Most of the results are summarised in *Chapter 2* of this report and documented by the references given there. The issues studied concerned the effect of linear as well as non-linear phenomena in fibres on the capacity and reach of single- as well as multi-channel links and end-to-end paths in static but also in dynamic environment where the signals could be routed to their destination through different network nodes depending on the network state.

Dynamic behaviour of long chains of fibre amplifiers indispensable in flexible wide area ultra-high capacity photonic networks, where the number of wavelengths at the input to the amplifiers can rapidly vary or in some cases they could even handle bursty signals, have also been extensively studied.

The heavy traffic among network nodes in metro environment requires employment of very high density WDM systems. However, such systems suffer from four-wave mixing and place severe requirements on wavelength stability of IR sources. As demonstrated in this report, this problem can be substantially alleviated by employing polarisation multiplexing since, due to short distances involved, the effect of polarisation mode dispersion is rather limited in this case. The achievable optical channel count can then be doubled.

As Wavelength converters add flexibility to photonic networks, therefore, they also have been researched in COST 266 Action.

The photonic network elements such as optical ADMs and OXCs that will inevitably be used in the advanced photonic networks will substantially limit the size of all-optical networks, therefore, proper attention has been paid also to their physical properties. Extensive investigations on different OADM architectures have been done comparing the performance of two architectural options using an analytical optical signal to noise ratio model and specifications of commercially available optical components.

Ultimately, the reach and flexibility of all-optical networks are limited at the physical level by the analogue nature of optical signals. Therefore, novel devices enabling full (3R) regeneration of digital signals in the optical domain have been investigated and demonstrated.

Realisation of optical circuit switching based on wavelength switching and routing dealt with in *Chapter 3* is the next logical step on the way towards all-optical networks. This requires introducing intelligence into optical network nodes residing in their control plane. Extensive study has been done on the control plane architecture investigating the possibility of integrating the optical control plane with the control plane of client networks considering also



dynamic provisioning over multi-provider interconnected networks. For such novel operation scenarios the flexibility of the interconnection architecture model is crucial. Applying the new paradigm of dynamic on-demand provisioning at the points of domains' interconnections, a novel interconnection architecture model for the generalised multi-protocol label switching (GMPLS) network interconnections introducing multi-provider edge nodes has been proposed. A contribution to the discussion on the possibility to standardise the optical ATM-based Private Network to Network Interface as an alternative to GMPLS has been made.

Designing optimised reliable WDM networks that includes wavelength assignment requires efficient algorithms and tools able to perform this operation. Two wavelength assignment algorithms have been conceived and considered.

An adaptive routing mechanism for reducing the routing inaccuracy effects in an automatically switched optical network has been proposed to facilitate routing in optical networks under inaccurate network state information.

Great attention has been paid to the problem of protection in the optical layer for mesh networks, especially to the p-cycle span protection scheme where the protection capacity is organized in cycles which is much more capacity-efficient than protection schemes based on rings. It was shown that for all mesh protection schemes a unified formulation as an ILP-problem can be adopted. The objective of this research was to minimize the required capacity of the network that still has sufficient spare capacity available to establish the intended protection scheme and to protect all entities.

Since it is expected that IP traffic will soon prevail in telecommunication networks and WDM will become a ubiquitous optical transport technology, designing and dimensioning of static as well as dynamic IP-over-WDM networks has been a research subject of over-all importance that includes the problem of migration from the multi-layer solutions assuming intermediate layers like SDH/SONET. IP-over-WDM network modelling has been used to compare several data link layer topologies and routing and grooming in IP transport networks. The network has been modelled using two main types of Point of Presence nodes - the IP PoPs aggregating the IP traffic for transmission over the WDM network and the WDM PoPs serving as ingress and egress points for IP traffic entering and leaving a transmission network also providing switching functionality as intermediate nodes inside the WDM network. Traffic engineering concepts for dynamic WDM networks have been introduced and the influence of asymmetric IP traffic have been analysed. Protection and

restoration mechanisms have also been considered. The algorithm developed has been tested on a real network.

Connectivity of IP backbone routers was studied during one of the Action short-term missions. The traffic demands in the data layer are assumed routed along the shortest paths in the optical network layer. Three scenarios to obtain the data layer topology are presented and two scenarios for mapping IP on WDM, i.e., IP/POS/WDM and IP/GbE/WDM were considered. Dual link and Dual link with dual router configuration protection and restoration scenarios obtained from the first two network layers were introduced. The described algorithm has been applied to a network sample taken from the case study developed within COST 239 and COST 266 described in Chapter 1 of this report. It was concluded that in the IP/GbE/WDM scenario the optimal topology gives a substantial reduction in price compared to the other two proposed scenarios whereas in the IP/POS/WDM scenario the cost of the optimal topology is approximately equal to the cost of a fully meshed topology.

From the viewpoint of network evolution, one of the most probable architectures for future IP backbones is realisation of a dynamic SDH network on top of a dynamic optical network (IP/SDH/WDM). Efficient transport of dynamic traffic demands requires optimised multi-layer (electrical/optical) routing and grooming algorithms. Non-integrated routing schemes treat both layers separately while integrated schemes try to improve the performance by combining both layers. A multi-layer node model comprising an optical cross-connect on the optical layer and electronic cross-connect on the electrical layer has been proposed with the aid of which a simulation study was performed using a fictitious 9-node German network enabling to evaluate the blocking probability for different schemes.

The lightpaths, that support the logical IP topology in multi-layer data-centric optical networks, can be set up or torn down dynamically in Intelligent Optical Networks. Reconfiguring them on demand enables cross-layer traffic engineering. The logical network layer will use signalling to request new lightpaths from the optical layer or to remove existing lightpaths. It will also adjust IP routing to make sure the new logical topology is used efficiently. A case study has been performed to investigate the benefits of multi-layer traffic engineering when a reactive strategy was chosen solving problems after they have been detected.

When in the backbone network carrying asymmetric traffic unidirectional WDM line-systems instead of the current bi-directional ones are used, significant cost reduction can be achieved. This cost advantage has been quantified

for three of the COST 266 optical backbone topologies discussed in Chapter 1 of this report.

As a result of globalisation, the deployment of Virtual Private Networks (VPNs) by multinational corporations is increasing steadily. In Optical Virtual Private Networks (OVPN) only parts of the network resources of a public network of a provider serve closed group of users. These are built from wavelength paths within a multi-hop wavelength routed DWDM network using optical cross-connects. The open issue is how to set these OVPNs in such a manner that all user needs are satisfied while using as few network resources as possible. In the approach adopted in COST 266 Action the advantages of traffic grooming, where numerous MPLS traffic streams of a VPN can share a single wavelength path, have been exploited. Three types of OVPNs have been considered differing in granularity of the links provided - sub-wavelength, wavelength or fibre. With the objective to configure the VPNs and the light-path system optimally without separating the network layers, three methods have been evaluated in detail. Two protection alternatives of the three types of OVPNs and the three methods for the OVPN configuration have been considered - protection at the higher (electrical) layer and at the lower (wavelength) layer.

Data-centric photonic networks should be able to efficiently handle packets of data in the optical domain. Therefore, due attention has been paid to optical burst and optical packet switching (OPS/OBS) as evidenced by the overview of work done within the Action presented in *Chapter 4*. Advanced photonic are expected to have a highly dynamic connection pattern with a significant portion of bursty traffic. In such environment, OPS/OBS on one hand can reduce delays and improve the utilisation of the network's resources through statistical multiplexing, on the other hand it cannot provide explicit transfer guarantees.

In OPS/OBS forwarding is separated from switching in the network nodes. Forwarding decisions are taken by means of a burst control packet (BCP) or packet header that undergoes O/E conversion and electronic processing at the nodes while the burst/packet payload is optically switched. In this report all-optical processing of the headers is not considered. Semi-transparent networks optimised for a certain signal format and bit rates that after O/E/O conversion allow performing 3R regeneration, wavelength conversion and buffering in the electronic domain is assumed. A comparison of optical and electronic buffering is made. An overview of packet and burst handling schemes is given concluding that asynchronous and variable length data units are considered more suitable for OBS, while synchronous, fixed length packet (FLP) schemes

have been more popular in OPS although, more recently, variable length packets (VLP) with both asynchronous and synchronous operation have attracted some attention. In most cases, the synchronous, FLP mode or the asynchronous, VLP mode is considered to be the better choice for OPS.

The OPS and OBS are considered first for application in high-capacity DWDM mesh-based WANs or “core networks”. The network nodes combine edge router and core router functionalities. Edge (ingress/egress) routers modules are capable of communicating with the electrical layer, and may also form the interfaces to adjacent OPS/OBS layers. The main functionalities of an ingress router are to aggregate electrical packets into optical packets/bursts shaping the traffic flowing into the core. The egress router reads the optically encoded control information, defragments arriving packets/bursts, and forwards the client packets, after reassembly if required, to the client layer. A core router reads the control information after O/E conversion, processes this information electronically to configure the switch matrix and contention resolution resources. The impact of packet/burst length on node design is discussed and client packets representative of today’s Internet traffic are considered in the proposed schemes.

The design and analysis of OPS nodes is discussed in some detail in this work. The OPS/OBS network should be designed to offer its client layers throughput in the Tbit/s range to meet expected increases in traffic. Therefore, external blocking or “contention”, leading to loss of packets/bursts, which depends on the packet/burst switch architecture/designs, must be minimised including the overhead needed for control information, synchronisation pattern(s) and optical guard bands. Technology status of the two main building blocks - the control unit and the optical switching unit is briefly discussed. OPS/OBS architectures are then treated in detail describing a switch design suitable for asynchronous packet switching proposed in the framework of COST 266 Action. This design is suitable for low node degree since it does not scale well to a high node degree. Therefore, a scaleable design based on the same principles was proposed later.

The issue of contention resolution is central in OPS/OBS systems. Since they rely on statistical multiplexing of bursty traffic, overload situations occur that have to be resolved, otherwise burst or packet loss results. The problem of contention resolution in wavelength and time domains is discussed and results of a joint comparative performance evaluation are presented. Special consideration is given to electronic buffering that is an essential part of the proposed OPS scheme. The impact of different wavelength and delay selection algorithms and individual delays in a fibre delay line buffer is compared for OBS

and OPS both assuming asynchronous operation and variable length bursts or packets. Then, both approaches are compared to OPS with electronic buffers based on the number of buffer interfaces. Performance evaluation based on simulations from several partners in COST 266 Action is presented and different approaches for OPS/OBS contention resolution in time, assuming contention resolution also in the wavelength domain, is shown. It is concluded that both fibre delay line and electronic buffers can significantly reduce loss probability when used together with wavelength conversion. The comparison shows that electronic buffers need fewer but potentially more expensive O/E/O interfaces to reach the same loss rates.

Since the IP protocol itself does not support QoS differentiation, the OPS layer should be able to support QoS differentiation preventing over-dimensioning of the network nodes and delivering QoS differentiation to the IP-layer. The quality of the service offered will be influenced by the amount of resources available at the network nodes, like buffering and wavelength converters. Several approaches have been proposed and investigated for providing service differentiation directly in the optical layer by optical burst switching. They take advantage of burst reservation, burst assembly or a combination of both. These schemes are classified and analyzed in the report. Several alternatives of ensuring quality of service in optical packet switching have been investigated in the framework of COST 266 Action and are described in the report. Special attention is given to the OPS switching architecture employing electronic buffering. Simulations for quantifying to what extent the buffer inputs should be reserved when two traffic classes are assumed were used.

As a mature technology, MPLS can be applied in the first phase of building optical packet networks to support QoS, Class of Services and Traffic Engineering. MPLS core nodes can process labels rapidly thereby improving network performance. This is especially desirable in optical packet networks where buffering is a problem. By establishing label switched paths, resources can be reserved in order to provide QoS guarantees, e.g. buffers or wavelengths over predetermined paths. For these reasons, the issue of QoS in MPLS optical networks is discussed in the report.

Overview and comparison of selected types of ring based OPS for metro area networks (MAN) complete the chapter on OPS/OBS. Two new advanced architectures studied within the COST 266 action and the IST DAVID project are highlighted.

Reliability and availability are indispensable aspects when designing telecommunication networks. Therefore, due attention has been paid to this issue in COST 266 Action. **Chapter 5** overviews the problems addressed and sum-

marises the results achieved. Within the framework of COST 266 project two approaches have been taken to the reliability and availability analysis of all-optical networks - the probability based survivability measure, namely availability, and the traffic based survivability measures.

An overview of considered network architectures, protection switching schemes and of availability model together with the results of application of the model to the COST 266 case studies presented in Chapter 1 has been presented in the report.

Two types of optical networks have been considered in the model; passive optical WDM networks (PWN), without the possibility of automatic switching or rerouting of wavelength paths, and automatically switched optical WDM networks (ASWN). A WDM network is assumed to provide circuit switching service to the upper layer through wavelength paths (WP) or virtual wavelength paths (VWP). Availability model is developed and described in detail. Both types of networks considered have been modelled assuming no protection, 1+1 protection and 1:1 restoration. The modelling procedure is described step by step and results for the COST 266 case study topologies and different P&R architectures are presented. For the large topology, in the case of 1+1 protection and path restoration, the analytical procedure was not able to produce results in an acceptable amount of time due to high complexity in terms of both space and time. For the rest of the cases, good agreement between the results obtained analytically and those obtained by simulation was achieved.

The influence of component availabilities on the availability of different transport entities has also been studied. The goal of the study was to determine criticality of some components with respect to the unavailability of the transport entities and to draw conclusions on what should be important in the design process. Two concrete links were extracted from the European reference network - the shortest and the longest one. It is concluded that the amplifiers play a major role in determining unavailability of the photonic network with their influence increasing as the length of the communication links increases.

The impact of component sharing among several transport entities on logical connection availability has been studied by computer simulation. It is shown that in the case of component type unavailability variation the amplifiers would still be the component with the largest influence due to their large component count but the cable will be in the second place because of being shared by multiple logical channels. Several more case studies are included in the report to demonstrate the potential of the tools for network availability calculations.

The traffic based availability measure is then used to compare the reference pan-European optical backbone networks in terms of expected loss of traffic. The changing relative importance of voice and data traffic alters the traffic pattern. This the network operators an opportunity to save costs by applying different recovery schemes for voice and transaction data on one hand and IP traffic on the other hand. To illustrate this, a comparison is made between the use of 1+1 dedicated optical protection for all traffic on one hand and 1+1 protection for voice and transaction data traffic while IP data is routed unprotected, on the other hand. The networks used in the case study presented in this section are the ones introduced in Chapter 1 of this report. The three reference topologies are compared in terms of the integrity of the network design. The Average Expected Loss of Traffic (AELT) to be lost each year due to failures is used to express the availability of the network.

The relationship between distance and the traffic type is illustrated. This has a strong impact on the traffic based availability measure. In the extreme case of Internet traffic, where the geographical distance plays almost no role, the connections on average follow a longer path that is more prone to failures. The case study presented in the report shows that the recovery scheme applied has a big influence on the AELT. The trade-off between cost and AELT for the three reference topologies over time is illustrated. It is concluded that the triangular design is the most expensive one since it requires the longest ducts and the digging cost is the dominant cost component. On the other hand, the availability of the connections is best in the triangular network since the routes between origin and destination node are the shortest, and these routes have thus the highest availability. The other extreme - the RT design is the cheapest one but has the highest values of AELT. Hence, the network operators have to make a strategic decision whether the higher availability of the connections compensates the extra investment required to reach a lower AELT. It is also shown that when the IP traffic is transported as unprotected, the difference in AELT between the three studied network topologies is quite small.

Some of the methods and algorithms resulting from the numerous studies performed in the framework of COST 266 Action have been materialised in the form of tools for designing photonic networks. Two such network tools are described in detail and a short overview of the third one is given in *Chapter 6* of this report. All these three tools and some additional ones just mentioned in this report have been used in evaluating various network concepts proposed during the COST 266 Action.

The COSMOS tool main capabilities include topology description, behaviour description, simulation and optimisation that can be used in junction with

simulation or analytic procedures. Network topology or system structure consists of a list of network elements, and a list of connections between the network elements. A simulation model represents module behaviour that is more or less simplified behaviour of a real-world element. The simulation mechanism during simulation execution calls different run methods belonging to different modules. Network behaviour is based on behaviour of the modules it contains as well as on module communications. Simulation mechanisms are the core of simulation. They run user provided behaviour description and give some meaning to the code. Simulation domain implements a simulation mechanism. Different projects require different simulation domains that are the most efficient ones in dealing with a particular problem. The user must choose the best simulation domain relevant to the purpose. The discrete event domain fits most simulation demands when it comes to telecommunication network simulation but electronic devices are described best by differential equations. Several examples on the use of the tool illustrate its capabilities. These include:

- Availability and reliability simulation
- Modelling, optimisation and analysis of a WDM all-optical network based on configurable cross-connects and circuit switching
- Protection and restoration schemes in the automatic circuit switched optical network
- Gigabit Ethernet network modelling, simulation and analysis
- Broadband IP-router network modelling and analysis
- Traffic modelling
- Transmission Control Protocol layer modelling
- Logical topology optimisation
- Burst Switching.

The Computer Aided Network Planning Cockpit CANPC has been developed for designing multi-service networks. Within the CANPC network editor, nodes, links and traffic demands can be placed graphically. For larger networks, scripts can be used to automate the placement of nodes and links, the set-up of traffic demands and the configuration of all network elements. The network may be structured at different levels, each node representing a sub-network. The network model contains all input parameters that are taken into account by applications and algorithms to perform evaluations. The network model is edited by placing and interconnecting the network objects and speci-



fyng the object properties. When the network is defined in this way and specific parameter values are assigned, the evaluation may start. Some basic applications are already implemented into the framework and can be selected from a menu. However, the developers may build and integrate new user applications. The computations are split into algorithms that are called by applications each to perform a well-defined single task. The application library grows permanently by adding new applications in different fields. Many of them are already available in CANPC. These include

- Network editing applications (network generator, scripting application, link editor, traffic editor)
- Routing applications (shortest path routing, hierarchical routing, routing with load balancing)
- Dimensioning of WDM photonic networks (wavelength assignment, dimensioning of optical packet switched networks, dimensioning optical transport network with optical add-drop multiplexers, planning a logical topology for IP over WDM networks).

Integration of new applications, protocols and algorithms into the CANPC framework is controlled by the extensions manager. It allows the user to extend the set of active applications without any programming.

The third network tool extensively used in COST 266 Action is an advanced multi-layer network planning solution called OPNET WDM Guru that enables to design resilient, cost-effective optical networks. The availability studies presented in Chapter 5, Section 3 were obtained using this network design and evaluation software tool. It:

- creates four-layered network models,
- can model opaque as well as transparent network architectures,
- allows the user to generically model different types of WDM line systems and node equipment at the different layers,
- enables the user through its graphical interface to easily create and modify his network designs,
- allows to multiple scenarios per project,
- supports SONET/SDH ring architectures,
- can consider various protection mechanisms at the optical and electrical layers,
- performs the link design based on the transmission characteristics of the different WDM line system types,

- dimensions the nodes and links to accommodate a particular traffic demand,
- can perform routing at both the optical-channel and the digital-client layers,
- allows the user also to manually establish connections in the network on a connection-per-connection basis,
- enables user design and evaluate networks that have limited transparency reach and selective regeneration,
- can groom the traffic at the digital-client layer into optical wavelengths,
- can perform availability calculations and failure analysis, and
- provides through its browsers intuitive access to relevant node, link and connection information at the different layers.

In *Chapter 7* several feasible evolutionary scenarios have been discussed without making a conclusion on which might be the best one. Future research is expected to find a consensus on this, based on an elaborated comparison between the different plausible scenarios. It will be of utmost importance that the migration from the present to the future network is seamless and gradual also allowing the coexistence of two.

*To recapitulate:*

Broad range of issues has been addressed in the framework of COST 266 Action relevant to the development of advanced infrastructure for photonic networks:

- Physical layer limitations that in photonic networks determine the maximum transmission capacity, all-optical reach and resilience to dynamic traffic conditions
- Architecture of flexible, dynamic and intelligent photonic networks optimised for the multi-service/client/user/provider environment offering differentiated quality of service
- Optical burst and packet switching as the optimum method for transmitting bursty data characteristic for the forthcoming Internet era
- Availability/reliability of photonic networks that is becoming a critical factor in their design due to the huge transmission capacity they carry
- Development of methods, algorithms and powerful network tools to facilitate the design of the increasingly complex networks

- A reference pan-European network has been devised that was used for validating and comparing the methods, algorithms and tools developed in the framework of COST 266 Action
- Possible evolution of optical transmission networks have been studied including migration scenarios from the existing to the future networks.

## **8.2 Future Work - Proposal for New COST Action: *Towards Digital Optical Networks***

### **8.2.1. Introduction**

The explosive growth of data, particularly internet traffic has led to a dramatic increase in demand for transmission bandwidth imposing an immediate requirement for broadband networks. An additional driving force for higher capacity, enhanced functionality and flexibility networks is the increased trend for interactive exchange of data and multimedia communications. Due to the unpredictable and ever growing size of data files and messages exchanged over global distances the Future Communication Grid must be agile in time able to react rapidly to support end-to-end bandwidth requirements for transmission of messages and data files of any conceivable size encountered in real-life communications.

In order to address these requirements, telecommunications networks currently widely employ wavelength division multiplexing (WDM) in single-mode optical fibres to interconnect discrete network locations and offer high capacity and long reach transmission capabilities. The presence of dark fibre in existing networks can accommodate a large percentage of the capacity requirements, however, novel solutions are needed as equipping the already installed infrastructure with conventional technologies is not the most cost efficient solution.

The recent advances in optical technologies had a significant impact in telecommunications network solutions deployed world wide leading to a web of optical fibres that interconnect the globe offering advanced services to the end users. However, further optimisation of the existing solutions involving not only the physical implementation but also the control and management is required, in order to address the continuous evolution of services and applications that are becoming available to the users in a resilient and secure manner.

Technical breakthroughs in research are expected to further accelerate the realisation of transparent optical networks to offer increased transmission bandwidth, integrated transmission and switching capabilities and optical sig-

nal processing functionality. This progress is not only expected in the core networks, but also in the metropolitan area and the access networks to provide scaleable, transparent and flexible end-to-end solutions. These will enable accessibility of new services and applications to the end user offering full access to the global information network for all, with improved system performance and reduced cost.

### **8.2.2. Objectives**

The primary objective of the proposed COST Action *Towards Digital Optical Networks* is to focus on novel network concepts and architectures exploiting the features and properties of photonic technologies, to enable future telecommunications networks. It is aiming to propose a new generation of systems and networks that will accommodate the unpredictable and growing size of data files and messages as well as real time services (e.g. voice, video etc) exchanged over global distances requiring an agile Communication Grid supporting quality of services. This needs to provide end-to-end bandwidth for transmission of traffic for applications such as information retrieval, downloading (often multimedia) web software, exchange of various type of software (hundreds of Mbytes) and data models (Gbytes) etc as well as real time multimedia applications.

These systems need to be very flexible and rapidly reactive to efficiently accommodate the abrupt and unpredictable changes in traffic statistics introduced by current and future applications with low end-to-end latency. They will enable advanced features such as efficient and simple multicasting and broadcasting of broadband signals. In general, they need to support a future proof, flexible, efficient and bandwidth-abundant fiber-optic network infrastructure capable of supporting ubiquitous services in a resilient manner offering protection and restoration capabilities as well as secure services to the users.

Transparency to various digital signals and protocols is required to eliminate the need for multilayer complex network architectures suffering by poor scalability for data services, high latency, complicated network management and high cost. This migration can be gradually achieved by removing and/or integrating intermediate layers. Flat and upgradeable network architectures supporting photonic core and access technologies with intelligent edge nodes at the interfaces will form a universal infrastructure offering a variety of services supported by multiple operators. The ease of maintenance, provisioning and resilient operation required in this type of networks will be achieved through advanced routing and management mechanisms, eliminating the requirement

for ever increasing amounts of complex software raising the cost and limiting the network reliability and availability.

This advanced photonic infrastructure will employ optical signal processing and dynamic impairment management to eliminate the limitations of the analogue nature of traditional optical networks, dense wavelength division multiplexing technologies for signal transmission and routing, and optical packet and/or burst switching to provide fine bandwidth granularity, network efficiency and flexibility.

### **8.2.3. Technical Programme**

The technical programme will include research in several areas, the outcome of which will lead to a proposal of novel networking concepts and designs for a new generation of systems and networks that will accommodate the unpredictable and ever increasing traffic requirements over global distances in the future communications grid infrastructure. These areas will include:

- Design and implementation of systems based on innovative technologies supporting flexible, scaleable, upgradeable and performance optimised networks through improved functionality and characteristics. Such networks may form a universal infrastructure supporting a variety of services supported by multiple operators.
- These systems and networks will employ novel devices and materials part of which have been studied within the previous COST actions (267 and 268). Examples include transmission fibres, novel transmitters, receivers and amplifiers as well as passive or active technologies for signal processing based on enhanced semiconductor or silica designs. This action will not focus on technology, but will study its use and performance in system and networks. Interaction and cooperation with COST 288 will be very valuable and will continue.
- The performance of proposed realistic photonic network scenarios will be evaluated. These network scenarios will aim to support QoS, flexibility, ease of maintenance and provisioning, resilient, reliable and secure operation with low operational and capital network cost.

At the initial phase of the proposed Action it is planned to set-up three working groups focusing in the following domains:

- *Optical processing for digital network performance.* This workgroup will focus on physical layer and implementation related issues of transparent optical networks and will cover advanced topics such as optical signal per

bit processing, optical switch architecture designs and implementations as well as transmission related issues.

- *Novel network architectures.* This workgroup will focus on the evolution of network scenarios including the study of novel network architectures. This workgroup will also study different node architectures and technologies. Three different architectures will be studied and compared: circuit (wavelength, waveband etc), optical burst and optical packet switched networks.
- *Unified control plane, network resilience and service security.* This workgroup will focus on two directions. The former will deal with the impact of transparency on photonic network architectures and the associated control and protocol issues, which include topics related to wavelength routing, GMPLS control plane, OPS/OBS, traffic grooming-traffic engineering, OBGPs, VPNs etc. The latter will focus on network survivability and security issues, covering topics such as protection and restoration, its impact on routing and wavelength assignment algorithms, fault isolation, disaster recovery, etc.