Modeling and performance analysis of IP with MPLS/GMPLS over backbone networks

Shao, Xuguang

2009

Shao, X. (2009). Modeling and performance analysis of IP with MPLS/GMPLS over backbone networks. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/18686>

<https://doi.org/10.32657/10356/18686>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**MODELING AND PERFORMANCE ANALYSIS OF
IP WITH MPLS/GMPLS OVER
BACKBONE NETWORKS**

SHAO XUGUANG

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2009

MODELING AND PERFORMANCE ANALYSIS OF IP WITH MPLS/GMPLS OVER BACKBONE NETWORKS

SHAO XUGUANG

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2009

Acknowledgements

This research work would not have been finished without the help, advice, and encouragements from many people.

First of all, I want to express my sincere thanks to my supervisors Professor Shum Ping and Professor Zhang Liren for their constant supports, mentoring and feedback. Throughout the term of my study, their understanding and encouraging words have greatly motivated me; their invaluable guidance and thorough comments lightened the right direction during the progress of the research. Their persistence in believing my own ability for accomplishing the project provides me full motivation and flexibility.

It is a great pleasure to thank and acknowledge many professors whom I have the honor to know and associate with. Among them, I am especially grateful to Professor Kam Chan Hin, Professor Kot Chichung and Professor Cheng Tee Hiang for giving me such a precious chance to fulfill my dream of pursuing my postgraduate study in such a distinguished school of my beloved alma mater. Much gratitude also extends to Professor Koh Soo Ngee, Soong Boon Hee and Poo Gee Swee for their invaluable suggestions and comments during the review session of my first year M.Eng candidacy and evaluation session of my conversion from M.Eng to Ph.D candidacy. Finally, I sincerely want to give special thanks to the Joan and Irwin Jacobs Professor Vincent W. K. Chan from MIT for the remarkable discussion, encouragement and great inspiration.

Next, I wish to show my heartfelt appreciation to my fellow research scholars in Network Technology Research Centre (NTRC), especially Dr. Tang Ming, Dr. Chen Bin, Dr. Fu Song Nian, Dr. Xia Li, Dr. Li Xuejun, Dr. Yin Yongning, Xiao Shi, Dr. Wang Haibo, Hu Juan Juan, Dr. Yu Yang, Dr. Liu Ning, Dr. Yu Xia, and many others for their friendship, valuable discussion and encouragement.

Furthermore, I am grateful to all the technicians from NTRC, especially, Mr. Poh Khoon Yong, Mr. Chong Lui Tat, Mrs. Wee–Chew Lee Foong and Ms. Thung Lina for their kind support.

Finally, this thesis is dedicated to my family members who have been giving the love, trust, sacrifice and support to me over the years, especially for my dearest baby boy Michael.

Contents

<u>ACKNOWLEDGEMENTS</u>	<u>I</u>
<u>CONTENTS</u>	<u>III</u>
<u>LIST OF FIGURES</u>	<u>VII</u>
<u>LIST OF TABLES</u>	<u>XI</u>
<u>LIST OF ACRONYMS AND ABBREVIATIONS</u>	<u>XII</u>
<u>ABSTRACT</u>	<u>XV</u>
<u>CHAPTER 1. INTRODUCTION</u>	<u>1</u>
1.1. MOTIVATION	1
1.2. MAJOR CONTRIBUTIONS	6
1.3. OUTLINE OF THE THESIS	7
<u>CHAPTER 2. BACKGROUND AND LITERATURE REVIEW</u>	<u>9</u>
2.1. INTRODUCTION	9
2.2. QoS SCHEMES	10
2.2.1. INTSERV/RSVP VS. DIFFSERV	11
2.2.2. MULTI-PROTOCOL LABEL SWITCHING (MPLS)	14

2.2.3. LITERATURE REVIEW OF INTEROPERATION OF INTSERV/RSVP, DIFFSERV AND MPLS	16
2.3. LITERATURE REVIEW OF ANALYTICAL FRAMEWORKS FOR STATISTICAL MULTIPLEXING	22
2.3.1. MARKOV MODULATED POISSON PROCESS (MMPP) APPROACH	22
2.3.2. SRIRAM AND WHITT APPROACH	23
2.3.3. FLUID FLOW- BASED APPROACH	24
2.4. OPTICAL NETWORKING	25
2.4.1. NETWORK TRANSPARENCY	25
2.4.2. OPTICAL NETWORK COMPONENTS	27
2.4.3. OPTICAL IMPAIRMENTS	31
2.4.4. LITERATURE REVIEW OF BER PERFORMANCE DUE TO OPTICAL IMPAIRMENTS	35
2.5. OPTICAL PACKET SWITCHING (OPS) NODE	38
2.5.1. OPS NODE ARCHITECTURE	39
2.5.2. MAJOR CONTENTION RESOLUTION TECHNOLOGIES	40
2.5.3. LITERATURE REVIEW OF OPTICAL BUFFERING TECHNIQUES	43
2.6. GENERALIZED MULTI-PROTOCOL LABEL SWITCHING (GMPLS)	47
2.6.1. MERITS AND ISSUES	48
2.6.2. LITERATURE REVIEW OF IMPAIRMENT-AWARE ROUTING AND SIGNALING	48
2.7. SUMMARY	52
 <u>CHAPTER 3. MODELING AND PERFORMANCE ANALYSIS OF IP/MPLS BACKBONE NETWORK</u>	 <u>54</u>
 3.1. INTRODUCTION	 54
3.2. MPLS LABEL	55
3.3. PROPOSED TRAFFIC CLASSIFICATION	56
3.4. PROPOSED MAPPING FOR INTSERV, DIFFSERV & MPLS	57

3.5. MODELING OF NODE ARCHITECTURE	59
3.6. ASSUMPTIONS OF MATHEMATIC MODELING	62
3.7. MATHEMATIC MODELING AND ANALYSIS	62
3.7.1. PERFORMANCE OF PREMIUM QUEUE	62
3.7.2. PERFORMANCE ANALYSIS OF GOLD QUEUE	64
3.7.3. PERFORMANCE ANALYSIS OF SILVER QUEUE	70
3.7.4. PERFORMANCE ANALYSIS OF BONZE QUEUE	71
3.7.5. END-TO-END PERFORMANCE ANALYSIS	72
3.8. RESULTS AND DISCUSSIONS	73
3.8.1. EFFECT OF GOLD QUEUE LOAD ON THE LOSS PROBABILITIES AND THROUGHPUTS OF THE TRAFFIC CLASSES	75
3.8.2. EFFECT OF THRESHOLD CONTROL ON THE LOSS PROBABILITIES AND THROUGHPUTS OF THE TRAFFIC CLASSES	85
3.8.3. EFFECT OF GOLD QUEUE LOAD ON THE DELAY PERFORMANCES OF THE TRAFFIC CLASSES	98
3.9. EXTENSION TO ALL-OPTICAL OPS NODES	101
3.9.1. PROPOSED OPTICAL BUFFER ARCHITECTURE	102
3.10. SUMMARY	113

CHAPTER 4. MODELING AND TRANSMISSION PERFORMANCE ANALYSIS OF ALL-OPTICAL WDM BACKBONE NETWORK **115**

4.1. INTRODUCTION	115
4.2. ASSUMPTIONS OF NETWORK CHARACTERISTICS	117
4.3. PROBLEM HIGHLIGHTS	118
4.4. MODELING	119
4.4.1. PER-HOP MODELING	120
4.4.2. SPACE SWITCH MODELING	122

4.5. PERFORMANCE ANALYSIS	124
4.5.1. ENUMERATION PROCESS	125
4.5.2. BER CALCULATION	128
4.5.3. PACKET ERROR CALCULATION	131
4.5.4. EFFECTIVE NETWORK UTILIZATION CALCULATION	131
4.6. RESULTS AND DISCUSSIONS	133
4.6.1. OPTICAL LAYER PERFORMANCE	135
4.6.2. EFFECT OF NETWORK LSP SETUP ON OPTICAL LAYER PERFORMANCE	155
4.6.3. ADVERSE EFFECT OF OPTICAL IMPAIRMENTS ON NETWORK PERFORMANCE	158
4.7. GUIDELINES TO NEXT GENERATION NETWORK DESIGN	160
4.7.1. ULTIMATE GOAL: ALL–OPTICAL TRANSPARENT NETWORK	160
4.7.2. INTERMEDIATE STEP: TRANSLUCENT NETWORK	161
4.7.3. OUR CHOICE OF NGN ARCHITECTURE	163
4.8. SUMMARY	171

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

173

5.1. CONCLUSIONS	173
5.2. RECOMMENDATIONS FOR FUTURE WORK	175

REFERENCES

178

APPENDIX A: FLUID FLOW MODEL

A

APPENDIX B: BIT ERROR RATE (BER)

D

List of Figures

FIGURE 1–1: INTER-NETWORKING OF INTSERV, DIFFSERV AND MPLS HYBRID NETWORK	3
FIGURE 1–2: EVOLUTION TOWARDS PHOTONIC NETWORKING	5
FIGURE 2–1: RSVP SIGNALING	11
FIGURE 2–2: INTSERV ARCHITECTURE.....	12
FIGURE 2–3: DIFFSERV ARCHITECTURE.....	13
FIGURE 2–4: ROUTING FUNCTIONAL COMPONENTS: CONTROL AND FORWARDING	14
FIGURE 2–5: PACKET TRAVERSING A LABEL SWITCHED PATH	16
FIGURE 2–6: INTSERV/DIFFSERV INTEROPERATION SCENARIO	17
FIGURE 2–7: INTEROPERATION OF INTSERV AND MPLS WITH SINGLE INGRESS LSR	19
FIGURE 2–8: TWO STATE MMPP	22
FIGURE 2–9: OPAQUE NODE, WITH FIXED ROUTING, AND NO WAVELENGTH CONVERSION	26
FIGURE 2–10: TRANSPARENT NODE, WITH FIXED ROUTING, AND NO WAVELENGTH CONVERSION	27
FIGURE 2–11: ERBIUM DOPED FIBER AMPLIFIER (EDFA).....	29
FIGURE 2–12: TRANSLUCENT NODE, COMPOSED OF TRANSPARENT SECTION (O) FOR PASSING THROUGH SIGNALS AND OPAQUE SECTION (E) FOR ADDING/DROPPING AND REGENERATING SIGNALS	30
FIGURE 2–13: CROSSTALK IN A 2X2 SWITCH	33
FIGURE 2–14: OPTICAL FREQUENCY ARRANGEMENT OF FWM	34
FIGURE 2–15: GENERIC NODE ARCHITECTURE IN OPS NETWORK.....	40
FIGURE 2–16: DEFINITION OF OPTICAL BUFFER	41

FIGURE 2–17: OPTICAL BUFFERING IN AN OPS NODE.....	45
FIGURE 2–18: SCHEME (A) CENTRALIZED RWA+ CENTRALIZED SIGNALING	49
FIGURE 2–19: SCHEME (B) CENTRALIZED RWA + DISTRIBUTED SIGNALING.....	50
FIGURE 2–20: SCHEME (C) DISTRIBUTED APPROACH WITH MODIFICATION TO ROUTING PROTOCOLS.....	51
FIGURE 2–21: SCHEME (D) DISTRIBUTED APPROACH WITH MODIFICATION TO SIGNALING APPROACH	51
FIGURE 3–1: MPLS HEADER FIELDS.....	56
FIGURE 3–2: TRAFFIC CLASSIFICATION	57
FIGURE 3–3: OUTPUT PORT MODELING OF THE LSR.....	60
FIGURE 3–4: PROPOSED OUTPUT–PORT FOUR–QUEUE SYSTEM.....	61
FIGURE 3–5: MODELING OF INTSERV GS TRAFFIC	63
FIGURE 3–6: MODELING OF DIFFSERV SOURCE.....	65
FIGURE 3–7: MODELING OF INTSERV CLS TRAFFIC	72
FIGURE 3–8: QUEUEING MODEL FOR NUMERICAL ANALYSIS	74
FIGURE 3–9: EFFECT OF CLASS 1 TRAFFIC LOAD ON LOSS PROBABILITY OF GOLD QUEUE, WITH AND WITHOUT THRESHOLD CONTROL.....	77
FIGURE 3–10: EFFECT OF CLASS 1 TRAFFIC LOAD ON THROUGHPUT OF GOLD QUEUE, WITH AND WITHOUT THRESHOLD CONTROL	79
FIGURE 3–11: EFFECT OF GOLD QUEUE TRAFFIC LOAD ON LOSS PROBABILITY OF SILVER QUEUE, WITH AND WITHOUT THRESHOLD CONTROL.....	81
FIGURE 3–12: EFFECT OF GOLD QUEUE TRAFFIC LOAD ON LOSS PROBABILITY OF SILVER QUEUE, WITH AND WITHOUT THRESHOLD CONTROL.....	84
FIGURE 3–13: EFFECT OF GOLD QUEUE THRESHOLDS CONTROL ON LOSS PROBABILITY OF GOLD QUEUE	86
FIGURE 3–14: EFFECT OF GOLD QUEUE THRESHOLDS CONTROL ON LOSS PROBABILITY OF SILVER QUEUE.....	88
FIGURE 3–15: EFFECT OF GOLD QUEUE THRESHOLDS CONTROL ON THROUGHPUT OF GOLD QUEUE	89

FIGURE 3–16: EFFECT OF GOLD QUEUE THRESHOLDS CONTROL ON THROUGHPUT OF SILVER QUEUE.....	91
FIGURE 3–17: EFFECT OF SILVER QUEUE THRESHOLDS CONTROL ON LOSS PROBABILITY OF SILVER QUEUE.....	93
FIGURE 3–18: EFFECT OF SILVER QUEUE THRESHOLDS CONTROL ON LOSS PROBABILITY OF GOLD QUEUE	94
FIGURE 3–19: EFFECT OF SILVER QUEUE THRESHOLDS CONTROL ON THROUGHPUT OF SILVER QUEUE.....	96
FIGURE 3–20: EFFECT OF SILVER QUEUE THRESHOLDS CONTROL ON THROUGHPUT OF GOLD QUEUE	97
FIGURE 3–21: EFFECT OF CLASS 2 TRAFFIC LOAD ON AVERAGE DELAY OF OTHER FOUR CLASSES	100
FIGURE 3–22: PROPOSED OPTICAL BUFFER FUNCTIONING AS FIFO BUFFER AT ONE OUTPUT PORT – SINGLE WAVELENGTH CASE	103
FIGURE 3–23: EQUIVALENT ELECTRONIC BUFFER OF FIGURE 3–22	106
FIGURE 3–24: PROPOSED OPTICAL BUFFER SYSTEM TO BE USED AT THE OUTPUT PORT OF OPS TO PROVIDE SERVICE DIFFERENTIATION.....	108
FIGURE 3–25: EQUIVALENT ELECTRONIC BUFFER OF FIGURE 3–24	112
FIGURE 4–1: SIMPLE NETWORK.....	118
FIGURE 4–2: PER–HOP MODELING OF INTERMEDIATE NODES – COMPONENTS AND THEIR LOSS/GAIN PARAMETERS	121
FIGURE 4–3: CROSSTALK EXAMPLE IN 2X2 AND 4X4 SWITCHES	123
FIGURE 4–4: MORE SWITCHES – 8X8 AND 16X16 STRUCTURES.....	124
FIGURE 4–5: TWO PARTS OF AN IP PACKET	131
FIGURE 4–6: AN N HOP CONNECTION. CONNECTION LINKS ARE LABELED. THE OTHER LINKS ARE INTERFERING LINKS.....	132
FIGURE 4–7: SAMPLE OPTICAL MESH NETWORK	136
FIGURE 4–8: MODELING OF INTERMEDIATE NODES IN DSF 8 SCENARIOS	137
FIGURE 4–9: Q PARAMETER OF SYSTEM A AT 1549.4NM (4TH).....	140
FIGURE 4–10: Q OF SYSTEM B AT 1549.4NM (8TH).....	141

FIGURE 4–11: Q OF SYSTEM C AT 1549.4NM (8TH).....	142
FIGURE 4–12: BER OF SYSTEM A AT 1549.4NM (4TH).....	143
FIGURE 4–13: BER OF SYSTEM B AT 1549.4NM (8TH).....	144
FIGURE 4–14: BER OF SYSTEM C AT 1549.4NM (8TH).....	145
FIGURE 4–15: COMPARISON OF Q PARAMETERS FOR SYSTEM A, B, AND C	146
FIGURE 4–16: COMPARISON OF BER FOR SYSTEM A, B AND C.....	146
FIGURE 4–17: MODELING OF INTERMEDIATE NODES IN SMF+DCF SCENARIOS	147
FIGURE 4–18: Q OF SYSTEM D: SMF+DCF 40GBPS 0.8NM SPACING.....	150
FIGURE 4–19: Q OF SYSTEM E: SMF+DCF 40GBPS 0.4NM SPACING	151
FIGURE 4–20: BER OF SYSTEM D: SMF+DCF 40GBPS 0.8NM SPACING	152
FIGURE 4–21: BER OF SYSTEM E: SMF+DCF 40GBPS 0.4NM SPACING	153
FIGURE 4–22: Q COMPARISON OF SYSTEM E AND F.....	155
FIGURE 4–23: BER COMPARISON OF SYSTEM E AND F.....	155
FIGURE 4–24: NETWORK TOPOLOGY	156
FIGURE 4–25: COMPARE LSP Q.....	157
FIGURE 4–26: COMPARE LSP BER.....	158
FIGURE 4–27: EFFECTIVE NORMALIZED THROUGHPUT FOR SYSTEM A, B, AND C	159
FIGURE 4–28: EXAMPLE OF A TRANSLUCENT NETWORK WITH THREE TRANSPARENT ISLANDS.	162
FIGURE 4–29: RECALL RESULTS IN FIGURE 4–11	165
FIGURE 4–30: Q PERFORMANCE OF SYSTEM C WITH DIFFERENT NUMBERS OF WAVELENGTHS ON	166
FIGURE 4–31: ASSIGN OEO NODES AT BOUNDARY OF ISLAND OF TRANSPARENCY	168
FIGURE 4–32: TWO CHOICES OF THE PROPOSED APPROACH	170

List of Tables

TABLE 2–1: DIFFSERV PHB TO MPLS SERVICE CLASS AND EXP MAPPING	18
TABLE 2–2: EXP TO MPLS SERVICE CLASS MAPPING	21
TABLE 2–3: SERVICE MAP FROM INTSERV/DIFFSERV TO MPLS	21
TABLE 2–4: COMPARISON OF THE IMPAIRMENT–AWARE LIGHTPATH SETUP SCHEMES	53
TABLE 3–1: SERVICE CATEGORIES MAPPING OF TYPICAL APPLICATIONS	58
TABLE 3–2: SERVICE MAPPING FOR INTSERV, DIFFSERV AND MPLS SERVICE CATEGORIES	58
TABLE 3–3: PROPOSED QOS MAPPINGS	59
TABLE 4–1: OPTICAL COMPONENTS AND THEIR LOSS/GAIN PARAMETERS	121
TABLE 4–2: COMMON PARAMETERS SETTING	134
TABLE 4–3: FIBER PARAMETERS.....	135
TABLE 4–4: PARAMETER SETTING FOR THREE CONFIGURATIONS OF DSF 10 GBPS SYSTEM	138
TABLE 4–5: PARAMETER SETTING FOR TWO CONFIGURATIONS OF SMF+DCF 40 GBPS SYSTEMS	148
TABLE 4–6: PARAMETER SETTING FOR TWO 40GBPS SYSTEMS	154

List of Acronyms and Abbreviations

AF:	Assured Forwarding
ASE:	Amplified Spontaneous Emission
ATM:	Asynchronous Transfer Mode
BA:	Behavior Aggregate
BE:	Best Effort
BER:	Bit Error Rate
BGP:	Boarder Gateway Protocol
CD:	Chromatic Dispersion
CLS:	Controlled – Load Service
CR–LSP:	Constraint–Routed Label Switched Path
CSO:	Composite Second–Order
DCF:	Dispersion Compensating Fiber
DF:	Default Forwarding
DLCI:	Data Link Connection Identifier
DRUM:	DiffServ and RSVP/intserv Use of MPLS
DiffServ:	Differentiated Service
DSCP:	DiffServ Code Point
DSF:	Dispersion Shifted Fiber
EDFA:	Erbium Doped Fiber Amplifier
EF:	Expedited Forwarding
E/O:	Electro–optic
E – OXC:	Electronic module of an OXC

EXP:	Experiment (Experiment Bit in MPLS header)
FEC:	Forwarding Equivalent Class
FDL:	Fiber Delay Line
FR:	Frame Relay
FWM:	Four Wave Mixing
Gbps:	Gigabits per second
GMPLS:	Generalized Multi-Protocol Label Switching
GS:	Guaranteed Service
GVD:	Group Velocity Dispersion
IETF:	Internet Engineering Task Force
IntServ:	Integrated Service
IP:	Internet Protocol
IS-IS:	Intermediate System to Intermediate System
ITU:	Internet Telecommunication Union
ITU – T:	ITU Telecommunication Standardization Sector
LDP:	Label Distribution Protocol
LER:	Edge LSR
LSP:	Label Switched Path
LSR:	Label Switch Router
MF:	MicroFlow
MPLS:	Multi-Protocol Label Switching
NGN:	Next Generation Network
NZDSF:	Non-Zero Dispersion Shifted Fiber
NZDSF+:	NZDSF with positive dispersion at 1550 nm
NZDSF–:	NZDSF with negative dispersion at 1550 nm
O/E:	Opto-electronic
OEO:	Optical-Electrical-Optical
O – OXC:	Optical module of an OXC
OSPF:	Open Shortest Path First

OSNR:	Optical Signal-to-Noise Ratio
OXC:	Optical Crossconnect
PER:	Packet Error Rate
PHB:	Per Hop Behavior
PMD:	Polarization Mode Dispersion
PSB:	Partially Shared Buffer
QoS:	Quality of Service
RMS:	Root Mean Square
RSVP:	Resource Reservation Protocol
RWA:	Routing and Wavelength Assignment
S:	Stack (Stack bit in the MPLS header)
SBS:	Stimulated Brillouin scattering
SDH:	Synchronous Digital Hierarchy
SL:	Slow Light
SLA:	Service Level Agreement
SMF:	Single Mode Fiber
SPM:	Self-Phase Modulation
SNR:	Signal-to-Noise Ratio
SOA:	Semiconductor Optical Amplifier
SONET:	Synchronous Optical NETWORK
SRS:	Stimulated Raman Scattering
Tbps:	Terabits per second
TE:	Traffic Engineering
TTL:	Time-To-Live
VC:	Virtual Circuit
VCI:	Virtual Circuit Identifier
VPI:	Virtual Path Identifier
WDM:	Wavelength Division Multiplexing
XPM:	Cross-Phase Modulation

Abstract

Backbone network is at the core of the computer network infrastructure and we aim to model and analyze the performance of the backbone network under various conditions. In a copper-based or opaque fiber-based backbone network, Integrated Service (IntServ) and Differentiated Service (DiffServ) should be viewed as complementary technologies in the pursuit of end-to-end Quality of Service (QoS). Multi-Protocol Label Switching (MPLS) is a good candidate that facilitates the interoperation of the IntServ and DiffServ network. Meanwhile, with the introduction of Generalized Multi-Protocol Label Switching (GMPLS) as the common control plane between optical crossconnects (OXC) and electronic switching devices, the integration of Internet Protocol (IP) and Wavelength Division Multiplexing (WDM) can be efficiently achieved. In the case of all-optical IP over WDM network, traffic flows can be directly mapped to optical channels without the intermediate layers such as Asynchronous Transfer Mode (ATM) and Synchronous Optical Networking (SONET) / Synchronous Digital Hierarchy (SDH). Therefore, the bursty nature of IP traffic may have significant effect on the performance of the optical channels. On the other hand, the optical impairments of the optical paths and optical nodes may greatly influence the performance of IP traffic at the network layer. In this thesis, new traffic classification and service mapping among the IntServ, DiffServ and MPLS/GMPLS are proposed to enable the interoperation of these technologies. More importantly, a novel four-queue system is proposed to be located at every output port of the MPLS Label Switch Router (LSR) or GMPLS OXC, but in different manners respectively. The functionally equivalent electronic queueing system and optical queueing system are illustrated in great details. In this way, the network node in the backbone is able to ensure the QoS of both the IntServ

and DiffServ edge customers. The performance analysis of the four-queue system is elaborated and the performance parameters in terms of throughput, loss probability and delay are derived analytically. We have demonstrated the desired QoS can be achieved for such a heterogeneous network with traffic from both IntServ and DiffServ customer edges over the MPLS/GMPLS transit network by using the proposed traffic classification, service mapping and queueing system. For the case of fully transparent all-optical backbone network, where the underlying infrastructure is WDM-based, performance analysis with the considerations of optical impairments along the optical paths and at the optical nodes is also conducted and performance is evaluated. We find that nonlinear impairments and linear impairments impact the optical network performance in different manners. Because of the non-availability of product that can perform optical reamplification, reshaping and retiming (optical 3R) in near term, the ultimate next generation network (NGN) architecture that is fully transparent all-optical IP over WDM network can be hardly achieved in near or medium term. Hence, we take one step back and adopt an intermediate-step network architecture, which is translucent with a collection of islands of transparency. We present a novel approach in designing our translucent network of interest, where linear impairments are used to define the larger islands of transparency and nonlinear impairments are used for impairments aware routing and wavelength assignment (RWA) and signaling.

Chapter 1. Introduction

A backbone network or network backbone is a part of computer network infrastructure that interconnects various pieces of network, providing paths for the exchange of information between different subnetworks. A backbone can tie together diverse networks in the same building, in different buildings, in a campus environment, or even national boundaries. Usually, in the backbone, high capacity communication facilities connect primary nodes in a mesh manner. Many main service providers in the world have their own backbone networks that are interconnected. Some large enterprises have their own backbone network, which are typically connected to the public networks. Backbone networks can be either copper-based or fiber-based, which raises new issues and poses distinct challenges for the design, implementation, modeling and performance analysis in the backbone.

1.1. MOTIVATION

Internet Protocol (IP) has been a key factor to propel the dramatic growth of data traffic in the past few decades. The trend of developments in IP networks is towards supporting quality of service (QoS) [1]. QoS generally describes the assurance of sufficiently low packet loss and delay for certain types of applications or traffic. The requirements can be given by human factors, e.g., tolerance to poor quality of an image as a result of loss of pixels, or by business needs, e.g., the need to complete a transaction within a given time horizon. It is expected that in many cases, traffic of clients with different traffic requirements, coming from various kinds of edge networks, will travel through the backbone network. Different priority service classes are hence necessary to meet different

QoS requirements from different clients. For example, some companies that providing web service would like to provide predictable Internet services. Such companies will be willing to pay a certain price to make their service reliable. They may offer a single service or multiple services that contain Gold category service, Silver category service and Bronze category service, with decreasing priority. Some other companies require low delay and low loss services to their applications, such as Internet Telephony and Video Conferencing. They are willing to pay a premium price to run a high quality videoconference to save travel time and cost. Finally, the Best Effort Service will remain for those who only need basic connectivity. IETF has proposed many service models and mechanisms to meet the demand for QoS, among which are Integrated Service (IntServ)/ Resource Reservation Protocol (RSVP) [2]-[3], Differentiated Service (DiffServ) [4], and Multi-Protocol Label Switching (MPLS) [5]. With the introduction of these QoS mechanisms, Internet is now able to support voice and video applications [6]. Hence IP-based traffic volumes have been continuously growing at a fast speed.

IntServ [2] offers end-to-end QoS, which is based on RSVP [3] signaling. On the other hand, DiffServ [4], which is based on prioritization and operating on flow aggregates, provides a more scalable solution to backbone networks. MPLS architecture, originally presented as a way of improving the forwarding speed of routers, is now emerging as a crucial standard technology that offers new QoS capabilities for large-scale IP backbone networks [5]. Consequently, MPLS attracted a lot of interest in research and applications.

Figure 1–1 shows how these different technologies would interoperate in today's Internet environment. In this scenario, senders from the IntServ and DiffServ networks need to communicate with receivers in the other IntServ and DiffServ networks respectively through the MPLS transit network. The MPLS backbone network consisting of Label Switch Routers (LSRs) provides transport and QoS guarantees for these edge customers.

Figure 1–1: Inter-networking of IntServ, DiffServ and MPLS Hybrid Network

To support different traffic requirements, QoS control is expected to be implemented in the backbone network. The network is expected to have heterogeneous architecture with traffic coming from different kinds of edge networks (i.e. IntServ edge network or DiffServ edge network) and traversing through the MPLS backbone network. How to implement QoS support for such heterogeneous network architectures poses new challenges for the design, modeling and performance analysis.

The development of optical communication technologies, especially the Wavelength Division Multiplexing (WDM) technology [7]-[8], has provided huge bandwidth up to terabits to meet the bandwidth requirement of the IP traffic. A lot of research and development works in recent years have been concentrating on finding an efficient way to integrate IP and WDM technologies [9]-[12]. Under such circumstances, Generalized Multi-Protocol Label Switching (GMPLS) was developed and finally standardized by IETF [13]-[16], which extends the concept to support devices that perform switching in time, wavelength and space domain, in addition to packet switching devices. Currently, GMPLS is widely regarded as the most suitable approach that integrates IP and WDM efficiently because it is especially suitable as the control plane of OXCs.

Meanwhile, in order not to let electronic nodes be the bottleneck of the backbone that utilizes IP over WDM infrastructure, all-optical network nodes are getting more and more attentions from academia as well as industry. In the most strict sense, all-optical network transports the traffic from source to destination (here refers to the source edge and destination edge) entirely in optical domain without doing any optical-electrical-optical (OEO) conversion in the intermediate nodes.

Although widely regarded as the most suitable approach that integrates IP and WDM efficiently, GMPLS raises some new issues in the integration of IP and WDM in the all-optical environment. One of these issues is the coupling of the optical characteristics in the physical layer and the IP traffic patterns at the network layer. Figure 1-2 shows the expected technology layer evolution of backbone IP networks from IP over asynchronous transfer mode (ATM) over synchronous optical network (SONET) /synchronous digital hierarchy (SDH) over WDM [17] to IP with GMPLS over SONET/SDH over WDM [18]; and, finally, to IP with GMPLS directly over WDM [15]. The removal of the intermediate layers, such as ATM [19]-[20] and SONET/SDH [21]-[22], implies that the traffic patterns carried in the optical domain and IP domain are correlated [23]. The performance of optical channels is affected by the bursty nature of IP traffic, while the existence of optical impairments along the optical paths and at the optical nodes may also greatly influence the performance of IP traffic. Furthermore, without OEO process at the intermediate nodes, optical impairments cannot be cleaned up, and worse, they get accumulated. Some works have been done on impairment-aware wavelength routing and assignment (RWA) [24]-[25], which try to choose a candidate route with bit error rate (BER) higher than a threshold (e.g., 10^{-12}) for making the traditional assumption of perfect optical transport still valid. However, with the ever-increasing end-to-end distance, data rate per wavelength, number of wavelengths in a fiber, and/or input optical power, we will reach a stage where we have to live with the imperfections in the ultimate fully transparent all-optical network. Hence we shall carry on the analysis or design by counting in the adverse effects of optical characteristics into the performance evaluation.

Figure 1–2: Evolution towards Photonic Networking

Therefore, in our thesis, we are firstly motivated to do the modeling and investigate the performance of the heterogeneous network architecture with QoS control, where traffic coming from different kinds of edge networks (i.e. IntServ edge network or DiffServ edge network) and traversing through the MPLS/GMPLS backbone network. The main purposes of this part of research include: (i) proposing the proper traffic classification and service mapping of the IntServ, DiffServ and MPLS/GMPLS networks that can be used to facilitate the interoperation of IntServ and DiffServ customer networks over MPLS/GMPLS core networks; (ii) proposing a queueing system that can be used at output ports of every modified MPLS/GMPLS LSR/OXC to ensure the desired QoS requirements of different traffic categories in the heterogeneous networks; (iii) modeling mathematically and deriving analytically the QoS performance parameters of the proposed queueing system.

Another motivation of our work is to study the cross-layer issues of the IP over WDM backbone network. The main purposes of this part of research include: (i) illustrating a per-hop model and the detailed node architecture that will be used thereafter; (ii) demonstrating a cross-layer approach for the analysis of coupling effect between the optical layer and network layer in an IP over WDM all-optical network; (iii) Performing comprehensive study of scenarios with various transport links and nodes configurations to show how the difference in physical infrastructures of next generation network (NGN) systems will impact differently on the transmission performance and effective network performance.

1.2. MAJOR CONTRIBUTIONS

Our work concentrates on large – scale backbone network and dedicates for the modeling and performance analysis of the network at such scales. The major contributions of our work can be summarized as follows:

1) We propose a model for analyzing the performance of IntServ and DiffServ traffic over the MPLS/GMPLS backbone network. This model is important to investigate how QoS at the network layer in terms of packet loss probability and delay can be achieved for the hybrid of IntServ and DiffServ traffic across the MPLS/GMPLS transit network.

We propose the IP traffic classification and the service mappings among IntServ, DiffServ and MPLS/GMPLS networks. In this way, the MPLS/GMPLS header is able to incorporate the loss priority scheme and delay priority scheme.

In the proposed approach, each output port of an MPLS/GMPLS enabled IP router includes four output queues representing different delay priorities. Among these queues, the Premium Queue is designed to carry the IntServ Guaranteed Service (GS), the Gold Queue is able to carry the DiffServ high priority traffic, the Silver Queue accepts DiffServ low priority traffic and the Bronze Queue aims to carry the traffic from IntServ Controlled-Load Service (CLS). Multiple thresholds are implemented in the Gold queue and Silver queue for packet loss control.

We also present the detailed model of an optical buffer system that can be used at the output port of an optical packet switching (OPS) node in an all-optical fiber optics – based network. This enables our proposed node not to limit its usage in copper-based network. What encourages us most is that, the proposed optical buffer system can be implemented by the state-of-the-art physical technology of recirculation-enabled fiber delay line (FDLs) in near term. By emulating the electronic queueing system, our proposed optical buffer

system makes its modeling and performance analysis essential for the future OPS network study.

2) Another important contribution of this research project is the model for analyzing the performance of optical layer impairments in the IP/GMPLS over WDM all-optical network. This model is important to derive analytically the performance parameters in terms of BER and Q parameter at the physical layer. The derived BER is used in evaluating the effective performance of the IP/GMPLS over WDM infrastructure. This is also so far the most comprehensive model to analyze the effect of physical layer characteristics to the network layer performance, which includes all of the major linear and nonlinear impairments at both the optical paths and the optical nodes. Also comprehensive study of networks with various fiber link types and system configurations are investigated to make our work more valuable for those who aim to design an all-optical backbone network.

Last but not the least, instead of going all the way to all-optical network for large scale network because of the non-viability of all-optical regeneration, reshaping and retiming (3R), we are inspired by the uniqueness of our numerical findings to propose high-level principles for designing the intermediate-step architecture of NGNs.

1.3. OUTLINE OF THE THESIS

This thesis is organized into five chapters. Chapter 2 provides the relevant background knowledge and critical literature review for our major work presented in Chapter 3 and Chapter 4. Chapter 3 emphasizes the modeling and performance analysis of QoS in IP/MPLS backbone network. Firstly, the fields in MPLS label are illustrated in details. Then the novel traffic classification and service mapping among the IntServ, DiffServ and MPLS are proposed to enable the interoperation of these technologies in such heterogeneous operation environment. More importantly, a four-queue system located at

each output port of the MPLS LSR is proposed. Then the mathematic modeling of the IntServ sources and DiffServ sources are carried out, followed by the modeling of the proposed queueing system. Furthermore, the performance analysis of the system in terms of throughput, loss probability and delay are elaborated. Following that, numerical results obtained demonstrates that the desired QoS can be achieved by using the proposed traffic classification, service mapping, and queueing model. Finally, we further extend copper-based network node to its equivalent optical counterpart in an MPLS/GMPLS based all-optical OPS network. In Chapter 4, a novel cross-layer approach to model and analyze the optical impairments at the physical layer of the IP/GMPLS over WDM backbone infrastructure is introduced considering the linear impairments as well as nonlinear impairments both along the optical paths and at the optical nodes. Moreover, the QoS parameters in terms of BER and Q parameter at the optical layer are derived analytically. Taking into consideration of the BER contributed from the optical layer, the end-to-end performance parameters of the IP/GMPLS over WDM including effective loss probability and effective delay are derived analytically. Lastly, inspired by the uniqueness of our findings, we present guidelines for designing the intermediate-step NGNs. Chapter 5 summarizes the content of this thesis with suggestions for further research.

Chapter 2. Background and Literature Review

2.1. INTRODUCTION

This chapter provides the necessary background knowledge and literature review of related works for the entire thesis. Section 2.2 highlights three major schemes to the interest of QoS study in networks, i.e., IntServ/RSVP suit, DiffServ and MPLS. The characteristics of all of the three QoS schemes are introduced, operations are illustrated, comparisons are made and the advantage of MPLS is emphasized. Then we present a literature review of the efforts in enabling the interoperation of these QoS models. Section 2.3 conducts a literature review of analytical frameworks for statistical multiplexing. Three approaches are examined closely, which are Markov Modulated Poisson Process (MMPP) - based approach, Sriram and Whitt approach, and stochastic fluid flow-based approach. Their merits and issues are discussed. These two sections provide a solid foundation for our work in Chapter 3. Section 2.4 firstly introduces some background knowledge that is necessary for developing our Chapter 4. Subsection 2.4.1 introduces three models of optical network, i.e. opaque, translucent and transparent and highlights the importance of transparency in optical network. Subsection 2.4.2 introduces the major components to be used in Chapter 4 and the related optical impairments produced by these components are illustrated in Subsection 2.4.3. Finally, subsection 2.4.4 presents a literature review of previous work on bit error rate (BER) performance due to optical impairments in the optical network. Section 2.5 emphasizes on the research on the optical packet switching (OPS) nodes. A literature review of optical buffer development is

presented after an introduction of the architecture of the OPS node and major contention resolution techniques. Several schemes are discussed and their pros and cons are mentioned. The current state-of-the-art technologies of the physical layer implementations of optical buffering are also investigated. This is our inspiration to build a realizable optical buffer, which is functionally equivalent to its electronic counterpart that is used in our Chapter 3. Finally, as the extension to MPLS, GMPLS is introduced in section 2.6, which can not only switch in label but also switch in other forms, such as time, space or even fibers. Moreover, the need for considering optical impairments in the GMPLS based WDM transparent networks is highlighted and the latest research approaches and schemes to incorporate optical impairment considerations into the GMPLS are reviewed and compared.

2.2. QOS SCHEMES

The goal of QoS is to provide some levels of predictability and control beyond the current Internet Protocol (IP) “best effort” service. The performance attributes in the QoS metrics generally comprise service availability, delay, delay variation (jitter), throughput and packet loss probability. QoS can be applied to either per flow (individual, unidirectional streams) or per aggregate (two or more flows having something in common). The fundamental principle of deploying QoS is to leave the complexity at the “edges” and keep the network “core” simple. The QoS protocol currently available contains Integrated Service (IntServ) / ReSerVation Protocol (RSVP) suit, Differentiated Service (DiffServ) and Multi-Protocol Label Switching (MPLS) etc. In this section, following the introductions of IntServ/RSVP suit, DiffServ and MPLS, we present a literature review of the efforts in enabling the interoperation of these QoS models.

2.2.1. IntServ/RSVP vs. DiffServ

In year 1994, RFC 1633 defined Integrated Service (IntServ) as a set of extensions to the traditional best effort model of the Internet with the goal of allowing end-to-end QoS to be provided to applications [2]. One of the key components of the architecture is a set of service definitions in addition to best-effort service, which consist of the Guaranteed Service (GS) [26] and Controlled-Load Service (CLS) [27]. GS aims to offer transfer with no loss, low delay and low jitter. CLS approximates the network performance in lightly loaded conditions. IntServ allows resource reservation and Internet Engineering Task Force (IETF) defined the resource reservation protocol (RSVP) [3], which is a setup protocol for quality integrated services on Internet. RSVP is used by a host to request specific qualities of service from the network for particular application data streams or flows. RSVP is also used by routers to deliver QoS requests to all nodes along the path(s) of the flows and to establish and maintain state to provide the requested service. RSVP requests will generally result in resources being reserved in each node along the data path.

The signaling process is illustrated in Figure 2–1. The sender sends a PATH message to the receiver specifying the characteristics of the traffic. The next hop of every intermediate hop is determined by the routing protocol. Upon receiving a PATH message, the receiver responds with a RESV message to request resources for the flow. Every intermediate router along the path can reject or accept the request of the RESV message. If the request is rejected, the router will send an error message to the receiver, and the signaling process will terminate. If the request is accepted, link bandwidth and buffer space are allocated for the flow. Moreover, the related flow state information will be installed in the router.

Figure 2–1: RSVP Signaling

As illustrated in Figure 2–2, IntServ assumes that some explicit setup mechanism is used to convey information to routers so that they can provide requested services to flows. The problems with this architecture are (1) scalability issue due to the per-flow processing, which impedes the deployment of IntServ architecture in the Internet at large; (2) stringent requirement on routers since all routers must have RSVP, admission control, microflow (MF) classification, and packet scheduling [28]. Besides, only a small number of hosts currently generate RSVP signaling. Worse, many applications may never generate RSVP signaling [2].

Figure 2–2: IntServ Architecture

In order to resolve the scalability issue and difficulty in implementation and deployment of IntServ, IETF has developed Differentiated Service (DiffServ) architecture in 1998 [4]. DiffServ provides QoS in IP networks by aggregating traffic based on DiffServ code point (DSCPs). Packets are classified and processed by traffic conditioners at the edge of a DiffServ domain. A DiffServ domain is typically equivalent to an administrative domain, that is, a customer network or a network of an ISP. Service level agreements (SLAs) must be established among the various DiffServ domains. These SLAs form the basis for traffic conditioning actions such as shaping, policing, and remarking at the edge routers.

At each DiffServ router, packets are subjected to a “per-hop behavior” (PHB), which is invoked by the DSCP. The Expedited Forwarding (EF) PHB can be used to build a low latency assured-bandwidth end-to-end service through DiffServ domains [29]. A typical

SLA for such a service might include the ingress and egress points of the DiffServ domain that shall provide the service and a peak rate that can be guaranteed to the traffic stream. The Assured Forwarding (AF) PHB group is a means for a DiffServ provider to offer different levels of forwarding assurance for IP packets received from a customer DiffServ domain. Four AF classes are defined with three levels of drop precedence each [30]. A typical SLA includes rates for low and medium drop priority packets, and might also specify ingress and egress points.

In contrast to the per-flow orientation of IntServ/RSVP, DiffServ network, as shown in Figure 2-3, classifies packets into one of the small number of aggregated flows or “classes”, based on the DSCP field in each packet’s IP header. The primary benefit of DiffServ is its scalability. DiffServ eliminates the need for per-flow state and per-flow processing and therefore scales well to large networks [4]-[5].

Figure 2-3: DiffServ Architecture

IntServ and DiffServ are in fact complementary technologies in the pursuit of end-to-end QoS. In year 2000, RFC 2998 described a framework, where IntServ may be supported over DiffServ networks [28]. Under such circumstance, a network that supports DiffServ should be viewed as a network element in the total end-to-end path.

2.2.2. Multi-Protocol Label Switching (MPLS)

In January 2001, the “Multi-Protocol Label Switching Architecture” was standardized in IETF RFC 3031 [31], which is an evolutionary step of routing/forwarding technology for the backbone of the Internet. MPLS delivers a solution that seamlessly integrates the control of IP routing with the simplicity of Layer 2 switching [32].

2.2.2.1. Separation of Control and Forwarding Components

MPLS solution is composed of two distinct functional components, which are a control component and a forwarding component as shown in Figure 2–4. The control component uses standard routing protocols (Open Shortest Path First (OSPF) [33], (Intermediate System to Intermediate System (IS-IS) [34], and Boarder Gateway Protocol (BGP) [35]) to exchange information with other routers to build and maintain a forwarding table. When packets arrive, the forwarding component searches the forwarding table maintained by the control component to make a routing decision for each packet.

Figure 2–4: Routing Functional Components: Control and Forwarding

2.2.2.2. Label-Swapping Forwarding Algorithm

The forwarding component of MPLS is based on a label-swapping forwarding algorithm. This is the same algorithm used to forward data in ATM and Frame Relay switches. A label is a short, fixed-length value carried in the packet’s header to identify a Forwarding Equivalence Class (FEC). An FEC is a set of packets that are forwarded over the same

path through a network even if their ultimate destinations are different. A label is analogous to a connection identifier, such as the Data Link Connection Identifier (DLCI) of Frame Relay (FR) [39] or the Virtual Path Identifier (VPI) / Virtual Circuit Identifier (VCI) [40] of ATM, because it has only link-local significance, which does not encode information from the network layer header. MPLS needs a protocol to distribute labels to set up label switched paths (LSPs) [36]-[38]. MPLS labels can also be piggybacked by routing protocols. An LSP is similar to an ATM virtual circuit (VC) and is unidirectional from the sender to the receiver. MPLS Label Switched Routers (LSRs) use the protocol to negotiate the semantics of each label, that is, how to handle a packet with a particular label from the peer. A forwarding table indexed by labels is constructed as the result of label distribution. Each forwarding table entry specifies how to process packets carrying the indexing label.

Figure 2-5 illustrates how the traffic is forwarded in an MPLS network. The ingress edge LSR (LER) receives an unlabeled packet with a destination address of 191.3.4.3. LER 1 performs a longest-match routing table lookup and maps the packet to an FEC—191.3/16. LER 1 then assigns an initial label – 7 to the packet and forwards it to the next hop – LSR1 in the LSP. In the MPLS backbone network, LSRs ignore the packet's network layer header and simply forward the packet using the label-swapping algorithm. When the labeled packet arrives at LSR 1, the forwarding component uses the input port number and label to perform an exact match search of its forwarding table. When a match is found, the forwarding component of LSR 1 retrieves the outgoing label, the outbound interface, and the next-hop address from the forwarding table. LSR 1 swaps the incoming label “7” with the outgoing label “2” and directs the packet to the outbound interface for transmission to LSR 3. Similarly, LSR 3 swaps the incoming label “2” with the outgoing label “9” and directs the packet to the outbound interface for transmission to egress LER 3. When the labeled packet arrives at the egress LER 3, the forwarding component searches its forwarding table. Since its next hop is not a label switch, LER 3 removes the label “9” and IP address of the packet – 191.3.4.3 will take effect from now on.

Figure 2–5: Packet Traversing a Label Switched Path

Furthermore, MPLS provides a foundation that supports the deployment of advanced routing services because it solves a number of complex problems: (1) MPLS addresses the scalability issues associated with the existing IP-over-ATM overlay model; (2) MPLS significantly reduces the complexity of network operation; (3) MPLS facilitates the delivery of new routing capabilities that enhance conventional IP routing techniques; (4) MPLS offers a standards-based solution that promotes multivendor interoperability. All these reasons have made MPLS very popular among network and router designers. Companies like Cisco, Nortel, Lucent and Juniper have been developing routers which support MPLS protocols, [41]-[44].

2.2.3. Literature Review of Interoperation of IntServ/RSVP, DiffServ and MPLS

2.2.3.1. Interoperation of IntServ/RSVP and DiffServ

Before the standardization of MPLS, researchers focused on the interoperation of IntServ and DiffServ networks [28] [45]-[47]. In year 2000, RFC 2998 [28], described a framework, where IntServ may be supported over DiffServ networks. They regarded IntServ/RSVP and DiffServ as complementary technologies in the pursuit of end-to-end QoS. They

presented a framework, where end-to-end, quantitative QoS is provided by applying the IntServ model across a network containing one or more DiffServ regions. The DiffServ regions may, but are not required to, participate in end-to-end RSVP signaling for the purpose of optimizing resource allocation and supporting admission control. Together, IntServ/RSVP and DiffServ can facilitate deployment of applications such as IP-telephony, video-on-demand, and various non-multimedia mission-critical applications. IntServ enables hosts to request per-flow, quantifiable resources, along end-to-end data paths and to obtain feedback regarding admissibility of these requests. DiffServ enables scalability across large networks. Two specific realizations of the framework were discussed. In the first one, resources within the DiffServ regions of the network are statically provisioned and these regions include no RSVP aware devices. In the second framework, resources within the DiffServ region of the network are dynamically provisioned and selected devices within the DiffServ network regions participate in RSVP signaling.

In the same year, [47] presented a prototype implementation for the IntServ operation over DiffServ networks under the Linux operating system. To show the rational for the interworking of IntServ and DiffServ, they showed a high-level scenario as shown in Figure 2–6, where DiffServ is used in the core network and IntServ/RSVP is used in the access/edge network. The edge devices, placed at the borders between these domains, play a key role in the interoperation. Around the same time, [45] showed both qualitative and quantitative mapping of loss and delay performance between IntServ and DiffServ.

Figure 2–6: IntServ/DiffServ Interoperation Scenario

2.2.3.2. Interoperation of DiffServ and MPLS

In Jan 2001, RFC 3031 [31] standardized MPLS, which was soon regarded as the most promising candidate to offer new QoS capabilities for large-scale IP backbone networks [5]. Thereafter, QoS-enabled MPLS became a new research topic.

As the first step towards enabling QoS in MPLS network, RFC 3270 [32] in 2002, “Multi-Protocol Label Switching (MPLS) Support of Differentiated Services”, allows the MPLS network administrators to select how DiffServ Behavior Aggregates (BAs) are mapped onto Label Switched Paths (LSPs) so that they can best match the DiffServ, traffic engineering (TE) and protection objectives within their particular networks. Later, [48] presented a policy-based QoS management architecture for DiffServ-aware MPLS-based IP network, where the overall structure and all the sub-modules were described in details. They showed mapping scheme among MPLS service class, EXP field value and DiffServ PHB as shown in Table 2–1. They defined four MPLS classes: Gold, Silver, Bronze and Best Effort. Gold class consists of low loss, low delay and low jitter service for delay-sensitive traffic. Silver and Bronze classes consist of throughput-sensitive traffic. Best Effort class has no guarantee from the network QoS performance.

Table 2–1: DiffServ PHB to MPLS Service Class and EXP Mapping

MPLS Service Class	EXP field value		Drop Precedence	DiffServ PHB
	FCI	DPI		
Gold	11	1 or 0	N/A	EF
Silver	10	1	Low	AF11 or AF21
	10	0	High	AF12 AF13 or AF22 AF23
Bronze	01	1	Low	AF31 or AF41
	01	0	High	AF32 AF33 or AF42 AF43
Best Effort	00	0 or 1	N/A	BE

There are quite a few works on the DiffServ and MPLS operation focusing on different research perspectives, such as path selection in DiffServ-MPLS network [49], packet loss performance for DiffServ-MPLS multicast network [50], quantitative QoS guarantee for

some particular PHB in a DiffServ-MPLS network [51], etc. However, these works did not mention the IntServ aspect at all.

2.2.3.3. Interoperation of IntServ/RSVP and MPLS

On the other hand, some researchers still preferred the IntServ/RSVP QoS model and concentrated on means to interoperate IntServ/RSVP and MPLS. Since an LSR is not a RSVP router, the entire MPLS domain is a non-RSVP region, hence RSVP cannot reserve resource of the MPLS backbone network. In 2002, Franco Tommasi, Simone Molendini and Andrea Tricco proposed an efficient combination of the application-oriented IntServ QoS with the power of MPLS label switching, which defined IntServ-like QoS services in MPLS domains [52]. As shown in Figure 2–7, after the receipt of an RSVP message, which reserves the required resources, the Ingress LSR establishes a Constraint Routed LSP (CR-LSP) towards the Egress LSR. This CR-LSP is endowed with the QoS carried by the RSVP message. The Ingress LSR is responsible to perform message translation to create this CR-LSP. The RSVP sender's traffic crosses the MPLS domain using the CR-LSP. LSRs not on the edge of the MPLS domains will only process Constraint Routed label distribution protocol (CR-LDP) messages.

Figure 2–7: Interoperation of IntServ and MPLS with single ingress LSR

Using the mechanisms defined, a homogeneous end-to-end QoS was achieved. This work also depicted how the scaling issue could be solved by using the Caching of CR-LSPs or Aggregation of CR-LSP. Unfortunately, this research work ignored the DiffServ model just as what section 2.2.3.2 has done to IntServ model.

2.2.3.4. Interoperation of IntServ/RSVP, DiffServ and MPLS

With the introduction of IntServ/RSVP and DiffServ QoS schemes, Internet is now able to support voice and video applications [6]. Hence IP-based traffic volumes have been continuously growing at a fast speed [5]. MPLS, which is able to offers new QoS capabilities for large-scale IP backbone networks, is regarded as a good candidate to enable the interoperation of IntServ/RSVP network and DiffServ network. Together, these mechanisms can facilitate deployment of real-time applications. In backbone networks where MPLS architecture is widely deployed, a combination of DiffServ and IntServ may provide a more robust service to applications [32].

In fact, as early as 2000, Nicolas Rouhana and Eric Horlait have realized the ultimate need to interoperate all of the three technologies – IntServ/RSVP, DiffServ and MPLS [53]. Figure 1–1 has shown us how these different technologies would interoperate in today’s Internet environment. There are two reasons for the emergence of such network structure. Firstly, due to the terrific advantages that MPLS offers, such as low cost and high forwarding speed, it will replace the existing DiffServ backbone network. Consequently, the DiffServ network will be pushed aside to the customer edge. Secondly, the assumption that the edge networks consist of a hybrid of both IntServ architecture and DiffServ architecture is more general than those assuming that the edge customers belong to homogeneous IntServ or DiffServ networks.

Nicolas Rouhana and Eric Horlait have proposed DRUM (DiffServ and RSVP/intserv Use of MPLS), an architecture that delivers end-to-end service guarantees for both DiffServ and IntServ networks, where part of the underlying technology used for IP transport is

MPLS using DiffServ-like mechanisms for QoS provision. However, only qualitative service mappings, as shown in Table 2–2 and Table 2–3, were proposed and the simulation results for a simple scenario demonstrated the effectiveness of the MPLS traffic engineering.

How to implement QoS support for such heterogeneous network architectures, where IntServ clients and DiffServ clients both exist at the edge of the MPLS backbone network, poses new challenges for the design, implementation, modeling and performance analysis. Chapter 3 will demonstrate our achievement in this field of research.

Table 2–2: EXP to MPLS Service Class Mapping

MPLS Service Class	MPLS EXP field value		Drop Precedence
	FCI	DPI	
Network Control Gold	11	1	N/A
	11	0	N/A
Silver	10	1	Low
	10	0	High
Bronze	01	1	Low
	01	0	High
Best Effort	00	1	Low
	00	0	High (less than BE)

Table 2–3: Service Map from IntServ/DiffServ to MPLS

Incoming Service Requirement		MPLS matching service
IntServ	DiffServ	
Guaranteed Service (GS)	Expedited Forwarding (EF)	Gold
Controlled-Load Service (CLS)	Assured Forwarding (AF)	Silver/Bronze
Best Effort	Best Effort	Best Effort

2.3. LITERATURE REVIEW OF ANALYTICAL FRAMEWORKS FOR STATISTICAL MULTIPLEXING

In this section, a literature review of analytical frameworks for statistical multiplexing is carried. Three approaches are examined closely, which are MMPP-based approach, Sriram and Whitt approach, and stochastic fluid flow-based approach. Their merits and issues are discussed. Finally, we choose one from the three frameworks to be used in Chapter 3.

2.3.1. Markov Modulated Poisson Process (MMPP) Approach

The first framework is based on Markov Modulated Poisson Process (MMPP). MMPP is most frequently seen in queueing theory [54] but it has other interesting applications, including environmental, medical, industrial, and sociological research. As shown in Figure 2–8, MMPP is a doubly stochastic process, whose rate varies according to a Markov process [55]. Packet generated by each source is a Poisson process with exponentially distributed length. The transitions between the ON state and the OFF state are controlled by an underlying continuous-time Markov chain [56]. This modulation introduces correlations between successive interarrival times in the process. The MMPP can be identified as a special case of the Markovian arrival process (MAP).

Figure 2–8: Two state MMPP

An MMPP can be classified by the number of states the modulating Markov chain contains. For example, a Markov chain with two states and two intensities is denoted MMPP-2, or sometimes Switched Poisson process (SPP). If the two intensities are equal

the model becomes an ordinary Poisson process. The special case, where one intensity is zero, is called Interrupted Poisson process (IPP). Much work has been done in finding methods to match the characteristics of MMPP-2s and/or IPPs to recorded data or more complex models [57]-[58]. Comparisons of these methods were performed in [59].

The main advantage of using MMPP as traffic model is that it is relatively simpler when doing analysis than other competing approaches. For example, important properties of queueing systems like MMPP/G/1 can, although not very easily, be derived. However, the main disadvantages of the MMPP approach are its slow convergence of the conventional Matrix Geometric methods in conditions of high burstiness and traffic intensity, and the computational complexity in typical applications [60].

2.3.2. Sriram and Whitt Approach

The second approach is based on the work done by Sriram and Whitt [61]. They analyzed a model of statistical multiplexer for packetized voice and data.

Their work characterized the aggregated packet arrival process resulting from the superposition of independent renewal processes, i.e., separate voice streams in this case. But the superposition arrival process makes modeling and analysis of queueing model very difficult, hence they used two-parameter approximations as in the queueing network analyzer (QNA) software package [62]-[63]. With this technique, the superposition arrival process is approximately characterized by two parameters, one representing the average arrival rate and the other the variability.

However, the key issue of the applicability of their work is the questionable assumption for the service discipline. Although they considered packets from both voice and data sources, yet the service discipline they considered was only a FIFO model.

2.3.3. Fluid Flow- based Approach

The third approach is called fluid flow model [64]-[67], in which each individual source is modeled as a Markov Modulated Fluid Source consisting of the ON state and the OFF state. In the ON-OFF process, the ON state is uniformly distributed and the transition between the ON state and the OFF state is controlled by a continuous-time Markov chain which determines the rate of fluid generating. The transport of packets over a transmission link is operated in the same manner.

The fluid flow technique has been applied successfully to a variety of problems in ATM networks, which shows superior ability to handle cells at burst level [68]-[71]. This is difficult for queueing frameworks, such as MMPP, to handle. Furthermore, the numerical complexity of solving fluid flow models with finite buffers does not depend on buffer size while with queueing model the complexity increases as buffer size increases. Finally, the fluid approximation presumes the interarrival time of cells/packets is small with respect to the time between changes in the rate, which is a feature of the high speed backbone network. Several comparative evaluations of techniques for modeling and analyzing statistical multiplexer showed that this approach is effective at the burst level in its accuracy and it is able to solve large systems [72]-[74].

For the above reasons, we choose stochastic fluid flow model for our modeling and analysis of proposed queueing system in Chapter 3 because our network of interest is high speed backbone network with traffic coming from large population of sources. For the mathematic details of the stochastic fluid flow model, please refer to Appendix A [65].

2.4. OPTICAL NETWORKING

In this section, following the introductions network transparency, optical components to be used in Chapter 4 and their associated impairments, a literature review is performed on bit error rate (BER) performance due to optical impairments in the optical network.

2.4.1. Network Transparency

Wavelength Division Multiplexing (WDM) optical networks provide the huge transmission capacity demanded by current and forthcoming applications. A WDM network consists of optical crossconnects (OXC) interconnected by fiber segments. Some OXCs, referred to as nodes in this thesis, are also attached to access stations where data from several end-users could be multiplexed onto a single WDM channel. An access station provides optoelectronic (O/E) conversion and electro-optic conversion (E/O) (in total the OEO conversion) to connect the optical network with conventional electronic equipment. The access station, at an intermediate node, may also be used for signal regeneration on a lightpath. Different kinds of networks are currently identified by their utilizations of OEO devices: opaque, transparent, and most recently, translucent networks, [75]-[77].

2.4.1.1. Opaque Model

A fully – opaque network [78] is characterized by OEO regenerations at every intermediate node between source and destination [79]. Figure 2–9 shows an example of an opaque WDM network node with fixed routing and no wavelength conversion, where E/O conversion is done for every signal added, O/E conversion is done for every signal dropped and OEO conversion is done for every signal passed. Opaque node prevents signal impairments by signal regeneration from accumulating through the network, thus alleviating the tight performance requirements that must otherwise be placed on the network's optical components [80]. But with the improving transmission performance of new generation systems, those regenerations are less and less necessary. Furthermore,

the frequent OEO conversions require a large number of OEO devices greatly increasing the total network cost.

Figure 2–9: Opaque Node, with Fixed Routing, and No Wavelength Conversion

2.4.1.2. Transparent Model

In a transparent network [81], signal bypasses the OEO devices at intermediate node. An example of transparent WDM network node is shown in Figure 2–10. In this transparent node, transit signals from wavelengths 1–4 are allowed to bypass expensive electronic signal processing. This offers potential cost savings over the opaque node of Figure 2–9 [80]. However, the quality of an optical signal degrades as it travels through several optical components along its lightpath from its source to destination due to optical impairments. To overcome these impairments, “long-distance” lightpaths may require signal regeneration, at one or more intermediate locations in the network, to “clean up” the signals [79]. Although it is still very demanding and costly to regenerate signals entirely in optical domain, with the fast pace of research on optical components, we believe that all-optical model will be the ultimate infrastructure of next generation network (NGN). The major work in Chapter 4 will be based on the belief of the ultimate all-optical transparent network.

Figure 2–10: Transparent Node, with Fixed Routing, and No Wavelength Conversion

2.4.1.3. Translucent Model

Translucent optical network is considered as an interesting option for future optical transport network, which does not require OEO regeneration at every intermediate node. In the case of insufficient signal quality or wavelength blocking, OEO regenerations are required [82]–[83]. Nevertheless a translucent optical network requires higher performance devices, to ensure a minimum number of OEO conversions. Although, some researcher viewed the translucent network is probably a good alternative to both fully transparent and fully opaque networks [79], yet we regard it as a short-term compromising infrastructure since the high speed of the network will only be constrained within the islands of transparency and the speed will still be up to the electronic limit when signal has to traverse across the boundaries. Our work in section 4.7 will propose an NGN architecture that will be treated as an intermediate step towards the ultimate all-optical transparent network, based on the translucent model.

2.4.2. Optical Network Components

In this part, we will give a quick introduction of components to be used in Chapter 4, which are three types of single mode fibers, erbium doped fiber amplifier (EDFA) and

transparent optical crossconnect (OXC). The way of introduction is more application-based rather than physics-based. More quantitative studies of the components will be included in modeling and analysis of IP over WDM network transmission systems in Chapter 4.

2.4.2.1. WDM systems

A WDM system uses a multiplexer at the transmitter to join the signals together and a demultiplexer at the receiver to split them apart. WDM systems are popular with telecommunication companies because they are allowed to expand the capacity of the network without laying more fibers. Modern systems can handle up to 160 signals and can thus expand a basic 10 Gbps fiber system to a theoretical total capacity of over 1.6 Tbps [84]-[86].

Here we introduce three types of single mode fibers to be used in Chapter 4, including the important physical parameters and operating regions. Standard single mode fiber (SMF) is the most commonly deployed fiber, which has a simple step-index structure and is optimized for operation in the 1310-nm band. It has a zero-dispersion wavelength at 1310 nm and can also operate in the 1550-nm band, but it is not optimized for this region. The typical chromatic dispersion at 1550 nm is as high as 17 ps/(nm-km); hence dispersion compensation must be employed for high-bit-rate applications. The attenuation parameter is typically 0.25 dB/km at 1550 nm, and the effective area is $80 \mu\text{m}^2$, [87]. Dispersion shifted fiber (DSF), on the other hand, is a type of single-mode optical fiber with a core-clad index profile tailored to shift the zero-dispersion wavelength from the natural 1300 nm in silica-glass fibers to the minimum-loss window at 1550 nm. This is advantageous because it allows a communication system to possess both low dispersion and low attenuation [88]. However, when used in WDM, DSF can suffer from four-wave mixing (FWM) which causes intermodulation of the independent signals. In addition, the typical attenuation of DSF is 0.2 dB/km and the effective area is as small as $55 \mu\text{m}^2$, [87]

[23]. Finally, dispersion compensating fiber (DCF) is used to compensate for dispersion for a network. Because typical DCF has negative dispersion and produces a negative slope, with nominal values that are typically much higher than those of the fiber to compensate for [87]. At 1550 nm, the standard SMF has dispersion around 20 ps/(nm.km) and the DCF has dispersion with value -100 ps/(nm.km). Hence as a rule of thumb, the length of the DCF is 1/5 of the standard SMF it is to compensate for. In addition, the typical attenuation of DCF is 0.4 dB/km and the effective area is $25.8 \mu\text{m}^2$, [89].

2.4.2.2. Optical Amplifier

An optical amplifier is a device that amplifies optical signals directly, without the need to first convert it to an electrical signal. Compared with other types of optical amplifiers, such as semiconductor optical amplifier (SOA) [88] and Raman amplifier [90], Erbium-doped fiber amplifier (EDFA) is the most deployed fiber amplifier as its amplification window coincides with the third transmission window of silica-based optical fiber [91]. Two bands have been developed in this window – the Conventional, or C-band, from approximately 1525nm to 1565nm, and the Long, or L-band, from approximately 1570nm to 1610nm. As its name implied, EDFA is a fiber heavily doped with Er^{+} to act as an active medium. External energy is delivered optically by pumping. Figure 2–11 shows how an EDFA works. Signal lights and pump light are put in the erbium doped fiber by a WDM coupler. As they propagate through the fiber section, the information signals get amplified and the pumping signal loses its power. Since EDFA has a relatively wide gain bandwidth, it can amplify many wavelengths simultaneously but it amplifies signals as well as noises [92].

Figure 2–11: Erbium Doped Fiber Amplifier (EDFA)

2.4.2.3. Optical Crossconnect (OXC)

OXC is a device used by telecommunications carriers to switch high-speed optical signals in a fiber optic network. There are three ways to realize an OXC. One can implement an OXC in the electronic domain as shown in Figure 2–9. This is known as an OEO OXC design which can be found in opaque network with a key limitation: the electronic circuits limit the maximum bandwidth of the signal. The second approach is to switch optical signals in an all-optical device, as shown in Figure 2–10. Such a switch is often called a transparent OXC, which keeps the features of data rate and protocol transparency. However, because the signals are kept in the optical format, this architecture does not allow easy optical signal quality monitoring. The last approach is a compromise between opaque and transparent OXCs – translucent OXC, as shown in Figure 2–12. Optical signals passing through the switch stage can be switched either by the optical switch module (O–OXC) or the electronic switch module (E–OXC). In most cases, the optical switch module is preferred for the purpose of transparency. When the optical switch module's switching interfaces are all busy or an optical signal needs signal regeneration through an OEO conversion process, the electronic module is used. But the criterion is sometimes difficult to define and open to debate.

Figure 2–12: Translucent Node, composed of Transparent Section (O) for Passing Through Signals and Opaque Section (E) for Adding/Dropping and Regenerating Signals

We will make use of the transparent OXC in Chapter 4, since we are interested in all-optical transparent network. Please note that the OXCs introduced here do not contain optical buffers. If the optical packets enter these OXCs and fall in the contention to the same output, they will either be converted into electrical domain then stored electronically (at nodes with OEO conversion capability), or be dropped (at nodes with all-optical processing).

2.4.3. Optical Impairments

In this part, we will give a brief introduction of optical impairments to be considered in Chapter 4, including both linear and nonlinear impairments.

An effective optical transmission system may make use of some of the components we introduced in section 2.4.2, where WDM is used as a cost-effective multiplexing technique to increase the bandwidth, OXCs are likely to emerge as the preferred option for switching multi-gigabit or even terabit data streams, and EDFA at strategic locations in the network are used to compensate for the progressive losses incurred [93]. In the end, this results in a simpler, more cost-efficient network that will transport a wide range of data streams and very large volumes of traffic. Unfortunately, the WDM channels, OXCs and EDFAs, while offering enormous capacity, transparent switching and loss compensation, respectively, may introduce noticeable impairments especially when transmission rate is greater than 10Gbps.

2.4.3.1. Linear Impairments

Attenuation

Transmission of light by optical fiber is not 100% efficient. The reduction of signal strength or light power over the length of the light-carrying medium is known as fiber attenuation, which is measured in decibels per kilometer (dB/km). Attenuation is caused

by several different factors, but primarily scattering and absorption. Attenuation in optical fiber is very low, as compared to other transmission media (i.e., copper, coaxial cable, etc.), with a typical value of 0.35 dB/km at 1300 nm for standard SMF. Attenuation at 1550 nm is even lower, with a typical value of 0.25 dB/km. This gives an optical signal, transmitted through fiber, the ability to travel more than 100 km without regeneration or amplification.

Chromatic Dispersion (CD)

Dispersion is due to frequency dependence of fiber refractive index. It is the time of flight differences of different components of that signal, typically resulting in pulse broadening. CD consists of two kinds of dispersions, which are balanced to produce a wavelength of zero dispersion shifting within the 1310 nm to 1650 nm operating window. In digital transmission, dispersion limits the maximum data rate, the maximum distance, or the information-carrying capacity of a single-mode fiber link. CD varies with wavelength and is controlled by fiber design. The wavelength at which dispersion equals zero is called the zero-dispersion wavelength (λ_0), where fiber has its maximum information-carrying capacity, e.g., $\lambda_0 = 1310\text{nm}$ for standard single mode fiber.

Amplifier Spontaneous Emission (ASE) Noise

The dominant source of noise in the EDFA is ASE, which is produced over a range of wavelengths exactly corresponding to the gain spectrum of the EDFA. ASE is light, produced by spontaneous emission, that has been optically amplified by the process of stimulated emission in a gain medium. In a cascaded system, ASE accumulates over many EDFAs. Because the optical signal is not converted back to the electrical domain until the end of the chain, the ASE noise continues to grow. As the ASE level grows, it begins to saturate optical amplifiers and reduce the signal gain. The net result is the drop of signal level and increase of ASE level along the transmission line, thus degrading signal-to-noise ratio (SNR) considerably at the receiver. For good performance, the signal should be at least 30 dB higher than the amplified noise.

Crosstalk

Crosstalk occurs in an OXC when a portion of a signal leaks into another signal as two or more optical signals copropagate through the switch fabric. In practice, the cumulative effect of homowavelength crosstalk in the space switches causes the dominant impairment as compared to its heterowavelength counterpart in multiplexing and demultiplexing devices [94]. The ratio of the power at the unselected output port over the total input power in a switch element is referred to as the crosstalk ratio of the switch element (χ). Consider the 2x2 switch shown in Figure 2–13. Suppose a connection has been set up in the switch between input port 2 and output port 1. Now a new connection is established between input port 1 and output port 2 and a portion of the optical power on the existing call is found to mix with the new connection at the combiner stage (depicted as the dashed lines). In this architecture, the crosstalk must occur in a single splitter element and in a single combiner element in order for an existing signal to interfere with the new signal. This is referred to as the first-order crosstalk effect and it results in a fraction of the interfering signal mixing with the new signal [95]. The first-order switch crosstalk ratio (χ_{sw}) is the square of the crosstalk ratio of the individual element, hence $\chi_{sw} = \chi^2$. This is because the interfering signal leaks into the unselected port at two intermediate switch elements before it encounters the new signal. Higher order crosstalk terms (χ^3, χ^4 , etc.), whose effects are negligible compared to the first-order terms, are not considered in this thesis.

Figure 2–13: Crosstalk in a 2x2 Switch

2.4.3.2. Nonlinear Impairments

Nonlinear electromagnetic phenomenon occurs when the response of a medium is a nonlinear function of the applied electric and magnetic field. One of the principal factors limiting system reach is the occurrence of nonlinear distortion, which decreases the optical signal-to-noise ratio (OSNR) at the receiver. The power penalty associated with nonlinear crosstalk can be represented by the increase in signal power required to restore the OSNR to its value in the absence of nonlinearities. However, increasing signal power has diminishing returns since higher signal power generates even more crosstalk, which could further degrade OSNR at the receiver. This also increases the overall system cost by requiring higher power amplifiers, for example. A common means of measuring signal robustness is to observe the Q-factor of the received signal. The Q-factor is directly related to the OSNR and signal bit-error rate (BER).

Four Wave Mixing (FWM)

FWM originates from the phenomenon that the polarization, induced in the electric dipoles of a medium by an electric field, E , is nonlinear in E . The third-order parametric process results in the lower order nonlinear effects in optical fibers. Among them, four-wave mixing (FWM) is the most influential factor for WDM network design. In multi-channel systems, a signal channel suffers from FWM effect, which generates various combinations of different frequencies and causes system degradation. Figure 2-14 schematically illustrates the optical frequency arrangement, where f_i , f_j , and f_k are the signal light frequencies, and f_{FWM} is the FWM lightwave frequency satisfying $f_{FWM} = f_i + f_k - f_j$. ([96] – [103])

Figure 2-14: Optical Frequency Arrangement of FWM

Self-Phase Modulation (SPM)

An interesting manifestation of the intensity dependence of the refractive index in nonlinear optical media occurs through SPM, a phenomenon that leads to spectral broadening of optical pulses. SPM is the temporal analog of self-focusing. The intensity dependence of the refractive index leads to SPM-induced nonlinear phase shift, resulting in chirping and spectral broadening of optical pulses. Clearly, SPM can affect the performance of lightwave system considerably. [104]-[106]

Cross-Phase Modulation (XPM)

When two pulses of different wavelengths propagate simultaneously inside optical fibers, their optical phases are affected not only by SPM but also by XPM. XPM is always accompanied by SPM. It occurs because the effective refractive index seen by an optical beam in a nonlinear medium depends not only on the intensity of that beam but also on the intensity of other copropagating beams. Fiber dispersion converts phase variations into amplitude fluctuations, affecting the SNR considerably. The XPM effect is quite important for WDM lightwave systems since the phase of each optical channel is affected by both the average power and the bit pattern of all other channels [107]-[109]. A proper understanding of the interplay between XPM and dispersion is important for WDM systems. [110]-[112]

2.4.4. Literature Review of BER Performance Due to Optical Impairments

Firstly, let us look at the BER due to the impairments along the optical path. Along the fiber path, linear impairments to be considered are group velocity dispersion (GVD) and fiber attenuation. We ignore polarization mode dispersion (PMD) because it can be compensated at the destination [122] with today's advance of optical layer technology.

On the other hand, the most significant fiber nonlinear effects include four-wave mixing (FWM), stimulated Raman scattering (SRS) and stimulated Brillouin scattering (SBS) in a WDM system design. However each of these effects manifests itself in a unique way. FWM appears at the lowest input power level in a DSF system, and causes cross-channel interference. In WDM system with tens of multiplexed channels, the launch power is restricted by FWM [113]. Whereas SRS causes power depletion in the lower wavelength channel and is the limiting factor of the launch power in a system with hundreds of channels [96] [114]-[116]. The launch power for individual channel is limited due to SBS [117]. Since the focus of our thesis is on WDM system with tens of channels, FWM effect is of our interest.

In literature, [118] studied the BER due to FWM effect. The output power P_{FWM} of the four-wave mixing product is given by [119],

$$P_{FWM}(f_i, f_j, f_k) = \frac{1024\pi^6}{n^4 \lambda^2 c^2} d^2 \chi^2 \frac{P_i P_j P_k}{A_{eff}^2} e^{-\alpha L} \frac{(1 - e^{-\alpha L})^2}{\alpha^2} \eta$$

where P_i , P_j and P_k are the input power of the frequencies f_i , f_j and f_k respectively.

P_{FWM} is the power of the signal from the FWM effect at the frequency f_{FWM} . n is the fiber refractive index, λ is the wavelength, A_{eff} is the effective mode area, α is the fiber attenuation coefficient, d is the degeneracy factor, χ is the third-order nonlinear susceptibility, and η is the mixing efficiency. The FWM light, detected at the receiver together with the signal light, induces the FWM noise with the power $N_{FWM} = 2b^2 P_s P_{FWM} / 8$, where P_s is the signal power and b is the quantum efficiency [120].

In [121], the FWM effect and BER performance in WDM systems were closely examined. Their work investigated the characteristics of network performance when IP traffic is directly mapped into WDM network with MPLS technique. They showed the effect of FWM and the effect of burstiness of IP traffic on the network performance.

Furthermore, due to high data rate per channel of the WDM system (e.g., 10Gbps and 40Gbps), nonlinear effects of phase shift due to self phase modulation (SPM) and cross phase modulation (XPM) become more and more apparent.

Next, let us focus on the BER performance due to impairments at the optical nodes. Two major impairments have been the focus of investigation, i.e., ASE noise and crosstalk.

In [123], noise and BER characteristics of lightwave systems with optical amplifier (e.g., EDFA) were calculated. The spontaneous emission power at the output from an optical amplifier has been given by $P_{sp} = N_{sp}(G-1)h\nu B_o^2$, where G is the optical gain, $h\nu$ is the photon energy, B_o is the optical bandwidth, and N_{sp} is the spontaneous emission factor for EDFAs. For an ideal amplifier, $N_{sp} = 1$, while for a non-ideal amplifier, $N_{sp} > 1$. Furthermore, noise terms, including signal to ASE spontaneous beat noise, ASE spontaneous to ASE spontaneous beat noise, were shown.

On the other hand, [95] reviewed and compared some architectures to be used as optical space switch fabrics and presented considerations involved in selecting the appropriate architecture. It showed that type I multiple substrate point-to-point architecture had strictly non-blocking characteristics and much better insertion loss characteristics for large switch dimensions compared with some other architectures they investigated. Insertion loss of this switch architecture is $(2\log_2 N)L + 4W$, where N is the number of input, L is the switch element loss, and W is the waveguide loss. [124] investigated the impact of crosstalk in an M by M arrayed-waveguide and obtained the relationship between crosstalk and power penalty. They have also derived signal to crosstalk beat noise and the BER performance due to the crosstalk effect.

In [120], researchers studied the WDM network scale, considering both the ASE noise and FWM noise. In their system, the minimum input power into a repeating node was

dependent on the ASE noise, while the maximum power launched into a fiber was restricted by FWM effect. They derived a simple expression to describe the possible number of channels and transmission distance in the WDM system. Reference [125] gave the node modeling and the enumeration process of a DSF system, with much lower channel rate (i.e., 1Gbps per channel). Hence while authors of [125] considered the impairments at the node consisting of the first order homowavelength crosstalk at the OXC and the ASE noise at the EDFA, the only impairment along the fiber path they considered is the fiber attenuation.

However, in Chapter 4, we are interested in the systems operating at much higher data rate (i.e. 10Gbps or 40Gbps per channel). Therefore, fiber nonlinearities (i.e., SPM, XPM and FWM) must be included in our study due to the high data rate and small channel spacing. As a result, for our transmission system studied, we will consider all the major linear and nonlinear optical impairments that will be introduced by the optical nodes and optical paths, which include phase shift due to SPM, XPM and noises due to FWM, crosstalk, and ASE.

2.5. OPTICAL PACKET SWITCHING (OPS) NODE

The favorite concept of broadband optical Internet urges the appearance of optical packet network, which can provide flexibility and efficiency as the electronics-based Internet but with tremendous improvement upon the capacity. In an all-optical transparent network, user data, transmitted in optical packets, are switched within each optical packet switch entirely in the optical domain. Thus, the user data remains as an optical signal in the entire path from source to destination. No O/E or E/O conversions are required. A major obstacle faced in building all-optical OPS networks is, however, the difficulty in implementing all-optical packet switching nodes. In a packet-switched network each packet has to go through a number of switches to reach its destination. When the packets

are being switched, contentions might occur due to limited resources. How contention is resolved greatly affect the network performance.

In this section, following an introduction of OPS node architecture and major contention resolution techniques, we present a literature review of optical buffer development. The current state-of-the-art technologies of the physical layer implementations of optical buffering are also investigated. Please note that the nodes investigated in this section involve contention resolution method, whereas the OXCs covered in section 2.4.2.3 do not involve such process.

2.5.1. OPS Node Architecture

OPS can handle fixed-length synchronous data cells as well as variable-length asynchronous data packets. A generic WDM optical packet switch consists of four parts as shown in Figure 2-15: the input interface, switching fabric, output interface, and control unit. The input interface is mainly used for packet delineation and alignment, packet header information extraction and packet header removal. The switch fabric is at the core of the switch and is used to switch packets optically. The output interface is used to regenerate optical signals and insert packet headers. The control unit controls the switch using the information in the packet headers. When a packet arrives at a WDM optical packet switch, it is first processed by the input interface. The header and payload of the packet are separated. The header is converted into the electrical domain and processed by the control unit electronically. On the other hand, the payload remains an optical signal throughout the switch. After the payload passes through the switching fabric, it is recombined with the header, which is converted back into the optical domain at the output interface.

Figure 2-15: Generic node architecture in OPS network

2.5.2. Major Contention Resolution Technologies

In an OPS network, contention to the same output port is inevitable for optical packets traveling through and how the contention is resolved has a great effect on network performance. There are mainly three different methods for contention resolutions: optical buffering, wavelength conversion, and deflection routing.

2.5.2.1. Optical Buffering

In electronic routers, contention is usually resolved by a store-and-forward technique by using electronic random access memory (RAM), which means that the packets in contention are stored in a queue and sent out one by one. However, without a ready-to-use optical RAM available, people have to find other approaches to temporarily store packets optically. Optical buffering is certainly a solution to hold the contending packets temporarily until later [126]-[135].

An all-optical buffer is defined as shown in Figure 2-16. The output data $f(L, t)$ is essentially a copy of the input data $f(0, t)$. $f(L, t)$ is equal to the multiple of a proportionality constant p with a time delay τ , which is variable controlled by an

external source V . The delay introduced by the optical buffer can be represented by $\tau = L/v_g$, where L is the length of the fiber being used as the optical buffer and v_g is the speed of the light traversing the fiber-based optical buffer medium.

Figure 2–16: Definition of Optical Buffer

Optical buffering was, to a great extent, inspired by its conventional electronic network counterparts. In electronic networks, the link bandwidth is much less than the capacity of today's optical fiber, and much effort were put into increasing link utilization. In a network deploying optical buffers, each packet is guaranteed to arrive at its destination along the shortest possible path, and for a given connectivity the expected number of hops is minimized. Implementing optical buffers involves a great amount of hardware and complex electronic controls. Another issue that arises with optical buffers is that the optical signal suffers from power loss in the delay lines, so optical amplifiers are often used. The accumulated noise from the cascaded amplifiers can severely limit the network size at very high bit rates, unless signal regeneration is applied.

2.5.2.2. Deflection Routing

Deflection routing is ideally suitable for switches that have little buffer space. In deflection routing, as the name implies, contention is resolved as follows: if two or more packets need to use the same output link to achieve minimum distance routing, only one will be routed along the desired link, while others are forwarded on paths which may lead to greater than minimum distance routing. Hence, for each source–destination pair the

number of hops taken by a packet is no longer fixed. Deflection routing does not necessarily exclude the use of optical buffers. However, the deflected packet may end up following a longer path to its destination. As a result, the end-to-end delay for a packet may be unacceptably high. Also, packets will have to be reordered at the destination since they are likely to arrive out of sequence [136]-[138]. One major problem of deflection routing is that it tends to collapse under the asynchronous operation or high load operation [139]. To solve this problem, deflection routing still needs to resort to optical buffering for improvement of performance [140].

2.5.2.3. Wavelength Conversion

Optical buffering and deflection routing could be regarded as deflection in general, one in the time domain and the other in the space domain. With today's enabling technology in WDM, the wavelength domain presents one more dimension of solution. Furthermore, wavelength conversion provides noise suppression and signal reshaping. This can help to minimize external blocking as follows. Let us assume that two packets are destined to go out of the same output port at the same time. Then they can be still transmitted out, but on two different wavelengths. This method may have some potential in minimizing external blocking, particularly since the number of wavelengths that can be coupled together onto a single fiber continues to increase.

Wavelength conversion combined with optical buffering can be incorporated in both synchronous as well as asynchronous network. Compared with a pure FDL-based optical buffer, wavelength conversion can reduce the number of optical buffers or reduce packet loss probability. Deflection routing with wavelength conversion becomes an interesting option for mesh topologies, such as ShuffleNet [141], when the nodes are provided with a number of optical receivers/transmitters equal to the number of wavelengths. People in favorite of this approach argue that in a network with a large number of wavelengths and full wavelength conversion, buffers may not be necessary.

In conclusion, both buffering and deflection have their advantages and disadvantages: buffering offers better network throughput but involves more hardware and controls; deflection is easier to implement, but cannot offer ideal network performance. When combined with wavelength conversion, their disadvantages could be overcome or minimized, therefore giving the network designer more choice and flexibility.

2.5.3. Literature Review of Optical Buffering Techniques

So far, there are two options to build an optical buffer: slow light –based (SL–based) [142]–[144] and fiber delay line – based (FDL – based) [145]–[147]. Both of them are based on fiber optics and aim to introduce a time delay τ to the optical packet that passes through.

2.5.3.1. *Slow Light-based Optical Buffer*

SL–based optical buffers make use of the dispersion property of optical fiber to slow down light (i.e. by decreasing v_g). The optical buffer built on SL technology outperforms FDL–based buffer in terms of much smaller physical size. Also, speed slowing effect varies for different wavelengths. Hence, for the same length of fiber, effective buffer lengths created by different wavelengths are different accordingly. It may contribute to more flexible design and control of the final optical buffer. However, SL–based buffer suffers from the property of constant delay – bandwidth product [148]. E.g., for a 12.6GHz system, the maximum delay introduced is 47ps, which cannot accommodate even a single optical packet in the buffer [149]. For a delay generated around 46ns, the maximum speed of the channel is found to be 40 MHz, which is well below today’s requirements on channel capacity [150]. So unless the delay–bandwidth product issue can be overcome, the SL–based optical buffer can hardly find a role in storing high–speed optical packets (in the order of Gbps transmission rate).

2.5.3.2. FDL-based Optical Buffer

In FDL-based optical buffers, contending packets are temporarily stored by a span of fibers attached to the switching node. The time that a packet can be stored is decided by the length of FDL and the times that the packet loops around the FDL. The techniques of optical buffering in the optical packet switch can be classified according to the position of the buffers, which are output buffering, input buffering and recirculation buffering.

Output buffering [151]-[152] consists of a space switch with an optical buffer at each output port, as shown in Figure 2-17 (a). In each time slot, zero or more packets with the same destination port arrive at the switch and they will be placed in the output buffer attached to that particular destined port simultaneously. If a packet arrives at the buffer, that is fully occupied, the packet will be dropped and the packet will be considered as lost. Theoretically speaking, output buffer achieves excellent performance in terms of loss rate, throughput and delay. However, it is difficult to construct a high-speed switch with large bundles of output buffers. Under the assumption of uniform Bernoulli traffic, a packet loss of 10^{-11} and a load of 0.8 implies a buffer depth requirement of 55 packets per output, whereas a load of 0.9 with the same packet loss requires a buffer depth of 110 packets.

Input buffering, as shown in Figure 2-17 (b), consists of a space switch with buffers at the input ports, hence there is no need for high-speed switching matrix [152]-[153]. However, input buffering suffers from head-of-line (HOL) blocking, which limits the throughput to a maximum of 58% for uniform traffic [151]. Specifically, a packet not at the head of buffering line could be blocked even if its destined output port is available. Because of its poor performance, input buffering is a never good choice for purely optical implementation. It is possible to improve the performance by selecting packets other than those at the head of a buffer to go to the output, but this is too complex for optical implementation. A combination of input-output queues is the well-known solution to achieve satisfactory performance by a switching matrix operating faster than data transmission speed but not as fast as that in an output-buffered switch [154].

In recirculation buffer, as shown in Figure 2–17 (c), a number of FDLs are installed to feedback from the outputs of the space switch to its inputs. If two or more packets arrive at the input for a particular output port, all but one are placed into the FDL recirculation loops and the contending packets then will be fed back to the input port to be handled together with the possible new arrivals. When implemented optically with unity delays, where the FDL only delay a single packet, many recirculations are required, implying high loss and accumulation of amplifier noise in the loops. For example, in a 64x64 switch, 237 recirculation loops are required for a loss of 10^{-10} if the load is 0.8 [152].

Figure 2–17: Optical buffering in an OPS node

Our work in section 3.9 makes use of output buffering technique due to its superior performance and the output optical buffer is able to be designed in such a way that it emulates the QoS controlled electronic queueing system proposed in Chapter 3.

Next, let us investigate the actual development of FDL technology. FDL can be of traveling type, where each packet can go through an FDL at most once, or of recirculation-enabled type, where packets can loop around FDLs multiple times.

The traveling type FDL-based optical buffers introduce the delay (τ) purely by manipulating the length of fiber (L), i.e. by increasing L . Hence it suffers the problem of bulky size and loss [134]. E.g., in order to achieve 100 Tbps Router with 2500 ports at

40Gbps, 250 ms buffering per port, the total length used for the FDL is about the distance between the earth and the sun.

The recirculation-enabled FDL optical buffer is an improvement of this basic traveling type FDL-based setup, which greatly reduces the length of the fiber for the same amount of delay required [135] [155]-[157]. In the actual implementation, if the recirculation circuit is purely passive, the limit of the recirculation will depend on the loss of the signal during recirculation process. To overcome the limit of the signal loss during the recirculation process, EDFAs can be used to recover the signal strength. However, in this case the limit of the recirculation cycles will be decided upon the noise accumulation by the EDFAs. Fortunately, the number of cycles, which the optical packets can be recirculated before being degraded below the level of acceptance, keeps improving throughout the years [127][135] [158]-[159].

In literature, there were some FDL-based optical buffer structures proposed and analyses of those buffers were carried. In [160]-[162], modeling and performance analysis of their optical buffer of interest were shown. In their work, FDLs introduce delays that are consecutive multiples of D . In a buffer made with B FDLs, the first introduces a delay of $0\mu s$, the second of $D\mu s$, the third of $2D\mu s$ and the B th a delay of $D_B = (B-1)D\mu s$. If packet arrives at time t and server is free at time t_f , the new packet is going to be

delayed by an amount $\Delta = \left\lceil \frac{t_f - t}{D} \right\rceil D$, where $\lceil x \rceil$ indicates the smallest integer greater

than or equal to x , due to limited time resolution related to the delay granularity D . However, the hardware requirement of their optical buffer can hardly be achieved based on today's technology or even near-future technology. E.g., in the performance analysis of these works, delay granularity D was taken as a measurement, which ranged from 0 to 1. This means their FDL or recirculation FDL should be able to delay part of a packet. However, in physical layer implementation, reading out and writing in the same optical

packet from one round of FDL at the same time is not possible. Furthermore, current research is having problem of storing multiple optical packets in one buffer at one time because of the unsolved addressing issue, [163].

Hence the state-of-the-art optical buffer that can be utilized is recirculation-enabled FDL-based buffer, where one cycle of the FDL should support the maximum size of optical packets in the network. This will be used in our work in section 3.9.

2.6. GENERALIZED MULTI-PROTOCOL LABEL SWITCHING (GMPLS)

As mentioned previously, WDM is a multiplexing technique that can efficiently accommodate various data traffic flows at the optical layer. Originally, the provisioning problem in WDM networks was solved at the management plane (MP) level by an operator-driven centralized approach. More recently, the introduction of the optical control plane (CP) allowed the simplification and automation of provisioning in WDM networks. In the context of wavelength routed networks, the ITU-T Automatic Switched Optical Network (ASON) standard [164] describes the set of CP components that are to be used to manipulate transport network resources in order to provide the functionality of setting up, maintaining, and releasing optical connections. GMPLS is an implementation of the CP that has been developed by the IETF to facilitate the establishment of label switched paths (LSPs) [165], involving signaling, routing, and resource management functions and protocols. It is a multipurpose control plane paradigm that supports not only devices that perform packet switching, but also devices that perform switching in the time, wavelength, and space domains [13]-[16].

2.6.1. Merits and Issues

GMPLS based IP over WDM system has several advantages: (1) Simplified management plane. All switching devices are managed by a single control plane [165], which results in reduced operational costs by solving the scalability issues associated with the overlay model. (2) Simplified protocol stack in data plane. IP traffic are directly mapped into underlying WDM channels without any intermediate protocol layers [15], such as ATM or SONET/SDH, which enables much more efficient usage of the optical bandwidth. (3) Flexible implementation. Under the framework of GMPLS, different deployment models can be used to meet different service requirements [5].

Despite all the merits GMPLS brings along, however, it also raises some new issues in the integration of IP and WDM. One of these issues is the coupling of the physical layer characteristics and the IP layer traffic pattern [23]. The removal of intermediate layers implies that the traffic patterns carried in the physical layer and network layer are correlated. In a transparent WDM networks, a proper resource allocation requires that the presence of physical impairments in routing and wavelength assignment (RWA) and /or lightpath provisioning be taken into account.

2.6.2. Literature Review of Impairment-Aware Routing and Signaling

We assume that full retiming, reshaping, and reamplification (3R) of optical channels is performed at the edges of transparent domains (also known as “islands of transparency”). For the considered all-optical scenario, there are two families of approaches for impairment-aware RWA when provisioning optical connections with stringent requirements on signal quality: (1) centralized approach, where the RWA is centralized while the wavelength provisioning can be either centralized or distributed; (2) distributed approach, where both RWA and provisioning are based on GMPLS protocol suite

extensions. In the centralized approach a single element, which is reachable by all the network elements (NEs) within the transparent domain, plays the role of a server wherein all the connections to be established are served [166]. This element is aware of the complete network topology, resource availability, and physical link (PL) parameters in a central repository known as the Traffic Engineering Database (TED). In this approach the performance monitoring system updates the TED information with the complete set of performance parameters within the entire domain that is to be considered by the routing computation process. In this way, this central element is able to guarantee and satisfy the specific set of lightpath requirements (e.g., bandwidth, diversity, optical signal quality, latency, etc.) requested by a particular NE. In the distributed approach, each NE's optical control plane (OCP) is responsible to set up lightpaths within the transparent domain. For this purpose, a common and distributed GMPLS-based control plane manages the required procedures for establishing end-to-end lightpaths (i.e., routing and signaling). Based on these two families of approaches, researchers are working on four types of schemes. Scheme (A) is shown in Figure 2–18 which utilizes a centralized RWA and a centralized signaling approach at MP. In this case, the network management system (NMS) is responsible for the administration, computation, and provisioning functionalities of the whole network [166].

Figure 2–18: Scheme (A) centralized RWA+ centralized signaling

As shown in Figure 2–19, scheme (B) makes use of a centralized RWA approach at MP and a distributed signaling approach at CP. In such an approach, all OCP modules send queries to the central path calculation element (PCE), with which all of them can communicate for route computation issues. Whenever the PCE receives a path computation request, it computes the required path, taking into account the TED information and the on-demand connection requirements. Then, the computed route is converted into an explicit route, which is returned to the source OCP for establishing the lightpath by means of a distributed signaling protocol [167]. Both scheme (A) and (B) belong to the centralized approach, where a server with the knowledge of the complete network topology, resource availability and physical parameters in the TED is present.

Figure 2–19: Scheme (B) centralized RWA + distributed signaling

On the other hand, scheme (C) in Figure 2–20 and scheme (D) in Figure 2–21 are the two alternatives of the distributed approach. In scheme (C), some extensions to the current GMPLS routing protocols (e.g., OSPF–Traffic Engineering (OSPF – TE), IS – IS – TE) need to be implemented with the aim of carrying PL information besides TE attributes. Thus, a global TED is built at each OCP, which maintains an up-to-date picture of both resource utilization and physical parameters concerning any link. When the connection request arrives to the source NE, first the RWA algorithm is executed, taking into consideration its own TED and the potential constraints required by the connection. The

on-line RWA algorithm computes a route considering network performance objectives (e.g., efficient network resource utilization) and the required end-to-end optical signal quality. The computed route is then passed to GMPLS signaling protocol (e.g., RSVP with traffic engineering (RSVP-TE) and Constraint – Routed LDP (CR – LDP)), which allocates the resources to set up the requested lightpath. If the requirements cannot be satisfied, the connection is blocked [168]. In scheme (D), existing GMPLS signaling protocols such as RSVP – TE are extended to perform on-line dynamic estimation of the signal quality during the lightpath setup process without the need of modifying routing protocols [24]. This mechanism allows guaranteeing the required signal-quality at the destination node along the complete route.

Figure 2-20: Scheme (C) distributed approach with modification to routing protocols

Figure 2-21: Scheme (D) distributed approach with modification to signaling approach

As a summary, Table 2–4 presents advantages and drawbacks of the four schemes for impairment–aware lightpath setup [169]. Since the distributed approaches fall entirely within the scope of the ASON/GMPLS control plane, GMPLS is the best candidate to facilitate the distributed impairment–aware lightpath setup in IP over WDM transparent network. Moreover, our work in section 4.7 makes use of scheme (D) after comparing the four schemes. Because scheme (D) is found to be the most cost–effective and performing scheme because of its ability to handle dynamic traffic, scalability and stability. Also, it can capture both the link impairments as well as the node impairments along the lightpath. Besides the high NE CP complexity, which we can do very little about it, the only disadvantage of this scheme is its relative long setup time.

2.7. SUMMARY

In this chapter, necessary background technical theories have been introduced and critical literature reviews have been performed as a solid foundation for developing our own work. Section 2.2 introduced three QoS models – IntServ, DiffServ and MPLS; and surveyed the interoperations among the three models. Section 2.3 surveyed three analytical frameworks for statistical multiplexing technique. These two sections provide a good foundation for developing our work in Chapter 3. Section 2.5 covered the background knowledge of OPS node architecture and contention resolution techniques. A survey of optical buffering technique has been done to extend our work in Chapter 3 from electronic packet switched network to OPS network. Qualitative as well as quantitative comparisons have been done and the advantage of using MPLS together with QoS control scheme in the backbone network was emphasized. Section 2.4 covered all the necessary information we need to acquire for producing Chapter 4, where different optical network architectures were compared, commonly used optical components were introduced, major optical impairments that need attention were investigated, and a survey of the BER performance due to optical impairments was performed. Finally, section 2.6 introduces GMPLS and the merits as well as issues of the GMPLS based IP over WDM system, with the presence of optical layer impairments.

Table 2–4: Comparison of the impairment-aware lightpath setup Schemes

Scheme	Traffic	RWA	Resource and lightpath QoS	TED Scope	Other Issues	
					Pros	Cons
A	Static traffic	<ul style="list-style-type: none"> • Offline routing • Highly intensive computation • Centralized 	<ul style="list-style-type: none"> • Optimal lightpath QoS • Optimal TE and resource usage 	<ul style="list-style-type: none"> • Global TE and physical layer info • High TED vulnerability • Poorly scalable 	<ul style="list-style-type: none"> • Highly adopted model • Lowest NE complexity 	<ul style="list-style-type: none"> • Interoperability problem
B	Dynamic traffic	<ul style="list-style-type: none"> • Online routing • Highly intensive computation • Centralized 	<ul style="list-style-type: none"> • Optimal lightpath QoS • Optimal TE and resource usage 	<ul style="list-style-type: none"> • Global TE and physical layer info • High TED vulnerability • Low scalable 	<ul style="list-style-type: none"> • Interdomain scenarios 	<ul style="list-style-type: none"> • Early deployment state • NE CP complexity
C	Dynamic traffic	<ul style="list-style-type: none"> • Online routing • Intensive computation • Source-based routing 	<ul style="list-style-type: none"> • Optimal lightpath QoS • Optimal TE and resource usage 	<ul style="list-style-type: none"> • Global TE and physical layer info • TED stability problems 	<ul style="list-style-type: none"> • Impairments link-oriented 	<ul style="list-style-type: none"> • Scalability • Stability • High NE CP complexity
D	Dynamic traffic	<ul style="list-style-type: none"> • Online routing • Moderate computation • Source-based routing 	<ul style="list-style-type: none"> • Optimal TE and resource usage • Physical link excluded from routing 	<ul style="list-style-type: none"> • Global TE info • Local physical layer info, no flooding • Scalable 	<ul style="list-style-type: none"> • Link and node-oriented impairments 	<ul style="list-style-type: none"> • High setup delay • High NE CP complexity

Chapter 3. Modeling and Performance

Analysis of IP/MPLS Backbone Network

3.1. INTRODUCTION

Multi-Protocol Label Switching (MPLS) is a technology that can address ever increasing network requirements. It solves problems to which no other solutions are known. Firstly, MPLS addresses the scalability issues associated with the IP over ATM overlay model. Secondly, MPLS significantly reduces the complexity of network operation. Furthermore, MPLS facilitates the delivery of new routing capabilities that enhance conventional IP routing techniques, which is called traffic engineering (TE). Last but not the least, MPLS offers a standards-based solution that promotes multi-vendor interoperability. [5] [31]

As described in Chapter 1, Integrated Service (IntServ), Differentiated Service (DiffServ) and MPLS would interoperate in today's Internet environment in a heterogeneous manner, Figure 1-1, where senders from the IntServ and DiffServ networks need to communicate with receivers in the other IntServ and DiffServ networks respectively through the MPLS transit network. In this chapter, we are interested in both the qualitative service mapping among the three technologies, and the mathematic modeling as well as the quantitative performance analysis of such heterogeneous network. The desired QoS parameters to investigate consist of throughput, loss probability and delay.

After illustrating the beauty of MPLS label in section 3.2, we propose the traffic classification and service mapping among the IntServ, DiffServ and MPLS to enable the

interoperation of these technologies in section 3.3 and 3.4. Section 3.5 proposes a four-queue electronic system to be located at each output port of the MPLS router. Section 3.6 illustrates the mathematic modeling of the IntServ sources, DiffServ sources and the queueing system. Furthermore, the performance analyses of the four-queue system in terms of throughput, loss and delay are elaborated in section 3.7. Section 3.8 obtains the numerical results that demonstrate the desired QoS can be achieved by using the proposed traffic classification, service mapping, and queueing model. Towards the end of this chapter, section 3.9 elaborates a realizable optical buffer system, functionally equivalent to the electronic queueing system proposed in Section 3.5, which can be used at the output ports of network nodes in optical packet switching (OPS) networks.

3.2. MPLS LABEL

The beauty of MPLS technology is its use of labels. As described in Chapter 2, the analysis of the packet header in the network layer is performed just once at the edge of MPLS domain, and causes the packet to be assigned a stack of labels, where the labels are usually much smaller than the packet headers themselves [31][170]. At each subsequent hop, Label Switched Router (LSR) examines the label at the top of the label stack, and makes the decision for the next hop based solely on that label. It can then pop this label off the stack if it so desires, and push on zero or more labels onto the stack, before sending it on its merry way. MPLS LSR uses this simple label-swapping mechanism to replace the standard destination-based hop-by-hop forwarding paradigm, which can save significant processing time by avoiding network layer analysis at each hop and soaring processing demands at each network node. In this way, LSRs can quickly forward packets and enable easy scaling to terabit rates. MPLS connection is the Label Switched Path (LSP) that is set up by Label Distribution Protocol (LDP) [37]. By using LSPs in a manner similar to a connection-oriented network, MPLS can provide many of the same advantages of a connection-oriented network while still retaining the underlying efficiency and operation of a datagram network [171]-[173].

MPLS LSR is employed in the network to combine Layer 3 routing with Layer 2 high-speed switching efficiently and transfer IP datagram fast by cut-through transmission via LSPs in Layer 2, where the control plane and data-forwarding plan are separated. In the data plane, MPLS presents a simple way to forward information through networks by installing and then examining short, fixed-length labels in packets.

Figure 3–1 illustrates the fields contained in the 32-bit MPLS header [170]. The 20-bit long “Label” carries the actual value of the MPLS label. We shall call it “Path Label” since it is very much related to the path that the packet should follow in the MPLS domain. The 3-bit long “EXP” (experimental use bits) field is reserved for experimental use, which can affect the queueing and discard algorithm applied to the packet as it is transmitted through the network. We shall call it “QoS Label” in our work since it denotes different QoS levels. Note that packets with same “Path Label” may have different “QoS Label”. The 1-bit long “S” (Bottom of stack bit), supporting a hierarchical label stack, is set to one for the last entry in the label stack, and zero for all other label stack entries. Finally, 8-bit long “TTL” (time-to-live) field provides conventional IP TTL functionality.

Figure 3–1: MPLS Header Fields

3.3. PROPOSED TRAFFIC CLASSIFICATION

Today’s Internet only provides best effort service, where traffic is processed as quickly as possible with no guarantee as to timeliness or actual delivery. With the rapid transformation of the Internet into a commercial infrastructure, demands for service quality have rapidly developed, which needs several service classes [1] [174]–[175].

For this reason, labeled packets have to be divided into separate traffic classes to support service differentiation. As proposed in Figure 3–2, the incoming IP packets are classified into five categories. The **Premium** service category will provide low - delay and low-jitter services to applications such as Internet telephony and videoconferencing. Another service category will provide predictable Internet services for companies that do business on the Web. This service category may contain a single service category, or as in our case, it contains multiple categories - **Gold**, **Silver**, and **Bronze**, with decreasing quality. Finally, **Best-Effort** service will remain for those customers who only need connectivity. The recommendation of particular applications corresponding to the proposed priority categories can be found in Table 3–1.

Figure 3–2: Traffic Classification

3.4. PROPOSED MAPPING FOR INTSERV, DIFFSERV & MPLS

In a typical heterogeneous network, Figure 1–1, at the edge of the MPLS backbone network, the ingress LSR map the incoming QoS requests from the neighboring DiffServ and IntServ networks to the corresponding service categories within the MPLS network. In this model, each incoming flow is assigned to one of the available categories for the duration of the flow and traverses the MPLS cloud in this category. The service mappings between the incoming service requirements of the IntServ traffic, DiffServ traffic and MPLS service categories are given in Table 3–2, which follow most naturally from the service definitions. Guaranteed Service (GS) and Expedited Forwarding (EF) map to

Premium and Gold categories; Controlled-Load Service (CLS) and Assured Forwarding (AF) map to Silver and Bronze categories [176].

Table 3–1: Service Categories Mapping of Typical Applications

Application	Priority Mapping	Categories
voice, videoconferencing	Low loss probability Low delay and jitter	Premium
real time applications, virtual lease line in VPN , video streaming	Low loss probability Low delay	Gold
mission-critical data, VPN outsourcing, similar to ATM VBR	Low loss probability Controlled delay and jitter	Silver
e-commerce similar to ATM VBR	Low loss probability	Bronze
non-critical data no minimum information rate guaranteed	Best-effort delivery Unmanaged performance	BE

Table 3–2: Service Mapping For IntServ, DiffServ and MPLS Service Categories

Incoming Service Requirements		Service
IntServ	DiffServ	Categories
Guaranteed Service (GS)	Expedited Forwarding (EF)	Premium/Gold
Controlled-Load Service (CLS)	Assured Forwarding (AF)	Bronze/Silver
Best Effort (BE)	Default Forwarding (DF)	Best Effort

After analyzing variety of mappings between IntServ and DiffServ, MPLS EXP field and DiffServ, the possible mapping for the service type/class of the IntServ flows/ DiffServ packets and “QoS Label” (EXP field in MPLS header) in the MPLS header of packets is proposed in Table 3–3 [28][32][170]. This QoS mapping provides the foundation for the performance analysis in section 3.7, which is used when modeling the input traffic to our modified MPLS LSR [176]. In this case, the IntServ GS traffic is best mapped to the Premium category, the DiffServ High Priority traffic is mapped to Gold category, the DiffServ Low Priority traffic is mapped to Silver category and lastly the IntServ CLS traffic is mapped to Bronze category ,[26]-[27] [29]-[30]. Taking into account that the 8

Behavior Aggregates (BAs) defined within the MPLS network offer less service granularity than the DiffServ classes, the network administrator can choose to group together traffic flows requiring similar service into a single MPLS service category, and configure the mappings of the Differentiate Service Code Point (DSCP) values [177] that DiffServ QoS uses, as well as the incoming IntServ service request, to the “QoS Label” for output port scheduling and congestion avoidance mechanisms.

Table 3–3: Proposed QoS Mappings

IntServ/DiffServ Service Type (Class)	MPLS QoS Label (EXP)	MPLS Service Category
IntServ GS	111	Premium
DiffServ EF	110	Gold
DiffServ AF1X	101	
DiffServ AF2X	011	Silver
DiffServ AF3X	010	
DiffServ AF4X	001	
IntServ CLS	100	Bronze
Best Effort	000	Best Effort

3.5. MODELING OF NODE ARCHITECTURE

In a heterogeneous network, as shown in Figure 1–1, the MPLS Label Switch Routers (LSRs) may be interconnected by several different input and output links. Each link may carry traffic from both IntServ and DiffServ traffic with different categories.

Graphically, the LSR can be modeled as shown in Figure 3–3. Each LSR has several input ports and output ports. Each input ports may accept many IP packets with different MPLS “Path Labels”. Even if some packets from the same input ports might follow the same path through the MPLS domain thus having the same “Path Labels”, they might belong to different MPLS service categories hence having different “QoS Labels”.

Before being processed at the router, these packets, previously statistically multiplexed when transmitted over the link will be demultiplexed by the demultiplexer (DEMUX). IP packets with different tuple (input port, “Path Label”) are then forwarded to corresponding output ports according to data forwarding table. At the output port, the packets are put in different queues according to their “QoS Labels” and multiplexed again before being transmitted to the output link. The multiplexer at the output port consists of four output queues with multiple thresholds operating according to non-preemptive priority discipline.

Figure 3-3: Output Port Modeling of the LSR

In Table 3-3, a service mapping among IntServ service types, DiffServ service classes, MPLS “QoS Labels” (EXP field) and the MPLS service categories was proposed. Figure 3-4 illustrates the detailed modeling of four-queue system located at the every output port of the modified LSR shown in Figure 3-3. The Premium queue is of highest priority and accepts traffic from IntServ Guaranteed Service (GS). When the Premium queue is empty, the Gold queue will get the service. Multiple loss priority [69] [178] is implemented using threshold control mechanism over the partitioned buffer of the two queues taking input from DiffServ traffic. The Gold queue having $L1$ levels of thresholds for packet loss control, denoted as $B_1, B_2, \dots, B_{L1}$, is fed with DiffServ High Priority traffic, where subscript 1 denotes the highest priority against the packet loss and subscript $L1$ denotes the lowest

priority against the packet loss at the Gold queue. Only when both of the Premium and Gold queues are empty, the Silver queue can be served. The Silver queue takes input from DiffServ Low Priority sources and has $L2 - L1$ levels of thresholds, denoted as $B_{L1+1}, B_{L1+2}, \dots, B_{L2}$, for packet loss control over the partitioned buffer, where subscript $L1+1$ denotes the highest priority against the packet loss and subscript $L2$ denotes the lowest priority against the packet loss at the Silver queue. The Bronze queue is of the lowest priority, which will only be served when all the other queues have no input traffic.

Multiple loss priority scheme at each of the DiffServ queue is implemented as follows: Here we take the Gold queue as an example. When the buffer occupancy of the Gold queue is x , where $B_l < x < B_{l-1} (l \in [2, L1])$, the only packets admitted to the buffer belong to priority levels $(1, 2, \dots, (l-1))$. For example, when the buffer occupancy is x , where $0 < x < B_{L1}$, packets of all priority levels are admitted. In another scenario, the buffer occupancy is x , where $B_2 < x < B_1$, only packets of priority level 1 are admitted because the buffer is almost full at that time. Moreover, at the boundary $B_l (l \in [1, L])$ packets of priority level l are admitted only if their presence does not cause the buffer occupancy to exceed B_l . That is, at the upper boundary of each buffer occupancy level, packets of the lowest priority, among those admitted at the level, may be lost and the loss rate is determined by the residual capacity of the channel.

Figure 3-4: Proposed Output-Port Four-Queue System

3.6. ASSUMPTIONS OF MATHEMATIC MODELING

The fluid flow technique [64]-[65] is used in the following performance analysis for DiffServ traffic with assumptions that (1) the population of DiffServ sources is large ($N \gg 1$), (2) the output transmission link is divided into equal time slots and each slot is equivalent to the maximum transmission time of packet such that the transmission time of IP packets is assumed to be uniformly distributed. This assumption is reasonable, because that when the population of DiffServ sources is large, the bit rate allocated to each DiffServ source comparing to the huge capacity of the broadband transmission link is negligible. Therefore, the variation of the transmission time of packet is also negligible.

3.7. MATHEMATIC MODELING AND ANALYSIS

3.7.1. Performance of Premium Queue

The Premium queue accepts traffic from IntServ Guaranteed Service (GS) customers, which is on per-flow basis. To model the Premium queue, we first take a look at the well-known M/M/1 queueing system. M/M/1 queueing system assumes a Poisson arrival process [179]-[180]. This assumption is a very good approximation for arrival process in real systems, such as road traffic system, telephony system, and etc., that meets the following criterion [181]: (a) The number of customers in the system is very large; (b) The impact of a single customer on the performance of the system is very small; (c) All customers are independent, i.e. their decisions to use the system are independent of other users [182]. Since we are interested in the backbone of a large-scale network, the customer bases of all the service categories are large. As a result, a single customer consumes a very small percentage of the system resources. Also, since the customers are well spread over the edge networks, independency of the customer profiles can be justified. Hence we can

assume that the arrival of the GS packets is of Poisson distribution with an average rate of λ_{GS} , and the service time is of exponential distribution with an average value of $1/C$.

Furthermore, since the premium queue has limited queueing capacity to take into account the fact that the percentage of premium traffic is usually small, we make use of the M/M/1 system with limited queueing capacity constraint. Hence, as shown in Figure 3-5, the $M/M/1/B_{GS}$ system, with limited queueing capacity B_{GS} , is used to model the Premium queue, as long as $\lambda_{GS} < C$, which is of the normal case [182]-[183].

Figure 3-5: Modeling of IntServ GS Traffic

The throughput of the Premium queue is

$$T_{premium} = \lambda_{GS} (1 - P_{b_premium}), \text{ if } \lambda_{GS} < C \quad (3-1)$$

where $P_{b_premium} = \frac{(1 - (\lambda_{GS}/C))(\lambda_{GS}/C)^{B_{GS}}}{1 - (\lambda_{GS}/C)^{B_{GS}+1}}$ is the blocking probability.

Hence the loss probability of the Premium queue is,

$$loss_{premium} = 1 - \frac{T_{premium}}{\lambda_{GS}} = P_{b_premium} \quad (3-2)$$

and the average delay of the Premium queue is,

$$D_{premium} = \frac{1}{(C - \lambda_{GS})} - \frac{B_{GS}(\lambda_{GS}/C)^{B_{GS}+1}}{\lambda_{GS} - C(\lambda_{GS}/C)^{B_{GS}+1}} \quad (3-3)$$

3.7.2. Performance Analysis of Gold Queue

3.7.2.1. DiffServ Source Modeling

Assuming that the DiffServ sources can generate an arbitrary number $L2$ of the priority classes, packets of priorities $l \in [1, L1]$ belong to Gold queue, while packets of priority $l \in [L1+1, L2]$ belong to the Silver queue. Each source has two states with generator G and also the traffic streams are coupled, i.e., mutually correlated [69] [178]. Loss priority is implemented by partitioning the buffer into as many levels of occupancy as there are priority classes. Packets of a particular priority class are either admitted or discarded depending upon the prevailing level of buffer occupancy. Each source when in state i ($i = 1, 2$) generates traffic of priority class l ($l = 1, 2, \dots, L2$) at rate $\lambda_i^{(l)}$. Hence each of the N sources is specified by $(\mathbf{G}; \lambda^{(1)}, \dots, \lambda^{(L2)})$ where $\lambda^{(l)} = [\lambda_1^{(l)}, \lambda_2^{(l)}]$. $\lambda_1^{(l)}$ denotes the arrival rate of the OFF state of class l for each source. $\lambda_2^{(l)}$ denotes the arrival rate of the ON state of class l for each source. Hence, the rate of generation of class l traffic $\gamma_n^{(l)}$ ($n \in [0, N]$) is defined to be $n\lambda_2^{(l)} + (N-n)\lambda_1^{(l)}$. Assume that the arrival rates of the OFF state of any of the priority classes are zero, therefore, $\lambda_1^{(l)} = 0$. In this case, the source is modeled as a simple ON-OFF process, with $\gamma_n^{(l)} = n\lambda_2^{(l)}$. For the convenience of representation, λ_l is used to represent $\lambda_2^{(l)}$.

As shown in Figure 3-6 (a), a single DiffServ is modeled as an ON-OFF process [184] [185], in which the transition rate from ON state to OFF state is α and the rate of transition from OFF state to ON state is β . Then the probability that the process is in the ON state is $\alpha/(\alpha + \beta)$. When N of such ON-OFF packet streams are multiplexed, the resultant stream can be represented by an $(N+1)$ -state Markov modulated process as shown in Figure 3-6 (b), where the state n represents that n ($n \in [0, N]$) packet streams are in the ON state.

Figure 3-6: Modeling of DiffServ Source

3.7.2.2. Gold Queue Modeling

The Gold queue can only get the service when the Premium queue is empty and the service rate for the Gold queue is reduced from C to $C' = C - T_{GS}$. Modeling of the Gold queue follows the fluid flow model introduced in [64]-[65] (please refer to section 2.3.3 and Appendix A for the details). Given the average OFF period of the source is denoted by $1/\alpha$, if at time t the number of ON sources equals n , two elementary events can take place during the next interval Δt , i.e., a new source can start or a source can turn off. Since the ON and OFF periods are exponentially distributed, the respective probabilities are $(N-n)\alpha\Delta t$ and $n\beta\Delta t$. Compound events have probabilities $0(\Delta t^2)$. The probability of no change is $1 - \{(N-n)\alpha + n\beta\}\Delta t + 0(\Delta t^2)$.

Let $P_n(t, x), n \in [0, N], t \geq 0, x \geq 0$, be the probability that at time t , n sources are ON and the buffer content does not exceed x . Now,

$$P_n^l(t + \Delta t, x) = (N - (n - 1))\alpha\Delta t P_{n-1}^l(t + \Delta t, x) + (n + 1)\beta\Delta t P_{n+1}^l(t + \Delta t, x) + [1 - ((N - n)\alpha + n\beta)\Delta t]P_n^l(t, x - n\lambda^l - C')\Delta t + 0(\Delta t^2) \quad (3-4)$$

where $\lambda^l = \sum_{k=1}^l \lambda_k$ (λ_k is the arrival rate of k th class DiffServ Traffic at ON state). Hence

λ^l is the sum of the arrival of priorities higher or equal to l .

Since only time-independent, equilibrium probabilities are interested, therefore, we take partial derivation of equation above to t , pass to the limit $\Delta t \rightarrow 0$ and set $\frac{\partial P_n^l}{\partial t} = 0$ and obtain, for $n \in [0, N]$,

$$(n\lambda^l - C') \frac{dF_n^l}{dx} = (N - (n-1))\alpha F_{n-1}^l + (n+1)\beta F_{n+1}^l - ((N-n)\alpha + n\beta)F_n^l \quad (3-5)$$

In the equation above, $F_n^l(x)$ denotes the equilibrium probability that n sources are in the ON state and buffer content does not exceed x of class l . Note that $F_n^l = 0$, if n is not in the stated interval. Equation above can be represented in the matrix notation as shown below,

$$\mathbf{D}^l \frac{d}{dx} \mathbf{F}^l(\mathbf{x}) = \mathbf{M} \mathbf{F}^l(\mathbf{x}), x \geq 0 \quad (3-6)$$

where $\mathbf{D}^l = \text{diag}\{-C', \lambda^l - C', 2\lambda^l - C', \dots, N\lambda^l - C'\}$ and $\mathbf{M}$ is the $(N+1) \times (N+1)$ tri-diagonal matrix shown below,

$$\mathbf{M} = \begin{bmatrix} -N\alpha & \beta & 0 & 0 & \dots & 0 & 0 \\ N\alpha & -\{(N-1)\alpha + \beta\} & 2\beta & 0 & \dots & 0 & 0 \\ 0 & (N-1)\alpha & -\{(N-2)\alpha + 2\beta\} & 3\beta & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 2\alpha - \{\alpha + (N-1)\beta\} & N\beta \\ 0 & 0 & 0 & 0 & \dots & \alpha & -N\beta \end{bmatrix}$$

Assuming $n\lambda^l - C'$ is not equal to zero for any n , $0 \leq n \leq N$, the general solution of equation above is:

$$\mathbf{F}^l(\mathbf{x}) = \sum_{n=0}^N a_n^l \mathbf{V}_n^l e^{z_n^l x} \quad (3-7)$$

where $\mathbf{F}^l(\mathbf{x}) = [F_0^l(x), F_1^l(x), \dots, F_N^l(x)]$. The elements in $\mathbf{z}^l(\mathbf{x}) = [z_0^l, z_1^l, \dots, z_N^l]$ are the eigenvalues of matrix $(\mathbf{D}^l)^{-1}\mathbf{M}$ and $\mathbf{V}_n^l$ is the eigenvector of matrix $(\mathbf{D}^l)^{-1}\mathbf{M}$. In the solution, the coefficients $\{a_n^l\}$ can be obtained from boundary conditions by defining $S_D^l = \{n \mid n\lambda^l < C'\}$ and $S_U^l = \{n \mid n\lambda^l > C'\}$.

The boundary condition of the Gold queue can be obtained,

$$\begin{aligned} F_n^1(B_1) &= P(n) & \text{If } n \in S_D^1 \\ F_n^l(B_l) &= F_n^{l-1}(B_l) & \text{If } n \in S_U^{l-1} \cup S_D^l, l \in [2, L1] \\ F_n^{L1}(0) &= 0 & \text{If } n \in S_U^{L1} \end{aligned} \quad (3-8)$$

where $P(n)$ is the probability that n DiffServ sources are in the ON state and P_{ON} is the probability that one DiffServ source is in the ON state.

$$P(n) = \binom{N}{n} P_{ON}^n (1 - P_{ON})^{N-n}, \quad n \in [0, N], \quad P_{ON} = \frac{\alpha}{\alpha + \beta} \quad (3-9)$$

3.7.2.3. Performance Analysis

Arrival rate of the l th class packet is $A_{Gold}^l = \sum_{n=0}^N \lambda_l n P(n)$, where $l \in [1, L1]$. As mentioned

before, Gold queue can only get service when the Premium queue has no input traffic.

Hence the throughputs of l th class are as follows,

For class $l=1$ traffic in this queue,

$$\begin{aligned} T_{Gold}^1 &= \text{Arrival of Class 1 Traffic} - \text{Drop Rate of Class 1 Traffic} \\ &= \sum_{n=0}^N \lambda_1 n F_n^1(B_1) + C' \left[1 - \sum_{n=0}^N F_n^1(B_1) \right] \end{aligned} \quad (3-10)$$

For the priority classes other than class 1, that is $l \in [2, L1]$,

$$T_{Gold}^l = \text{Arrival of Class } l \text{ traffic} - \text{Drop Rate of Class } l \text{ traffic}$$

$$= \sum_{n=0}^N \lambda_l n F_n^{l-1}(B_l) - \sum_{n=S_U^{l-1} \cap S_D^l} \left[(F_n^{l-1}(B_{l-1}) - F_n^l(B_l)) \times (n\lambda^l - C') \right] \quad (3-11)$$

Hence the loss probability,

$$loss_{Gold}^l = 1 - \frac{T_{Gold}^l}{A_{Gold}^l}, \quad l \in [1, L1] \quad (3-12)$$

For the Gold queue, delay is also an important QoS parameter because the traffic inputs to the Gold queue are delay sensitive traffic. Next, we will derive the equation for the delay of the different priority classes in the Gold queue. Let $Q^l(t)$ be the probability that the delay is less than t for the l th class packet.

$$Q_k^l(t) = \frac{1}{T_{Gold}^l} \sum_{n=0}^N n \lambda_l F_n^k(C't), \quad \frac{B_{k+1}}{C'} < t < \frac{B_k}{C'}, \quad k \in [l, L1-1] \quad (3-13)$$

$$Q_{L1}^l(t) = \frac{1}{T_{Gold}^l} \sum_{n=0}^N n \lambda_l F_n^{L1}(C't), \quad 0 \leq t < \frac{B_{L1}}{C'} \quad (3-14)$$

$$\Pr\left(\text{delay} = \frac{B_1}{C'}\right) = \frac{1}{T_{Gold}^1} \left[1 - \sum_{n=0}^N F_n^1(B_1) \right] C' \quad (3-15)$$

$$\Pr\left(\text{delay} = \frac{B_k}{C'}\right) = \frac{1}{T_{Gold}^l} \sum_{n=0}^N n \lambda_l \left[F_n^{k-1}(B_k) - F_n^k(B_k) \right], \quad k \in [l, L1] \quad (3-16)$$

where B_k is the k th threshold value and C' is the effective service rate for Gold queue.

Therefore, the average delays for different priority classes are calculated below from the delay distribution function as shown by Equations (3-13) – Equation (3-16),

For $l = 1$, the average delay is given by,

$$D^1 = \int_0^{\left(\frac{B_{L1}}{C'}\right)^-} t d(Q_{L1}^1(t)) + \sum_{k=1}^{L1-1} \int_{\left(\frac{B_{k+1}}{C'}\right)^-}^{\left(\frac{B_k}{C'}\right)^-} t d(Q_k^1(t)) + \sum_{k=2}^{L1} \frac{B_k}{C'} \Pr\left(\text{delay} = \frac{B_k}{C'}\right) + \frac{B_1}{C'} \Pr\left(\text{delay} = \frac{B_1}{C'}\right) \quad (3-17)$$

where $\int_0^{\left(\frac{B_{L1}}{C'}\right)^-} td(Q_{L1}^1(t))$ is the average delay of priority level 1, calculated when the buffer occupancy level is $0 \leq x < B_{L1}$. $\sum_{k=1}^{L1-1} \int_{\left(\frac{B_{k+1}}{C'}\right)^-}^{\left(\frac{B_k}{C'}\right)^-} td(Q_k^1(t))$ is the sum of the average delay of priority level 1, calculated when the buffer occupancy level is between any of the two neighboring thresholds i.e. $B_{k+1} < x < B_k$ ($k \in [1, L1-1]$). $\sum_{k=2}^{L1} \frac{B_k}{C'} \Pr\left(delay = \frac{B_k}{C'}\right)$ is the sum of the average delay of priority level 1, calculated when the buffer occupancy happens to be at the thresholds B_k ($k \in [1, L1]$) of the Gold queue. Finally, $\frac{B_1}{C'} \Pr\left(delay = \frac{B_1}{C'}\right)$ is the average delay calculated when the buffer is fully occupied.

The average delay for the l th traffic with the loss priority ($l \in [2, L1]$) can be obtained from the delay distribution function.

$$D^l = \int_0^{\left(\frac{B_{L1}}{C'}\right)^-} td(Q_{L1}^l(t)) + \sum_{k=l}^{L1-1} \int_{\left(\frac{B_{k+1}}{C'}\right)^-}^{\left(\frac{B_k}{C'}\right)^-} td(Q_k^l(t)) + \sum_{k=l}^{L1} \frac{B_k}{C'} \Pr\left(delay = \frac{B_k}{C'}\right) \quad (3-18)$$

where $\int_0^{\left(\frac{B_{L1}}{C'}\right)^-} td(Q_{L1}^l(t))$ is the average delay of priority level l , calculated when the buffer occupancy level is $0 \leq x < B_{L1}$. $\sum_{k=l}^{L1-1} \int_{\left(\frac{B_{k+1}}{C'}\right)^-}^{\left(\frac{B_k}{C'}\right)^-} td(Q_k^l(t))$ is the sum of the average delay of priority level l , calculated when the buffer occupancy level is between any of the two neighboring thresholds, i.e. $B_{k+1} < x < B_k$ ($k \in [l, L1-1]$). Finally, $\sum_{k=l}^{L1} \frac{B_k}{C'} \Pr\left(delay = \frac{B_k}{C'}\right)$ is the sum of the average delay of priority level l , calculated when the buffer occupancy happens to be at the thresholds B_k ($k \in [l, L1]$) of the Gold queue.

3.7.3. Performance Analysis of Silver Queue

The Silver queue takes input from DiffServ low priority traffic with input priority level $l \in [L1+1, L2]$. It implements the multiple priorities control scheme by setting thresholds in the buffer. The detailed modeling of the queue is similar to the modeling of the Gold queue with buffer threshold control $B_{L1+1}, B_{L1+2}, \dots, B_{L2}$. The Silver queue can only be served when both of the Premium and Gold queues are empty. In this case, we define $C_j'' = C_j' - j\lambda^{L1}$, recall that $\lambda^{L1} = \sum_{j=1}^{L1} \lambda_j$, and let $F_{jn}^l(x)$ denote the probability that content of Silver queue does not exceed x when there are j sources at Gold queue and n sources at the Silver queue.

$$(n\lambda^l - C_j'') \frac{dF_{jn}^l}{dx} = (N - (n-1))\alpha F_{j(n-1)}^l + (n+1)\beta F_{j(n+1)}^l - ((N-n)\alpha + n\beta)F_{jn}^l \quad (3-19)$$

where $l \in [L1+1, L2]$, $n, j \in [0, N]$, $x \in [0, B_{L1+1}]$

The solution to the equation above is,

$$\mathbf{F}_{jn}^l(\mathbf{x}) = \sum_{n=0}^N a_{jn}^l \mathbf{V}_{jn}^l e^{z_{jn}^l x}, \quad x \in [0, B_{L1+1}] \quad (3-20)$$

where, the coefficients $\{a_{jn}^l\}$ can be obtained from boundary conditions by defining

$$S_{jD}^l = \{n \mid n\lambda^l < C_j''\} \quad \text{and} \quad S_{jU}^l = \{n \mid n\lambda^l > C_j''\}.$$

The boundary conditions of the Silver queue are,

$$\begin{aligned} F_{jn}^{L1+1}(B_{L1+1}) &= P(n) & \text{If } n \in S_{jD}^{L1+1} \\ F_{jn}^l(B_l) &= F_{jn}^{l-1}(B_l) & \text{If } n \in S_{jU}^{l-1} \cup S_{jD}^l, l \in [L1+2, L2] \\ F_n^{L2}(0) &= 0 & \text{If } n \in S_{jU}^{L2} \end{aligned} \quad (3-21)$$

Now we are at the stage to calculate the throughput of the Silver queue traffic.

Firstly, the class $L1 + 1$ throughput is,

$$\begin{aligned} T_{jSilver}^{L1+1} &= \text{Arrival of Class } L1 + 1 \text{ Traffic} - \text{Drop Rate of Class } L1 + 1 \text{ Traffic} \\ &= \sum_{n=0}^N \lambda_{L1+1} n F_{jn}^{L1+1}(B_{L1+1}) + C_j'' \left[1 - \sum_{n=0}^N F_{jn}^{L1+1}(B_{L1+1}) \right] \end{aligned} \quad (3-22)$$

The throughputs of the classes other than class $L1 + 1$ ($l \in [L1 + 2, L2]$) are,

$$\begin{aligned} T_{jSilver}^l &= \text{Arrival of Class } l \text{ traffic} - \text{Drop Rate of Class } l \text{ traffic} \\ &= \sum_{n=0}^N \lambda_l n F_{jn}^{l-1}(B_l) - \sum_{n=S_{jU}^{l-1} \cap S_{jD}^l} \left[(F_{jn}^{l-1}(B_{l-1}) - F_{jn}^l(B_l)) \times (n\lambda^l - C_j'') \right] \end{aligned} \quad (3-23)$$

Hence, the average throughput of the Silver queue is,

$$T_{Silver}^l = \sum_{j=0}^{j < C'' / \lambda^{L1}} T_{jSilver}^l F_j^{L1}(0), \quad l \in [L1 + 1, L2] \quad (3-24)$$

Since $A_{Silver}^l = \sum_{n=0}^N \lambda_l n P(n)$ is the arrival rate of the l th ($l \in [L1 + 1, L2]$) class packets in the

Silver queue, the loss probability for the Silver queue can be obtained as,

$$loss_{Silver}^l = 1 - \frac{T_{Silver}^l}{A_{Silver}^l}, \quad l \in [L1 + 1, L2] \quad (3-25)$$

3.7.4. Performance Analysis of Bronze Queue

The Bronze queue contains the traffic from the IntServ CLS (Integrated Service Controlled-Load Service). Similar to the Premium queue, whose traffic is IntServ GS traffic, the traffic through the Bronze queue is also on per-flow basis. This queue can only get the service when none of the previous three queues has input traffic. Therefore, the

service rate of the Bronze queue reduces further to $C''' = C' - \sum_{l=1}^{L1} T_{Gold}^l - \sum_{l=L1+1}^{L2} T_{Silver}^l$. Assume

that the arrival rate λ_{CLS} of the CLS traffic is of Poisson distribution and the service time $1/C'''$ is of exponential distribution. As long as $\lambda_{CLS} < C'''$, which is of the normal case, and the buffer is big enough, this queue can be modeled as $M/M/1/B_{CLS}$ system [183] as shown in Figure 3–7.

Figure 3–7: Modeling of IntServ CLS Traffic

Hence the throughput of the Bronze queue is,

$$T_{Bronze} = \lambda_{CLS} (1 - P_{b_bronze}), \text{ if } \lambda_{CLS} < C''' \quad (3-26)$$

where $P_{b_bronze} = \frac{(1 - (\lambda_{CLS} / C'''))(\lambda_{CLS} / C''')^{B_{CLS}}}{1 - (\lambda_{CLS} / C''')^{B_{CLS}+1}}$ is the blocking probability of bronze

queue. Therefore, the loss probability of the Bronze queue is,

$$loss_{Bronze} = 1 - \frac{T_{Bronze}}{\lambda_{CLS}} = P_{b_bronze} \quad (3-27)$$

3.7.5. End-to-End Performance Analysis

Assuming that the Label Switched Path (LSP) starts from node 1 until node K , hence knowing the loss probability and delay of each node for every service category, we can obtain the end-to-end loss probability as,

$$Loss_c^l = 1 - \left(\prod_{k=1}^K (1 - Loss_c^l(k)) \right) \quad (3-28)$$

where $Loss_c^l(k)$ denotes the loss probability at k th nodes along the path for service category c and class l . Meanwhile, the end-to-end delay is obtained,

$$Delay_c^l = \sum_{k=1}^K D_c^l(k) \quad (3-29)$$

where $D_c^l(k)$ is the delay at k th nodes along the path for service category c and class l .

3.8. RESULTS AND DISCUSSIONS

The following numerical results focus on the effect of priorities on the steady-state performance including the different IntServ traffic types (GS and CLS) and different DiffServ classes (EF and AF).

For the illustrative purpose only, as shown in Figure 3–8, the queues are classified according to service mapping proposed in Table 3–3. The Premium queue of length B_{GS} accepts traffic from Premium Category, i.e. IntServ Guaranteed Service (GS). The Gold queue handling Gold category traffic of length B_1 consists of 2 classes of packet loss priorities, named Class 1 and Class 2, where Class 1 traffic is the DiffServ Expedited Forwarding (EF) traffic, with higher priority than Class 2 traffic with input from DiffServ Assured Forwarding (AF1x). The threshold B_2 is used for the packet loss priority control in the Gold queue. When B_2 reaches, only Class 1 traffic is allowed to get in the Gold queue. When B_1 reaches, traffic from neither of the two Gold classes is allowed to enter the Gold queue. The Silver queue takes input from DiffServ low priority traffic and it consists of 3 classes of packet loss priorities, named Class 3, Class 4 and Class 5. The Class 3 traffic has higher priority than Class 4 traffic, which in turn has higher priority than Class 5 traffic. B_3 is the queue length and thresholds B_4 and B_5 are used for

priority control in this queue. A proposed scheme is that the DiffServ AF2x, AF3x, AF4x traffic are assigned to Class 3, 4 and 5 respectively. When B_5 reaches, only traffic from Class 3 and 4 are allowed to get in the Silver queue. When B_4 reaches, only Class 3 traffic is allowed to enter the Silver queue. Of course, when B_3 reaches, traffic from all the Silver category classes will be dropped. For the DiffServ sources, $\alpha = 0.4$ and $\beta = 1$, so the probability that the process is in the ON state is $\alpha/(\alpha + \beta) = 0.4/1.4$. Similar to the Premium queue, the Bronze queue of length B_{CLS} accepts per-flow traffic from IntServ Controlled-Load Service (CLS).

In this section, the term “Offered Load” is defined to be the total arrival traffic as normalized to the link capacity or the service rate, which does not have unit. The queue capacities/buffer sizes used for calculating the packet loss probability shown in the figures are measured as normalized to the link capacity. The buffer sizes of both the Gold queue and Silver queue are 36% of the output link capacity for Figure 3–9 ~ Figure 3–12. The buffer size of the Gold queue is 24% of the output link capacity and the buffer size of the Silver queue is 36% of the output link capacity for Figure 3–13 ~ Figure 3–20. The buffer size of the Gold queue is 0.32% of the output link capacity and the buffer size of the Silver queue is 1.6% of the output link capacity for Figure 3–21.

Figure 3–8: Queueing Model for Numerical Analysis

Since the behavior of two queues storing per-flow IntServ traffic are very easy to analyze, therefore, we focus on the two queues comprising the DiffServ traffic, i.e. Gold queue and Silver queue. Hence, we assume that both of the IntServ queues have no incoming traffic for the ease of analysis.

3.8.1. Effect of Gold Queue Load on the Loss Probabilities and Throughputs of the Traffic Classes

In this subsection, we want to show the effect traffic loads of Gold queue on the loss probabilities and throughputs of all the traffic classes. Here, the buffer sizes of both the Gold queue and Silver queue are 36% of the output link capacity. The lines with filled marker show the results *with* threshold control, whereas the rest of the lines show the results *without* threshold control.

Firstly, Figure 3–9 and Figure 3–10 illustrate the effect of Class 1 traffic on the loss probability and the throughput of Gold queue classes, i.e. Class 1 and Class 2. The effect of threshold control in the Gold queue is also shown. In this scenario, traffic load from Class 2 is fixed to be 40% of the total offered load and Silver queue has no input traffic. Threshold is set to be 40% (B_2) of the total Gold queue capacity. Results are obtained by varying the Class 1 traffic load from 10% to 40 % of the total offered load at a step of 10%.

Figure 3–9 shows that, the loss probabilities of both Class 1 and Class 2 increase with the increase of Class 1 traffic load and total traffic load regardless whether there is threshold control scheme. For example, *with* threshold control, as offered load increases from 1 to 1.6, when Class 1 load is 0.4, the loss probability of Class 1 increases from 10^{-6} to 10^{-2} and the loss probability of Class 2 increases from 4.5×10^{-2} to 0.55. Another example is that, when offered load is fixed to be 1, as Class 1 traffic load increases from 0.1 to 0.4, the

loss probability of Class 1 increases from almost 0 to 10^{-6} and the loss probability of Class 2 increases from 4×10^{-4} to 4.5×10^{-2} .

Furthermore, Figure 3–9 also tells that **with** threshold control, loss probability of Class 1 improves significantly at the expense of slight increase of Class 2 loss probability. For example, when the total offered load is around 1 and traffic load from Class 1 is 0.4, the loss probability of Class 1 drops significantly from 4.5×10^{-2} to 10^{-6} while the loss probability of Class 2 increases from 4.5×10^{-2} to 10^{-1} . Without (**w/o**) threshold control, Class 1 and Class 2 are of no difference in loss probabilities.

Figure 3–9(a): Class 1 loss probability with varying Class 1 traffic Load

Figure 3-9(b): Class 2 loss probability with varying Class 1 traffic Load

Figure 3-9: Effect of Class 1 Traffic Load on Loss Probability of Gold Queue, With and Without Threshold Control

Figure 3-10 shows that, *with* threshold control, the throughput of Class 1 increases with the increase of total traffic load and Class 1 traffic load. For example, when total offered load increases from 1 to 1.6, with 0.4 Class 1 load, the throughput of Class 1 increases from 0.4 to 0.64. On the other hand, if total offered load is fixed to be 1.6, as Class 1 traffic increases from 0.1 to 0.4, the throughput of Class 1 increases from 0.16 to 0.64. However, Class 2 behaves unlike Class 1. Firstly, when offered load is small, with different Class 1 load, the throughput of Class 2 has no change because its load is set to be 40% of total offered load. However, as offered load increases further, the throughput curve with higher Class 1 load drops earlier and more severely than those with lower Class 1 load. This is because when offered load increases further, the threshold is reached and Class 2 cannot be accepted any more.

Meanwhile, as the Class 1 load increases, there is less chance that Class 2 will get through the queue. For example, the highest throughputs of Class 2 happen at (Class 1 load, offered load) = throughput are: $(0.1, 2.4) = 0.6697$, $(0.2, 1.8) = 0.5268$, $(0.3, 1.4) = 0.4285$ and $(0.4, 1.2) = 0.3603$.

Furthermore, Figure 3–10 also elaborates that **with** threshold control, in heavily loaded network environment, the throughput of Class 1 keeps increasing, which improves a lot from the case without (**w/o**) threshold control. On the contrary, Class 2 performs poorer **with** threshold control, this is because the Gold queue is occupied by Class 1 traffic most of time offered load is heavy. When Class 1 load is 0.4, which is the same as the fixed Class 2 load, **with** threshold, the throughput of Class 1 approaches 1 while the throughput of Class 2 approaches 0. Without (**w/o**) threshold control, the throughputs of both Class 1 and Class 2 approach 0.5 because they are of no difference. When Class 1 is 0.1, the throughput of Class 1 approaches 0.2 while that of Class 2 approaches 0.8. These results also prove that our derived equations are correct.

Figure 3–10(a): Class 1 throughput with varying Class 1 traffic Load

Figure 3-10(b): Class 2 throughput with varying Class 1 traffic Load

Figure 3-10: Effect of Class 1 Traffic Load on Throughput of Gold Queue, With and Without Threshold Control

Secondly, Figure 3-11 and Figure 3-12 illustrate the effect of Gold queue traffic, i.e. Class 1 and 2 on the loss probabilities and the throughputs of Silver queue classes, i.e. Class 3, 4 and 5. The effect of threshold control in the Silver queue is also illustrated. In this scenario, the traffic load from each of the Silver queue classes is fixed to be 20% of the total offered load. Gold queue threshold is set to be 40% (B_2) of the total Gold queue capacity and Silver queue thresholds are set to be 30% (B_5) and 60% (B_4) of the total Silver queue capacity. Results are obtained by varying the Class 1 and 2 traffic load from 5% each to 20% each of the total offered load at a step of 5%.

Figure 3-11 shows that, the loss probabilities of all the Silver queue classes increase with the increase of Gold queue traffic load and total traffic load despite whether there is threshold control scheme. For example, *with* threshold control, as offered load increases from 0.8 to 1, when traffic loads from Class 1 and Class 2 are 0.1 each, the loss probability of Class 3 increases from 8.6×10^{-13} to 10^{-6} , the loss probability of Class 4 increases from

2.9×10^{-4} to 1.5×10^{-2} and the loss probability of Class 5 increases from 0.056 to 0.198. Another example is that, when offered load is fixed at 0.8, as traffic loads from Class 1 and Class 2 increase from 0.05 each to 0.2 each, the loss probability of Class 3 increases from almost 0 to 1.7×10^{-3} , the loss probability of Class 4 increases from 5.85×10^{-6} to 0.033 and the loss probability of Class 5 increases from 0.018 to 0.203.

Furthermore, Figure 3–11 proves that **with** threshold control, loss probability of Class 3 improves significantly and Class 4 improves apparently at the expense of slight increase of Class 5 loss probability. For example, when the total offered load is 0.8 and traffic loads from Class 1 and 2 are 0.1 each, the loss probability of Class 3 drop significantly from 4×10^{-3} to 10^{-12} , the loss probability of Class 4 drop apparently from 4×10^{-3} to 3×10^{-4} , while the loss probability of Class 5 increases slightly from 4×10^{-3} to 4×10^{-2} .

Without (**w/o**) threshold control, Class 3, 4 and 5 are of no difference in loss probabilities.

Figure 3–11(a): Class 3 loss probability with varying load from Class 1 & 2

Figure 3-11(b): Class 4 loss probability with varying load from Class 1 & 2

Figure 3-11(c): Class 5 loss probability with varying load from Class 1 & 2

Figure 3-11: Effect of Gold Queue Traffic Load on Loss Probability of Silver Queue, With and Without Threshold Control

Figure 3–12 shows that, the throughputs of all the Silver queue classes degrade consistently with the increases of load from Gold queue classes, regardless whether the threshold control scheme is used. Since Class 3 has higher priority than Class 4, which in turn has higher priority than Class 5, the of peak throughput of Class 3 (e.g., 0.3693 when w/ 1&2: 10%) are shown to be higher than those of Class 4 (0.2266 when w/ 1&2: 10%), which in turn should be higher than those of Class 5 (0.1603 when w/ 1&2: 10%).

It is also observed that the throughput curves are similar to Figure 3–10(b). When offered load is small, with different loads from Gold classes, the throughput of Silver classes are of no change because their load is set to be 0.2 of total offered load each. However, as offered load increases further, the throughput curve with higher Gold classes load drops earlier and more severely than those with lower Gold classes load. This is because when offered load increases further, the thresholds B_5 , B_4 and even B_3 reach subsequently and the Silver classes traffic are dropped accordingly. Meanwhile, as the loads of Gold classes increase, there is less chance that Silver queue get the service, thus Silver queue is more likely to build up.

Furthermore, Figure 3–12 also elaborates that **with** threshold control, Class 3 always has higher throughput than that **without** threshold (Figure 3–12(a)), Class 5 always has lower throughput than that without threshold (Figure 3–12(c)). However, it is interesting to notice the behavior of Class 4 is not that ‘consistent’. For example, as shown in Figure 3–12(b), when Gold classes have loads of 0.05 each, when total offered load increases beyond 0.6, the curve **with** threshold outperforms that **without** threshold. But when total offered load increases further till 2, the curve **with** threshold start to appear below that **without** threshold. This is while packets from Class 3 are protected by both B_4 and B_5 , Class 5 are discriminated by B_5 , Class 4 are protected by B_5 but discriminated by B_4 . The complexity of the queueing effect causes Class 4’s unique behavior.

Figure 3-12(a): Class 3 throughput with varying load from Class 1 & 2

Figure 3-12(b): Class 4 throughput with varying load from Class 1 & 2

Figure 3-12(c): Class 5 throughput with varying load from Class 1 & 2

Figure 3-12: Effect of Gold queue Traffic Load on Loss Probability of Silver Queue, With and Without Threshold Control

As a summary, this subsection illustrates that priority indeed has effect on the packet loss probabilities and throughputs of the traffic from all the five different classes. Under the same load, classes with higher priorities less likely experience packet loss than the classes with lower priority. More importantly, it also can be seen that the traffic loads offered by the classes with higher classes have significant effects on the performance of packet loss probabilities and throughputs for all the other lower priority traffic classes. Therefore, in order to achieve desired packet loss probabilities for different traffic classes, both the overall traffic load in the network and the traffic loads offered by the higher-priority classes need to be properly controlled.

3.8.2. Effect of Threshold Control on the Loss Probabilities and Throughputs of the Traffic Classes

In this subsection, we focus on the effect of threshold control implemented in the queueing system on the loss probabilities and throughputs of the traffic classes. Here, the offered load of Class 1 to Class 5 are fixed at 10%, 15%, 25%, 25% and 25% of the total offered load respectively. The buffer size of the Gold queue is 24% of the output link capacity and the buffer size of the Silver queue is 36% of the output link capacity. The dotted lines show the results *without* threshold control, whereas the rest of the lines show the results *with* threshold control.

Firstly, Figure 3–13 and Figure 3–14 show the packet loss probabilities of all the five classes with different threshold values ($B_2 = 0.2, 0.4, 0.6$) in the Gold queue.

Figure 3–13 demonstrates that the effect of the different threshold values in Gold queue on the performance of packet loss probability is significant for all the classes in Gold queue (Class 1 and Class 2). In this case, when the threshold in Gold queue increases, the loss probability of Class 1 increases whereas the loss probability of Class 2 decreases. For example, when the offered load is fixed to be 4, as the threshold in the Gold queue increases from 0.2 to 0.6, the loss probability of Class 1 traffic increases from 6.72×10^{-9} to 2.12×10^{-7} and the loss probability of Class 2 traffic drops from 0.24 to 0.21. This is because when the threshold becomes larger, lower priority traffic in the Gold queue (Class 2 in this case) becomes less likely to be dropped. Hence it will affect the loss probability of the higher priority traffic (Class 1 in this case) adversely.

Furthermore, it can be observed that, *with* threshold control, the loss probability of Class 1 decreases significantly at the expense of moderate increase of loss probability of Class 2. For example, when the offered load is fixed to be 3, *with* the threshold $B_2 = 0.6$

implemented in the Gold queue, the loss probability of Class 1 drops from 0.016 to almost 0. On the other hand, the loss probability of Class 2 increases slightly from 0.016 to 0.038.

Figure 3–13(a): Class 1 loss probability with varying thresholds in Gold queue

Figure 3–13(b): Class 2 loss probability with varying thresholds in Gold queue

Figure 3–13: Effect of Gold Queue Thresholds Control on Loss Probability of Gold Queue

As expected, Figure 3–14 shows that the threshold control implemented in the Gold queue has no effect on the loss probabilities of all the traffic classes in the Silver queue, since the traffic composition of all the traffic classes as well as the thresholds in Silver queue are fixed in this scenario.

Figure 3–14(a) Class 3 loss probability with varying thresholds in Gold queue

Figure 3–14(b): Class 4 loss probability with varying thresholds in Gold queue

Figure 3–14(c): Class 5 loss probability with varying thresholds in Gold queue

Figure 3–14: Effect of Gold Queue Thresholds Control on Loss Probability of Silver Queue

Correspondingly, Figure 3–15 and Figure 3–16 show the throughputs of all the five classes with different threshold values ($B_2 = 0.2, 0.4, 0.6$) in the Gold queue. All the figures show the same trend. When the offered load is low, the throughputs *with* threshold control match the throughputs *w/o* threshold control. However, when the offered load increases beyond a certain value, the throughputs *with* threshold control differentiate themselves from the throughputs *w/o* threshold control.

We also observe that the higher priority the classes are, the later the deviation processes start. For example, when offered load goes beyond 2.5, deviation of the throughputs for those *with* and *w/o* threshold control appears for Class 2. In the same scenario, the deviation starts when the offered load goes beyond 0.75 for Class 3.

Figure 3–15 and Figure 3–16 also show that, *with* threshold control, throughput performance of higher class traffic within the same queue improves at the expense of degradation of the throughput performance of lower priority class traffic.

Figure 3-15(a): Class 1 throughput with varying thresholds in Gold queue

Figure 3-15(b): Class 2 throughput with varying thresholds in Gold queue

Figure 3-15: Effect of Gold Queue Thresholds Control on Throughput of Gold Queue

Figure 3-16(a): Class 3 throughput with varying thresholds in Gold queue

Figure 3-16(b): Class 4 throughput with varying thresholds in Gold queue

Figure 3–16(c): Class 5 throughput with varying thresholds in Gold queue

Figure 3–16: Effect of Gold Queue Thresholds Control on Throughput of Silver Queue

Next, Figure 3–17 and Figure 3–18 illustrate the packet loss probabilities of all the five traffic classes with different threshold values ($B_5 = 0.1, B_4 = 0.2$; $B_5 = 0.2, B_4 = 0.4$; $B_5 = 0.3, B_4 = 0.6$) in the Silver queue.

Figure 3–17 demonstrates that the effect of the different threshold values in Silver queue on the packet loss probability is significant for all of the classes in the Silver queue, i.e. Class 3, Class 4 and Class 5. In this case, when the thresholds in Silver queue increase, the loss probability of Class 3 increases, but the loss probabilities of Class 4 and Class 5 decrease. For example, when the offered load is at 0.8, as the thresholds in the Silver queue increase from ($B_5 = 0.1, B_4 = 0.2$) to ($B_5 = 0.3, B_4 = 0.6$), the loss probability of Class 3 increases from 1.05×10^{-12} to 4.27×10^{-11} , the loss probability of Class 4 drops from 2.5×10^{-3} to 8×10^{-4} and the loss probability of Class 5 drops from 0.1259 to 0.0895. This is because when the thresholds become larger, the lower priority traffic (Class 4 and Class 5 in this case) becomes less likely to be dropped. Hence it will affect the loss probability of

the higher priority traffic (Class 3 in this case) adversely. Furthermore, it can be observed that, with threshold control, the loss probability of Class 3 and Class 4 decreases significantly at the expense of moderate increment of loss probability of Class 5.

Figure 3-17(a): Class 3 loss probability with varying thresholds in Silver queue

Figure 3-17(b): Class 4 loss probability with varying thresholds in Silver queue

Figure 3-17(c): Class 5 loss probability with varying thresholds in Silver queue

Figure 3-17: Effect of Silver Queue Thresholds Control on Loss Probability of Silver Queue

Figure 3-18, as expected, shows that the threshold control occurring at the Silver queue has no effect on the loss probabilities of the traffic classes in the Gold queue, since the traffic composition of all the traffic classes as well as the thresholds in Gold queue are kept same in this scenario.

Figure 3-18(a): Class 1 loss probability with varying thresholds in Silver queue

Figure 3–18(b): Class 2 loss probability with varying thresholds in Silver queue

Figure 3–18: Effect of Silver Queue Thresholds Control on Loss Probability of Gold Queue

Correspondingly, Figure 3–19 and Figure 3–20 show the throughputs of all the five classes with different threshold values ($B_5 = 0.1$, $B_4 = 0.2$; $B_5 = 0.2$, $B_4 = 0.4$; $B_5 = 0.3$, $B_4 = 0.6$) in the Silver queue. Similar to our findings in Figure 3–15 and Figure 3–16, all the figures show the same trend. When the offered load is low, the throughputs *with* threshold control match the throughputs *w/o* threshold control. However, when the offered load increases beyond a certain value, the throughputs *with* threshold control differentiate themselves from the throughputs *w/o* threshold control. Furthermore, the higher priority the classes are, the later the deviation processes start. Finally, the figures also show that, *with* threshold control, throughput performance of higher class traffic within the same queue improves at the expense of degradation of the throughput performance of lower priority class traffic.

Figure 3-19(a): Class 3 throughput with varying thresholds in Silver queue

Figure 3-19(b): Class 4 throughput with varying thresholds in Silver queue

Figure 3-19(c): Class 5 throughput with varying thresholds in Silver queue

Figure 3-19: Effect of Silver Queue Thresholds Control on Throughput of Silver Queue

Figure 3-20(a): Class 1 loss probability with varying thresholds in Silver queue

Figure 3–20(b): Class 2 loss probability with varying thresholds in Silver queue

Figure 3–20: Effect of Silver Queue Thresholds Control on Throughput of Gold Queue

In summary, we have demonstrated in this subsection that threshold control indeed has effect on the packet loss probabilities and throughputs of all the five traffic classes. With threshold control, the loss probabilities and throughputs of class with higher priority improves significantly with moderate increase of lower priority class in the same queue. Furthermore, varying thresholds in one of the queue does not affect the performance of the other queues but causes apparent effect to the performance of the queue itself. More specifically, with larger threshold value, lower priority classes will experience less loss at the expense of increase in the losses of higher priority classes.

3.8.3. Effect of Gold Queue Load on the Delay Performances of the Traffic Classes

In this part of thesis, we would like to investigate the effect of Gold queue load on the delay performances of the traffic classes in both the Gold queue and the Silver queue. Figure 3–21 demonstrates the average delay versus the offered load with varying load of the second highest priority class – Class 2. In this scenario, the load from Class 1 is fixed to be 5% of the total offered load and Class 2 increases from 15% to 25% at a step of 5% of the total offered load. For the rest of the load, Class 3 takes 20%, Class 4 takes 30% and Class 5 take 50% each. The buffer size of the Gold queue is 0.32% of the output link capacity and the buffer size of the Silver queue is 1.6% of the output link capacity. Gold queue threshold is set to be 40% (B_2) of the total Gold queue capacity and Silver queue thresholds are set to be 30% (B_5) and 60% (B_4) of the total Silver queue capacity.

In this case, with same offered load, the improvement of average delay for Class 1 is significant when the load from Class 2 decreases. For example, when offered load is 1.5 and the load of Class 2 decrease from 25% to 20% to 15%, the delay of Class 1 decreases from 7.42×10^{-6} to 2.03×10^{-7} to 3.83×10^{-11} . Also, the delay of Class 1 appears consistently changing and slowly converging. On the contrary, the improvements of the average delays for Class 3, 4 and 5 are insignificant. Finally, it is obvious that the delay performance of Class 1 is much better than the three classes in Silver queue, because the Silver category customers can only get the service when the Gold queue is empty. For example, it can be seen that when the offered load reaches 0.36, the delays for the Silver classes all go beyond 1. However, the delay for Class 1 is still negligible (4.1×10^{-11}) when the offered load is 1 and the load of Class 2 is 0.25. But since delay is less important for Silver queue classes, we can have less concern on it for Silver–category customers generally have low expectation in delay.

Figure 3-21(a): Class 1 delay with varying Class 2 traffic load

Figure 3-21(b): Class 3 delay with varying Class 2 traffic load

Figure 3-21(c): Class 4 delay with varying Class 2 traffic load

Figure 3-21(d): Class 5 delay with varying Class 2 traffic load

Figure 3-21: Effect of Class 2 Traffic Load on Average Delay of Other Four Classes

3.9. EXTENSION TO ALL–OPTICAL OPS NODES

In this part of the research, we would like to extend our network node model to those to be used in all–optical OPS networks [186]–[187]. The reason is that all–optical networks are widely considered as the promising candidate for the next generation network (NGN) development, which offer line and switch capacities greatly exceeding those of existing facilities. As we mentioned in section 2.5, OPS networks provide tremendous improvement upon the capacity over the legacy electronics–based network. Moreover, the motivation of this part of research is because we believe that the goal of the OPS network should also include the ability to provide service differentiation and QoS control just as its electronics–based counterpart.

As we mentioned in section 2.5, the major obstacle faced in building all–optical OPS networks is the optical buffer development, although optical buffering is the most suitable candidate for solving contention at the OPS node. Among the two choices of the fundamental building blocks for optical buffer, i.e. slow light –based (SL – based) [142]–[143] and fiber delay line –based (FDL – based) [145]–[147], the bandwidth delay product issue of the SL – based buffer constrains us only resort to the FDL–based optical buffer [148]. Recirculation FDL, an improvement over the basic FDL, greatly reduce the length of the fiber for the same amount of delay required [155]–[157]. As we conclude in section 2.5.3, the state-of-the-art physical layer implementation of optical buffer, which is realizable for usage and correct operation, is recirculation–enabled FDL–based buffer, where one cycle of the FDL should support the maximum size of optical packets in the network.

In this part of thesis, we propose an optical buffer system to be used at the output port of the optical packet switch, which can be implemented by the state-of-the-art optical buffer technology in near term ([145]–[147], [135]–[163]). This architecture is designed in such a way that it exactly resembles its electronic counterpart as shown in section 3.5. In addition, we demonstrate how to implement the threshold control in such an optical buffer

to differentiate QoS at the OPS nodes. Here, we use optical queueing system and optical buffer system interchangeably.

Notice that the performance analyses done in section 3.7.1 and 3.7.4 are for variable-sized IP packets traveling through electronic nodes in copper network. Hence the models for the two IntServ queues, i.e. the Premium queue and the Bronze queue, are based on M/M/1 queueing model with limited queueing capacity. In this section, however, our operating scenario is that fixed-sized optical packets travel through OPS nodes in an optical network. Hence, the performance analyses for the Premium queue and Bronze queue, that handle the IntServ GS traffic and IntServ CLS traffic respectively, should be based on M/D/1 queue with limited buffer capacity, where the interarrival time is exponentially distributed and the service time is constant. The performance analysis for an M/D/1 with limited queueing capacity can be found in [153] [189]-[190]. The performance analyses of the Gold queue and Silver queue for the proposed architecture in this section are the same as the performance analyses done in section 3.7.2 and 3.7.3, which make use of fluid flow model to model the statistical multiplexer. This is because the fluid flow model has been successfully applied for the modeling and analysis of ATM networks, where each cell is of fixed size – 53 bytes (please refer to section 2.3.3 for more details). Therefore, it is perfect for the modeling and analysis of the fixed-sized optical packets in our operating scenario of interest.

3.9.1. Proposed Optical Buffer Architecture

In this section, we first propose an optical buffer system that is functionally equivalent to a First-In-First-Out (FIFO) electronic buffer. With this as a solid foundation, we then propose a more complicated optical queueing system that is functionally equivalent to its electronic counterpart as shown in Figure 3–4.

3.9.1.1. Equivalent First In First Out (FIFO) Optical Buffer – Single Wavelength Case

The first step towards building our proposed buffer is to build a buffer system, as shown in Figure 3–22. If it is used at the output port of an OPS node, it will behave the similar way as an FIFO electronic buffer used in an electronic network node. More importantly, the design of the optical buffer system only makes use of the technologies that are viable by today’s physical layer development. For the simplicity of illustration, the model shown is for single wavelength operation and the effective buffer size of the system is B .

Figure 3–22: Proposed optical buffer functioning as FIFO buffer at one output port – Single Wavelength Case

This optical queueing system takes optical packet from the “Input to Buffer Stage” and stores optical packets one by one from Line 0 down to Line $B-1$. D means the delay for one optical packet. MU denotes the monitoring unit. As shown in Figure 3–22, this buffer system consists of two stages of recirculation FDLs. Stage 1 FDLs, by default, recirculate packets as many cycles as they are capable of before forwarding them to the stage 2 FDLs along the corresponding line. That means without any trigger from the MU, the optical packets will be recirculated forever before being forwarded. This is very similar to the behavior of electronic buffer, which will not force the packets to output if the output line is

not available. When MU detects that stage 2 FDLs are all empty, and output link is available, stage 1 FDLs will be triggered to forward all the packets that are stored in stage 1 recirculation FDLs to their corresponding recirculation FDLs in stage 2. Stage 2 FDLs each has a fixed number of cycles ($0 \sim B-1$) that the optical packets have to be recirculated before their departure. In this way, optical packets that output from stage 2 FDLs can be ensured not to contend with each other.

In the actual implementation, if the recirculation circuit is purely passive, the limit of the recirculation cycles will depend on the loss of the signal during the recirculation process. To overcome the limit of the signal loss during the recirculation process, EDFAs can be incorporated to recover the signal strength. However, in this case the limit of the recirculation cycles will be decided upon the noise accumulation by the EDFAs. Fortunately, the number of cycles, which the optical packets can be recirculated before being degraded below the level of acceptance, keeps improving throughout the years [127] [135] [158]-[159].

Operation Principle

Let us consider three cases in order to fully understand the operation principle of the output optical queueing system.

Case (a): if the network is lightly loaded and the arrival rate is less than the service rate, the optical packets will always enter the line 0 of buffer, simply bypass recirculation stages and output directly. Hence there is no delay.

Case (b): when a burst of optical packets enters the buffer, they will be stored in stage 1 FDLs one by one from line 0 orderly down to line $B-1$. Note that according to the state-of-the-art technology, each recirculation line can only allow one packet to travel at one time (please refer to section 2.5.3.2). Hence, if the burst has more than B optical packets, the excess packets will have to be dropped. This means, the maximum capacity of this optical

buffer system is B . When MU detects that the output line is free for this queue now, it will trigger stage 1 FDLs to forward the optical packets to their corresponding stage 2 FDLs. From then on, the queue can output packets according to the sequence from 0 to $B-1$ one at each time because of the fixed number of cycles that a packet has to travel in stage 2.

Case (c): When stage 2 is outputting optical packets, MU, as a counting device, counts the number of FDL lines that has been outputted at stage 2 and sends back this information to the “Input to Buffer Stage”. When there are some new arrivals to be stored in the buffer, they will be stored in stage 1 FDLs upon arrival under the control of the “Input to Buffer Stage”. The “Input to Buffer Stage” will control the number of new arrivals that will be stored in stage 1 FDL, to ensure the total number of optical packets stored in the buffer system does not exceed the buffer size B . For example, if $B = 30$ and the first burst of 10 optical packet comes when output line is not available for this queue, which will result in line 0 to line 9 of stage 1 FDLs being filled up (i.e. 10 optical packets were stored). Later, MU detects that output line is available and stage 2 FDLs are all empty. So at time T_0 , stage 1 recirculation FDLs then are triggered to forward the 10 packets to line 0 ~ line 9 of stage 2 recirculation FDLs. Suppose that the packets stored in line 0, line 1, line 2 and line 3 then are outputted successfully. When line 4 is about to output its contents right after T_0+3D , a new burst comes. Then MU counts that the stage 2 FDLs have already outputted 4 optical packets since line 4 are being read, together with the information that stage 2 FDLs currently are occupied up to line 9. Hence MU then sends back to the “Input to Buffer Stage” that the number of packets being stored in stage 2 FDLs is N_2 ($N_2 = 9 - (4 - 1) = 6$). Hence the “Input to buffer stage” will only allow up to N_1 ($N_1 = B - N_2 = 30 - 6 = 24$) optical packets to enter the optical queue. Subsequently, as more packets are outputted from stage 2 FDLs, more packets can be stored in stage 1 FDLs. By operating in this way, the buffer capacity can be ensured to be B . Once MU finds all the packets in stage 2 have been sent, it will trigger those stored in stage 1 FDLs to be forwarded to their corresponding stage 2 FDLs.

Based on the described operation principle, it is obviously that our proposed optical buffer system operates in a manner that is analogous to an electronic buffer implementing FIFO with the maximum capacity being B packets.

Multiple Thresholds Control Implementation

The operation principle of our optical buffer together with the usage of monitoring unit uniquely provides the information of buffer usage, which can be fed back to the access unit inside the “Input to Buffer Stage”. Service differentiation in terms of multiple loss priority [69] [178] can be implemented by partitioning the buffer into as many levels as of occupancy as there are for the priority classes.

Therefore, its equivalent electronic buffer can be drawn as shown in Figure 3–23. In this case, the equivalent electronic queue having L levels of thresholds for packet control, denoted as $B_1, B_2, \dots, B_L$, is fed with packets with different priorities (with arrival rate being $A_1, A_2, A_3, \dots, A_L$, where the subscript 1 denotes the highest priority and the subscript L denotes the lowest priority in the queue). Obviously, for the highest priority traffic, its threshold is equal to the entire buffer size, i.e. $B=B_L$. For the simplicity of demonstration, we assume no wavelength conversion will be done at the node. Hence the service rate (C) will be the speed of a single wavelength (C_λ).

Figure 3–23: Equivalent electronic buffer of Figure 3–22

In fact, our work can be generalized to the case where tunable wavelength converters (TWCs) will be located at the “Input to Buffer Stage” to switch incoming traffic to a set of wavelengths within a wavelength band. In that case, if the wavelength converter has the information of output wavelength occupancy and is able to always switch the incoming

optical packets to the wavelength with the lowest buffer occupancy, the number of packets that can be simultaneously transmitted out of the buffer is increased. To the best case, the effective service rate would be increased to the product of the number of wavelengths in the band (N_w) and the speed of one wavelength (C_λ), i.e. $C = N_w C_\lambda$. If the control unit of the “Input to Buffer Stage” only allows to start to store optical packets to the wavelengths contained in line $b+1$ after all wavelengths have been filled up in line b , where $b \in [0, B-2]$, the effective buffer size will be the product of N_w and number of recirculation FDLs in each stage (B), i.e., $N_w B$.

3.9.1.2. Proposed Optical Buffer to be used at the Output Port

Based on the architecture of optical buffer system proposed in section 3.9.1.1, we will enhance the optical buffer system so that the OPS node, equipped with our proposed optical buffer system at the output port, will function similarly to its electronic counterpart as shown in Figure 3–3. In another word, we propose here an optical buffer system which will analogous to the electronic queueing system proposed in Figure 3–4. For the simplicity of illustration, let us consider the single wavelength operation.

IntServ traffic can be handled by some dedicated wavelengths or even fibers, which do not incorporate the multiple thresholds control mechanism to provide QoS differentiation. Hence, for the simplicity of the illustration, we emphasize on the two queues that handles the DiffServ traffic, which are more complicated. Those are the Gold Queue for DiffServ high priority traffic and the Silver Queue for DiffServ low priority traffic as shown in Figure 3–4. The buffer capacity for Gold Queue and Silver is B_I and B_{LI+1} respectively. As we know, Gold Queue has higher priority than Silver Queue so Silver Queue can only get the service when Gold Queue is empty, if we consider Premium Queue (taking input from IntServ GS traffic) to be always empty.

Buffer System Architecture

Figure 3–24 shows our proposed optical buffer system to be located at each output port of the OPS node. This shows the case where the largest number of FDLs is needed. Later, we will show the case where a smaller number of FDLs are needed, depending on the system requirement and hardware viability. In this case, still there are two stages of recirculation FDLs. The upper B_{b1} lines belong to the Gold Queue and the lower B_{bL+1} lines belong to the Silver Queue. Without TWCs at the input stage, the buffer system is for single wavelength operation. Hence the effective service rate is the rate of one wavelength (C) and the number of FDLs at each stage is equal to the corresponding buffer size, i.e., $B_{b1} = B_1$ and $B_{bL+1} = B_{L+1}$. For the simplicity of illustration, the model shown is for single wavelength operation.

Figure 3–24: Proposed Optical Buffer System to be used at the Output Port of OPS to Provide Service Differentiation

Optical Buffer to Handle DiffServ High Priority Traffic

This upper B_{b1} lines (i.e. the Gold Queue) take optical packets, belonging to DiffServ high priority, from the “Input to Buffer Stage”. In a single wavelength operation, it will store optical packets one by one from Line 0 orderly down to Line $B_{b1}-1$. Same as the one shown in Figure 3–22, this Gold Queue consists of two stages of recirculation FDLs. Stage 1 FDLs, by default, recirculate packets as many cycles as they are capable of until MU detects that stage 2 FDLs for the Gold Queue are all empty, and output link is available. Stage 1 FDLs of Gold Queue will be triggered to forward all the stored packets to their corresponding FDLs in stage 2 of the Gold Queue. Stage 2 FDLs each has a fixed number of cycles ($0 \sim B_{b1}-1$) that the optical packets have to be recirculated before their departure.

Optical Buffer to Handle DiffServ Low Priority Traffic

The lower B_{bL1+1} lines (i.e. Silver Queue) takes optical packets, belonging to DiffServ low priority, from the “Input to Buffer Stage”. In a single wavelength operation, it will store optical packets one by one from Line 0 orderly down to Line $B_{bL1+1}-1$. Similar to the Gold Queue buffer system, this Silver consists of two stages of recirculation FDLs. Stage 1 FDLs, by default, recirculate packets as many cycles as they are capable of before forwarding them to the stage 2 FDLs along the corresponding line. Stage 2 FDLs, by default, each has a fixed number of cycles ($0 \sim B_{bL1+1}$) that the optical packets have to be recirculated before their departure. Notice the number of recirculations in stage 2 FDLs is not fixed, which may be changed by the control unit of stage 2 FDLs in some operation scenarios.

Operation Principles

Firstly, for the operation principle and multiple thresholds control for different packet loss priorities of the Gold Queue optical buffer system, please refer to the operation principle stated in 3.9.1.1. They are exactly the same. Next, we consider several scenarios to fully understand the operation principle of the Silver Queue.

Case (a): If initially both Gold Queue and Silver Queue are empty and DiffServ low priority optical packets enter the Silver Queue. During the entire operation, there is no DiffServ high priority optical packets arrive, i.e. Gold Queue remains empty. In this case, the operation of the Silver Queue will be exactly same as the Gold Queue.

Case (b): If initially Gold Queue has content stored, Gold Queue will occupy the output link. DiffServ low priority packets arrive at the empty Silver Queue and start to fill up the Silver queue orderly. In this case, stage 1 FDLs of Silver Queue will hold the number of packets up to its maximum capacity, i.e. B_{bL1+1} . The control units of stage 1 FDLs of Silver Queue will only be triggered when MU detects that Gold Queue is empty and output port is available. Then stage 1 FDLs of Silver Queue will forward all it has stored so far to the stage 2 FDLs of Silver Queue along their corresponding lines. From then on, the queue can output packets according to the sequence from 0 to $B_{bL1+1}-1$ one at each time because of the fixed number of cycles that a packet has to travel in stage 2 by default.

Case (c): When Silver Queue is outputting and there are new arrivals at Gold Queue, the Gold Queue packets will get service immediately. In this case, MU will detect this non-empty state of Gold Queue. Then MU will record the position of the stage 2 FDL line of the Silver Queue that is outputting and trigger the Silver Queue stage 2 FDLs to forward the contents that all stage 2 FDLs contain to their corresponding backup lines. Please refer to case (c) of section 3.9.1.1 for the method of recording the position of outputting stage 2 FDL in the Silver queue. The backup lines will send back the optical packets to their corresponding stage 1 FDLs and keep them waiting there till the Gold queue is empty and output port is available for the Silver Queue again.

Case (d): When Gold Queue is empty again, stage 1 FDLs of Silver Queue will only forward the packets from the backup process to the stage 2 of FDLs of Silver Queue. Moreover, the number of cycles that the FDL will recirculate them before outputting them

will become the default number minus the pointer position that MU denoted during the backup process. In this way, the Silver Queue packets will not get further delayed.

Case (e): When the backup data is cleared, the pointer will be reset to 0 again and if the Gold Queue is still empty, MU will then trigger Silver Queue stage 1 FDLs to forward all the contents they store to their corresponding lines in stage 2.

Multiple Thresholds Control Implementation

As we can see the Gold Queue and Silver Queue are of different priority in terms of delay performance since Gold Queue will always get service before Silver Queue. Next, inside each queue, QoS differentiation in terms of multiple loss priority [69] [178] can be implemented by partitioning the buffer into as many buffer occupancy levels as there are for the loss priority classes.

Therefore, its equivalent electronic buffer can be drawn as shown in Figure 3–25. As stated in section 3.5, multiple loss priority [69] [178] is implemented using threshold control mechanism over the partitioned buffers. The Gold queue having $L1$ levels of thresholds for packet loss control, denoted as $B_1, B_2, \dots, B_{L1}$, is fed with DiffServ High Priority traffic (with arrival rate equal to $\lambda_1, \lambda_2, \dots, \lambda_{L1}$, where the subscript 1 denotes the highest priority and the subscript $L1$ denotes the lowest priority in Gold Queue). Only when both the Premium queue and Gold queue are empty (in this case, Gold Queue only) the Silver queue can get the service. The Silver queue takes input from DiffServ Low Priority sources (with arrival rate equal to $\lambda_{L1+1}, \lambda_{L1+2}, \dots, \lambda_{L2}$, where the subscript $L1+1$ denotes the highest priority and the subscript $L2$ denotes the lowest priority in Silver Queue) and has $L2 - L1$ levels of thresholds, denoted as $B_{L1+1}, B_{L1+2}, \dots, B_{L2}$, for packet loss control over the partitioned buffer.

Figure 3-25: Equivalent electronic buffer of Figure 3-24

3.9.1.3. Improvement of Our Proposed Optical Buffer System

The proposed model shown in section 3.9.1.2 uses the largest number of FDLs possible for the equivalent electronic buffer model as shown in Figure 3-25 and it is for single wavelength operation. On the other hand, the requirement of the physical layer technology is the simplest, which means it can be implemented in near future. However, with suitable hardware and control unit available, the number of FDLs required can be greatly reduced.

Let us recall that the upper B_{b1} lines belong to the Gold Queue and the lower B_{bL1+1} lines belong to the Silver Queue. Without TWCs at the input stage, the buffer system is for single wavelength operation. Hence the effective service rate is the rate of one wavelength (C) and the number of FDLs at each stage is equal to the corresponding buffer size, i.e., $B_{b1} = B_1$ and $B_{bL1+1} = B_{L1+1}$. With the help of TWCs located at the “Input to Buffer Stage”, WDM signals can be handled. Let us assume that TWCs will assign the packets to the whole wavelength range of a fiber (W). In that case, if the TWCs have the information of output wavelengths occupancy and are able to always switch the incoming optical packets to the wavelength with the smallest number of buffer occupancy, the number of packets that can simultaneously be transmitted out of the buffer is increased. To the best case, the effective service rate would be increased to product of the number of wavelengths in the band (N_w) and the speed of one wavelength (C_λ), i.e. $C = N_w C_\lambda$. The control unit of the “Input to Buffer Stage” only allows starting to store the optical packets to the wavelengths

contained in one line after all wavelengths have been filled up in the previous line. In order to provide Gold Queue with B_l capacity and Silver Queue with B_{Ll+1} capacity, the number of FDLs at each stage should be set as $B_{b1} = B_l / W$ and $B_{bLl+1} = B_{Ll+1} / W$. Therefore, by using TWC, the number of recirculation FDLs used is much less than the amount required for single wavelength operation with same buffer size requirement.

In fact, the output buffer system will look more compact if the buffer size of Gold Queue and Silver Queue are same. In that case, one set of 2-stage FDLs is sufficient to handle the traffic entering these two queues. However, in that case, the wavelengths contained in one fiber should be divided into two wavelength bands. Each queue will take one wavelength band. By dividing the number of wavelengths to be same, the effective service rates of these two queues are equal, which can be different if the number of wavelengths is divided otherwise. TWCs will need to be able to convert wavelengths to a specific set of wavelengths rather than the entire band contained in a fiber. The requirement for the monitoring unit and control unit will become more stringent because more intelligence are needed, which can be hardly implemented in near future.

3.10. SUMMARY

Considering the increasing expectations being placed on networks, the benefits provided by MPLS make it a promising technology. MPLS can enhance network scalability and performance while providing advanced services in a simple package. Therefore with desire to upgrade an existing network or plan a next-generation network (e.g., OPS network), MPLS is a concept to bear in mind.

This chapter focused on the issues to enable the interoperation of IntServ, DiffServ and MPLS while improving the QoS at IP packet level. We showed the significance of MPLS label, proposed the traffic classification and showed the possible service mapping among

the IntServ, DiffServ and MPLS network domains. The most important portions were the modeling and analysis of the four-queue system in the environment of mixed traffic from IntServ and DiffServ customer edges. We presented an efficient model for investigation of the related performance issues. However, we mainly focused on the MPLS network and the procedures were also analytical in nature. With this approach, one can obtain the results quickly without resorting to simulation. The numerical results illustrated the performance parameter in terms of loss probability, throughput and delay with combinations of different traffic load proportions and various traffic priorities. We concluded that in order to achieve the desired packet loss probability for different traffic classes, both the overall traffic load in the network and the traffic load offered by the higher-priority classes need to be properly controlled. Also, with threshold control, the loss probability of higher priority class improves significantly with moderate increase of lower priority class in the same queue. Last but not the least, we presented a realizable optical buffer system which can be used at the OPS node to extend our work in the QoS control in all-optical OPS backbone networks.

Chapter 4. Modeling and Transmission

Performance Analysis of All-Optical WDM Backbone Network

4.1. INTRODUCTION

The development of optical communication technologies, especially wavelength division multiplexing (WDM) technology, programmable optical crossconnects (OXC)s and optical amplifier, has provided huge bandwidth up to terabits per second for the transmission of Internet protocol (IP) traffic over optical fiber. A lot of research and development works for the next generation network (NGN) in recent years have been focusing on finding an efficient way to integrate IP and WDM technologies [9]-[12][191]. Generalized Multi-Protocol Label Switching (GMPLS [192]) has been standardized by IETF to be the most promising approach to the integration of IP and WDM. GMPLS is a multipurpose control plane paradigm in support of not only the devices that perform packet switching, but also the devices that perform switching in time, wavelength, and space domains [13]-[16].

As shown in Figure 1-2, the expected technology layer evolution of backbone IP networks evolves from IP over ATM over SONET/SDH over WDM [17] to IP with GMPLS over SONET/SDH over WDM [18], and finally, to IP with GMPLS directly over WDM [15]. This promising and ultimate network architecture should utilize an all-optical transparent model on its physical layer [79]. In this way, the network is able to offer lower cost and higher speed than the existing opaque model [80], where optical-electrical-optical (OEO)

conversion is employed at every node, and the translucent model [82], where OEO conversions are only performed at some of the nodes [193].

However, the removal of the intermediate ATM and/or SONET/SDH protocol layers, on the other hand, implies the lack of packet error detection and/or correction capability. Moreover, without the isolation provided by the intermediate layers, traffic patterns carried in the optical domain and IP domain are now correlated [23]. Under such circumstances, the bursty nature of IP traffic may have effect on the performance of the optical channels. On the other hand, the impairments of the optical paths and optical nodes may have noticeable influence on performance of IP traffic in the network layer. At the same time, optical signal degradation can no longer be cleaned up by the OEO conversion in the intermediate nodes due to the transparency of the physical layer. As a result, the architecture of the system itself raises the coupling issue of optical and network layers that just begins attracting the attention of researchers. Few works have been done on impairment-aware wavelength routing assignment (RWA), which made modifications by adding bit error rate (BER) due to optical layer impairment as the extra criteria to be satisfied after choosing a candidate route from the well-know RWA algorithm [24]-[25]. As a result, the candidates with BER higher than a threshold (e.g., 10^{-12}) will not be considered for making the traditional assumption of perfect optical transport still valid.

However, there might be the case that eventually no candidate will satisfy the criteria with the ever-increasing end-to-end distance, data rate per wavelength, number of wavelengths in a fiber, and/or input optical power. What shall we do then? Shall we give up the transparency by going back to the old days OEO memory or keeping data rate and wavelength number sufficiently low in spite of the ever-increasing bandwidth demand from the users, since we cannot change the physical distance between end nodes? Or shall we live with the imperfections and carry on the analysis or design by counting in the adverse effects of optical characteristics into the performance evaluation?

This chapter aims to analyze the coupling effect between the optical layer transmission and network layer performance. Section 4.2 states the major assumptions of our work, followed by section 4.3, which highlights the necessity of our work by illustrating a simple example. Section 4.4 shows the per-hop model and the detailed structure of optical cross connect (OXC). The major work lies in section 4.5, where a three-step cross-layer approach is elaborated for the analysis of coupling effect between the optical layer and network layer in an IP over WDM all-optical network. A comprehensive study of scenarios with different transport links and nodes configurations are demonstrated in section 4.6. The comparative results shown provide reader a flavor of how the difference in physical infrastructure of NGN system will impact differently on the effective network performance when we choose to live with the imperfection in the ultimate fully transparent all-optical network. Section 4.7 is inspired by the results obtained in the previous sections and presents a strategic view of intermediate – step design of NGN backbone networks. Last but not the least, section 4.8 concludes the entire chapter.

4.2. ASSUMPTIONS OF NETWORK CHARACTERISTICS

Before highlighting the problem in the existing analyses, four assumptions are made for the operating environment and functionality of components of the network to be analyzed.

- (1) This work is done for the performance analysis of NGN – IP/GMPLS over WDM backbone network without any intermediate protocol layers. In this work, the physical layer structure is not limited to dual-ring, the analysis can also be used for structure with more diversity, e.g., mesh.
- (2) In our NGNs, all backbone devices are used in electrical domain merely for network layer function, e.g., routing, setup of the label switched paths (LSPs), etc.

(3) In the actual traffic transmission, all the optical backbone devices only perform the simplest function as fast switching devices and all core processing is done in optical domain, e.g., switching, amplification, etc. Packet correction is done in network layer or even higher layer, thus this function is performed at the network layer of access ends or the application layer of users.

(4) At the moment, optical queueing effect is not considered; hence the loss due to contention for output port will be considered “blocked” and taken care by blocking probability P_b , which is a network layer metric.

4.3. PROBLEM HIGHLIGHTS

Consider one simple case, as shown in Figure 4–1, for the illustration of the necessity of considering optical impairments when evaluating the performance in IP over WDM network. $S_1, S_2, \dots, S_N$ are the sources, which send packets to destination D through the forwarding node F .

Figure 4–1: Simple Network

In order to evaluate the total throughput (TH) of the network, most people will follow the

normal procedure and get the result as $TH = \sum_{n=1}^N TH_n$. The throughput (TH_n) from source

node (S_n) to destination is a function of blocking probability (P_{bn}) and arrival traffic

(Arr_n) at S_n by the expression – $TH_n = f(P_{bn}, Arr_n)$, $n \in [1, N]$. If we assume that the

arrival is independent of the blocking probability, the expression can be further modified

as $TH_n = f(P_{bn})f_{arr}(Arr_n)$, $n \in [1, N]$, where $f_{arr}(Arr_n)$ is a function of Arr_n . However,

these analyses always take one assumption for granted – the data transmission in underlying optical network is perfect.

Consider one scenario in our simple network where this assumption fails functioning. Since the network is IP/GMPLS direct over WDM without intermediate layers such as SONET/SDH or ATM, error detection and correction capability of SONET/SDH encapsulation is no longer used. Now, one LSP has been successfully set up between S_n and D and one packet has been sent from S_n and D by using GMPLS label, but the content of the packet is corrupted, which results in useless information received by D . In this case, we should not consider this packet contributes to the throughput of the network. Hence the effective throughput in our example by taking into consideration of the optical impairment effect should be further modified as,

$$TH_E = \sum_{n=1}^N TH_E_n = \sum_{n=1}^N f(P_{bn})f_{arr}(Arr_n)f_{opt}(BER) \quad (4-1)$$

where $f_{opt}(BER)$ is a function of BER that is calculated resultant effect of the impairments along the optical paths and at the optical nodes.

4.4. MODELING

IP/GMPLS over WDM-based network node architecture mainly consists of two portions: GMPLS control module and one switching fabric [194]. Since we are more interested in the adverse effect of optical impairments to the performance of IP traffic transmission based on WDM, our focus lies in the switch fabric portion of the node structure, which is responsible for transmitting packets. Therefore, a comprehensive node model is needed to accurately capture all the optical events occurring during the actual data transmission.

4.4.1. Per-Hop Modeling

When IP/GMPLS over WDM is performed as a large-scale backbone optical network, traffic has to be switched from ingress nodes to egress nodes by OXCs at intermediate nodes. In such a large-scale network, since optical signal traverses numbers of intermediate nodes with long fiber segments, optical amplifiers such as EDFA are usually used to compensate the progressive losses incurred by the signal.

Although WDM fiber segments, OXCs and EDFAs are able to offer enormous capacity, transparent switching and loss compensation, yet they may introduce adverse effects such as phase shift due to SPM, XPM and noises due to FWM, crosstalk, and ASE.

Figure 4-2 shows the modeling of such intermediate hop that captures all the major adverse effects along the optical paths and at the optical nodes, where an LSP is viewed as a series of such nodes interconnected by fiber segments from source to destination.

Figure 4-2(a) shows the case when dispersion shifted fiber (DSF) is used as fiber segment (refer to Chapter 2 for the concept and characteristics of system utilizing DSF span). Each intermediate node consists of one pair of power taps, one EDFA at the input to compensate the fiber attenuation and one EDFA at the output to compensate the node loss. The small residual dispersion in DSF span can be managed with the dispersion equalization method [195].

On the other hand, Figure 4-2(b) shows the case when standard single model fiber (SMF) is used as the fiber segment (refer to Chapter 2 for concept and characteristics of system utilizing SMF+ DCF span). Each intermediate node consists of one pair of power taps, one EDFA at the input to compensate the fiber attenuation, dispersion compensation fiber (DCF) to compensate the GVD, one EDFA to compensate the loss due to DCF and one EDFA at the output to compensate the node loss. In both two models, the OXC is

composed by a stage of demultiplexers (DeMux), a space switch and a stage of multiplexers (Mux). The corresponding gain (G) and loss (L) variables of each optical component are listed in Table 4–1.

Figure 4–2: Per-hop modeling of intermediate nodes – components and their loss/gain parameters

Table 4–1: Optical components and their loss/gain parameters

Component		Loss(L)/Gain(G)
Fiber span (SMF or DSF)		L_f
Tap		L_{tap}
Erbium Doped Fiber Amplifier	Input (Preamplifier)	G_{in}, G_{in1}, G_{in2}
	Output (Postamplifier)	G_{out}
Dispersion-Compensating Fiber (including splicing loss at both ends)		L_{DCF}
Demultiplexer		L_{demux}
Switch		L_{SW}
Multiplexer		L_{mux}

4.4.2. Space Switch Modeling

The OXC is composed by a stage of demultiplexers (DeMux), a space switch and a stage of multiplexers (Mux). The space switch fabric can be of different configurations, which will result in different ways of calculating first-order homowavelength crosstalk. However, the idea and flow of our approach is independent of the configuration of the switch.

In our work, we use the $M \times M$ space switch to provide strictly non-blocking, point-to-point connectivity. The switch consists of a stage of $1 \times M$ active splitters followed by a stage of $M \times 1$ active combiners [95]. The active splitter and combiner portion of the switch consists of $\log_2 M$ stages of 1×2 and 2×1 switch elements respectively. Both the 1×2 and the 2×1 switch elements employ a directional coupler as the basic building block, which in this case we use titanium-diffused lithium niobate (Ti:LiNbO₃) based devices [196]. Note that M is the number of input/output fiber pairs not the number of wavelengths carried in a fiber. As we have mentioned in subsection 2.4.3.1, the ratio of the power at the unselected output port over the total input power in a switch element is referred to as the crosstalk ratio of the switch element (χ). In this kind of switch architecture, the crosstalk must occur in a single splitter element and in a single combiner element in order for an existing signal to interfere with the new signal. This is referred to as the first-order crosstalk effect and the first-order switch crosstalk ratio $\chi_{sw} = \chi^2$ [95]. We will not consider the higher order crosstalk terms (χ^3, χ^4 , etc.), whose effects are negligible compared to the first-order terms.

Figure 4-3(a) shows how first-order homowavelength crosstalk is resulted in a 2×2 switch, where the leaked power of existing LSP 1 from input port 2 to output port 1 at output port 2 interferes with the new LSP 2 from input port 1 to output port 2. Another example is shown in Figure 4-3(b), where there are four input/output port pairs. In this case, the leaked power of existing LSP 3 from input port 3 to output port 3 at output port 2 interferes with the new LSP 4 from input port 2 to output port 2.

Figure 4-3: Crosstalk Example in 2x2 and 4x4 Switches

Figure 4-4 (a) and Figure 4-4 (b) give the schematic views of the 8 x 8 switch and 16 x 16 switch configurations. For more input/output fiber pairs, the space switch block of Figure 4-2 can be modified by replacing the type of the switch blocks. For example, if the number of input/output pair is 4 and the number wavelengths per fiber is 8, the space switch should be composed by 8 blocks of 4x4 switches as shown in Figure 4-3(b).

Figure 4-4: More Switches – 8x8 and 16x16 Structures

4.5. PERFORMANCE ANALYSIS

A three-step approach is employed in the analysis of the performance of IP over WDM network. In the enumeration processes, the effects of signal and noise components at the destination through a set of recursive equations are shown. Next, the BER at the receiver is evaluated. Finally, the network performance is evaluated in terms of effective utilization.

4.5.1. Enumeration Process

Assume an LSP is set up on wavelength λ_i ($\lambda_i = c / f_i$) from node 1 to N . At output of the k th ($k \in [1, N]$) node, the powers of desired signal $p_{sig}(k, f_i)$, crosstalk $p_{xt}(k, f_i)$, ASE noise $p_{ase}(k, f_i)$, and SPM-XPM-aware FWM noise $p_{fwm}(k, f_i)$, can be expressed by a set of recursive equations. Equations (4-2) (a) ~ (4-5) (a) are the set for the systems that utilize node structure as shown in Figure 4-2 (a) [119] [121] [125]. Equation (4-2) (a) shows the output signal power of the k th intermediate node; equation (4-3) (a) shows the output crosstalk noise power of the k th intermediate node; equation (4-4) (a) shows the output ASE noise power of the k th intermediate node; equation (4-5) (a) shows the output SPM-XPM aware FWM signal power of the k th intermediate node. On the other hand, (4-2) (b) ~ (4-5) (b) are the set of equations for the systems with node structure as shown in Figure 4-2(b).

For the corresponding gain (G) and loss (L) variables of each optical component, please refer to Table 4-1.

$$p_{sig}(k, f_i) = p_{sig}(k-1, f_i) L_f(k-1, k) G_{in}(k) L_{demux}(k) \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 \quad (a)$$

$$p_{sig}(k, f_i) = p_{sig}(k-1, f_i) L_f(k-1, k) G_{in1}(k) L_{DCF}(k) G_{in2}(k) L_{demux}(k) \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 \quad (b)$$

(4-2)

where $L_f(k-1, k)$ denotes the attenuation loss from $k-1$ th hop to k th hop.

$$p_{xt}(k, f_i) = p_{xt}(k-1, f_i) L_f(k-1, k) G_{in}(k) L_{demux}(k) L_{SW}(k) \times L_{mux}(k) G_{out}(k) L_{tap}^2 + \sum_{j=1}^{J_k} X_{SW} p_{in}(j, k, f_i) L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap} \quad (a)$$

$$\begin{aligned}
 p_{xt}(k, f_i) &= p_{xt}(k-1, f_i) L_f(k-1, k) G_{in1}(k) L_{DCF}(k) G_{in2}(k) L_{demux}(k) \\
 &\quad \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 \\
 &\quad + \sum_{j=1}^{J_k} X_{SW} p_{in}(j, k, f_i) L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}
 \end{aligned} \tag{b}$$

(4-3)

where $p_{in}(j, k, f_i)$ is the power of j th copropagating signal at the switch shared by the desired signal, X_{SW} is the first – order switch crosstalk ratio [94]-[95], and J_k is the total number of such sources. The value of J_k depends on how many copropagating signals there are together with the signal of interest [124]-[125].

$$\begin{aligned}
 p_{ase}(k, f_i) &= p_{ase}(k-1, f_i) L_f(k-1, k) G_{in}(k) L_{demux}(k) \\
 &\quad \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 \\
 &\quad + 2n_{sp} [G_{in}(k) - 1] h f_i B_o L_{demux}(k) L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap} \\
 &\quad + 2n_{sp} [G_{out}(k) - 1] h f_i B_o L_{tap}
 \end{aligned} \tag{a}$$

$$\begin{aligned}
 p_{ase}(k, f_i) &= p_{ase}(k-1, f_i) L_f(k-1, k) G_{in1}(k) L_{DCF}(k) G_{in2}(k) L_{demux}(k) \\
 &\quad \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 + 2n_{sp} [G_{in1}(k) - 1] h f_i B_o \\
 &\quad \times L_{DCF}(k) G_{in2}(k) L_{demux}(k) L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap} \\
 &\quad + 2n_{sp} [G_{in2}(k) - 1] h f_i B_o L_{demux}(k) L_{SW}(k) \\
 &\quad \times L_{mux}(k) G_{out}(k) L_{tap} + 2n_{sp} [G_{out}(k) - 1] h f_i B_o L_{tap}
 \end{aligned} \tag{b}$$

(4-4)

where B_o is the optical bandwidth, h is plank constant, and n_{sp} is the spontaneous emission factor for EDFAs [123] [125].

$$\begin{aligned}
 p_{fwm}(k, f_i) &= p_{fwm}(k-1, f_i) G_{in}(k) L_{demux}(k) L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 \\
 &\quad + P_{FWM}(k, f_i)
 \end{aligned} \tag{a}$$

$$\begin{aligned}
 p_{fwm}(k, f_i) &= p_{fwm}(k-1, f_i) G_{in1}(k) L_{DCF}(k) G_{in2}(k) L_{demux}(k) \\
 &\quad \times L_{SW}(k) L_{mux}(k) G_{out}(k) L_{tap}^2 + P_{FWM}(k, f_i)
 \end{aligned} \tag{b}$$

(4-5)

where $P_{FWM}(k, f_i)$ is the summation of total FWM power generated by various combinations of interacting signals at f_i ([119]-[121]). It can be expressed as

$$P_{FWM}(k, f_i) = \sum_{f_m=f_j+f_n-f_i} \sum_{f_n} \sum_{f_j} P_{FWMi}(k, f_j, f_n, f_m) \quad (4-6)$$

where f_j , f_n and f_m are signal frequencies, and f_i is the FWM frequency satisfying $f_i = f_j + f_n - f_m$ and $j, n \neq m$. P_{FWMi} is the output power of the FWM product mixing at frequency f_i , given by [107] [119].

$$P_{FWMi}(k, f_j, f_n, f_m) = \frac{1024\pi^6}{n_c^4 \chi_i^2 c^2} d^2 \chi_{1111}^2 \frac{p_{sig}(k, f_j) p_{sig}(k, f_n) p_{sig}(k, f_m)}{A_{eff}^2} \times (-\alpha D_{eff}(k)) D_{eff}(k)^2 \eta_{FWM}(f_j, f_n, f_m, k) \quad (4-7)$$

where $D(k)$ is the fiber distance between $k-1$ th node and k th node,

$D_{eff}(k) = \frac{1 - e^{(-\alpha D(k))}}{\alpha}$ is the effective distance, α is the fiber attenuation coefficient

measured with unit of 1/m, n is the fiber refractive index, A_{eff} is the effective mode area

of the fiber, d is the degeneracy factor ($d = 3$ for $j = n$, $d = 6$ for $j \neq n$), $\chi_{1111} \approx \frac{8n_c n_2}{3}$

(n_2 is the nonlinear-index coefficient) is the third-order nonlinear susceptibility [114] of

the fiber, and $\eta_{FWM}(f_j, f_n, f_m, k)$ is the FWM efficiency given by [119],

$$\eta_{FWM}(f_j, f_n, f_m, k) = \frac{\alpha^2}{\alpha^2 + (\Delta\beta)^2} \left[1 + \frac{4(1 - \alpha D_{eff}(k)) \sin^2\left(\frac{\Delta\beta}{2D(k)}\right)}{\alpha^2 D_{eff}(k)^2} \right] \quad (4-8)$$

where $\Delta\beta$ denotes the phase mismatch and may be expressed in terms of signal frequency differences [99], with one important modification to take into account the phase shift effect of SPM and XPM,

$$\Delta\beta = \left(-2\pi\lambda^2 D_c / c \right) (f_j - f_m)(f_n - f_m) + \left(2\pi f_0 n_2 / c A_{eff} \right) \left(1 - \exp(-\alpha D_{eff}(k)) \right) \times (p_{sig}(k, f_j) + p_{sig}(k, f_n) - p_{sig}(k, f_m)) / (-\alpha D_{eff}(k)) \quad (4-9)$$

where D_c is the chromatic dispersion coefficient.

4.5.2. BER Calculation

Having iterated through the enumeration process till the destination ($k = N$) node, the powers of the composite electrical noises for binary zero and one receptions are obtained. Next, the BER at the receiver takes into account the composite electrical noise powers and the received photocurrent will be evaluated.

The complex amplitude of the electrical field received at the destination node in the presence of crosstalk, ASE and SPM/XPM-aware FWM contributions can be expressed as,

$$E(t) = A(t) \cos(2\pi f_i + \varphi(t)) + E_{xt}(t) + E_{ase}(t) + E_{fwm}(t) \quad (4-10)$$

where $A(t) \cos(2\pi f_i + \varphi(t))$ represents the desired signal component corresponding to frequency f_i , in which $A(t)$ and $\varphi(t)$ are the amplitude and the phase of the signal respectively. $E_{xt}(t)$, $E_{ase}(t)$ and $E_{fwm}(t)$ represent the lightwaves corresponding to the accumulated crosstalk, ASE and SPM/XPM-aware FWM noise components respectively. Here we assume that noises added by the receiver are shot noise and thermal noise only, which normally are most significant noises generated by the receiver. After photodetection, a photocurrent is produced as,

$$i_p(t) = R_\lambda \langle E^2(t) \rangle + i_{sh}(t) + i_{th}(t) \quad (4-11)$$

where $i_{sh}(t)$ is the shot noise produced by incident lightwaves, $i_{th}(t)$ is the receiver thermal noise, and $R_\lambda = \frac{\eta e}{hf_i}$ is responsivity of the photodetector, where η is the quantum efficiency, e is the electronic charge (1.6×10^{-19} C), h is the plank's constant (6.626×10^{-34} J.s) and f_i is the frequency of the incident light. Hence, the square-and-average response of the photodetector to the incident lightwave $E(t)$ can be expressed,

$$R_\lambda \langle E^2(t) \rangle = i_s(t) + i_{sx}(t) + i_{sf}(t) + i_{sa}(t) + i_{xx}(t) + i_{aa}(t) + i_{ff}(t) + i_{xf}(t) + i_{xa}(t) + i_{fa}(t) \quad (4-12)$$

where $i_s(t)$ is the desired signal component while the remaining terms account for the beat noise components between signal and crosstalk $i_{sx}(t)$, signal and SPM/XPM-aware FWM $i_{sf}(t)$, signal and ASE $i_{sa}(t)$, crosstalk and itself $i_{xx}(t)$, SPM/XPM-aware FWM and itself $i_{ff}(t)$, ASE and itself $i_{aa}(t)$, crosstalk and FWM $i_{xf}(t)$, crosstalk and ASE $i_{xa}(t)$, and SPM/XPM-aware FWM and ASE $i_{fa}(t)$. Since the dominant beat noise terms are contributed by the signal-crosstalk, signal- SPM/XPM-aware FWM and the signal-ASE combinations, $n_\kappa(t)$ is used to indicate all the noise components as a combined process.

$$\begin{aligned} i_p(t) &= i_s(t) + [i_{sx}(t) + i_{sf}(t) + i_{sa}(t) + i_{sh}(t) + i_{th}(t)] \\ &= I_{s\kappa} + n_\kappa(t) \\ &= R_\lambda p_{sig}(N, f_i) b_\kappa(t) + n_\kappa(t) \end{aligned} \quad (4-13)$$

where the subscript κ indicates the data bit (1 or 0) being received. $I_{s\kappa}$ with $\kappa=1$ or 0 represents the corresponding signal components of the photocurrent. $p_{sig}(N, f_i)$ is the average power of the received signal at frequency f_i . $b_\kappa=2$ or 0 for $\kappa=1$ or 0, respectively, is because the power $p_{sig}(N, f_i)$ is the average power of the received signal

instead of the instantaneous value or the peak value of the signal power. The combined noise $n_{\kappa}(t)$ is modeled as a zero-mean Gaussian random process with a variance,

$$\sigma_{\kappa}^2 = \sigma_{sx\kappa}^2 + \sigma_{sa\kappa}^2 + \sigma_{sf\kappa}^2 + \sigma_{sh\kappa}^2 + \sigma_{th}^2 \quad (4-14)$$

Equations (4-15) ~ (4-19) show the variances of signal-crosstalk beat noise [124], signal-FWM beat noise [120], signal-ASE beat noise [123], shot noise and thermal noise, where ξ_{pol} ($=1/2$) is the polarization mismatch factor between the signal and crosstalk lightwaves [124], B_o and B_e are the optical and electrical bandwidths of the receiver, η_{th} is the spectral density of the thermal noise current in the receiver.

$$\sigma_{sx\kappa}^2 = 2\xi_{pol} R_{\lambda}^2 b_{\kappa} p_{sig}(N, f_i) p_{xt}(N, f_i) \quad (4-15)$$

$$\sigma_{sf\kappa}^2 = 2R_{\lambda}^2 b_{\kappa} p_{sig}(N, f_i) p_{fwm}(N, f_i) / 8 \quad (4-16)$$

$$\sigma_{sa\kappa}^2 = 4R_{\lambda}^2 b_{\kappa} p_{sig}(N, f_i) p_{ase}(N, f_i) B_e / B_o \quad (4-17)$$

$$\sigma_{sh\kappa}^2 = 2eR_{\lambda} (b_{\kappa} p_{sig}(N, f_i) + p_{xt}(N, f_i) + p_{fwm}(N, f_i) + p_{ase}(N, f_i)) B_e \quad (4-18)$$

$$\sigma_{th}^2 = \eta_{th} B_e \quad (4-19)$$

The receiver BER is evaluated and minimized with a given decision choice,

$$I_D = \frac{\sigma_0 I_{s1} + \sigma_1 I_{s0}}{\sigma_0 + \sigma_1} \text{ (refer to Appendix B), the Q parameter is [115],}$$

$$Q = \frac{I_{s1} - I_{s0}}{\sigma_0 + \sigma_1} = \frac{I_{s1}}{\sigma_0 + \sigma_1} \quad (4-20)$$

and the minimized BER is found to be (refer to Appendix B),

$$BER = \frac{1}{4} \left[\operatorname{erfc} \left(\frac{I_1 - I_D}{\sqrt{2}\sigma_1} \right) + \operatorname{erfc} \left(\frac{I_D}{\sqrt{2}\sigma_0} \right) \right] = \frac{1}{2} \operatorname{erfc} \left(\frac{Q}{\sqrt{2}\sigma_1} \right) \quad (4-21)$$

where, erfc is the complimentary error function [197].

4.5.3. Packet Error Calculation

In this part of analysis, two different cases should be considered. In the first case, the erroneous bit appears in the header, e.g., destination IP address. The packet will be dropped and the transmission will be considered as a failure. For the second case, the erroneous bit appears in the content and the packet with the error will be forwarded to higher layer to handle, which is out of our interested region. Hence, without losing the generality of this work we consider this case as a success.

Figure 4–5: Two parts of an IP packet

Thus if we consider a packet with length L_P bits consists of a header with length L_H bits and content with length L_C bits as shown in Figure 4–5, the average packet error probability (PER) [198] should be calculated using L_H instead of L_P as,

$$PER(L_P) = 1 - (1 - BER)^{L_H} \quad (4-22)$$

where BER denotes the bit error rate of the packet at the destination.

4.5.4. Effective Network Utilization Calculation

To show how the optical impairments affect the IP layer performance adversely, for the network model and its performance metric, we make use Lee's well-known traffic model for circuit-switched networks [199]. Although the model is simple, later researchers have drawn the conclusion that path length is the key design parameter for networks without wavelength converters. In order to keep the blocking probability small, path length must be kept small since it becomes less likely to find a free wavelength on all the hops of a path as the number of hops increases. This conclusion has been emphasized some authors in different ways, by theorems and examples in [200], by simulations in [201], and by the

alternate blocking probability models in [202]-[204]. Some researchers argued that Lee's model overestimates the gain in efficiency that wavelength converters provide [204]. However, since in this chapter we assume that wavelength converters are not present, we would like to make use of this model and its analytical result due to its simplicity.

Figure 4–6 shows the situation considered in Lee's model. Source client A attaches to source node S and destination client B attaches to destination node D. A and B set up a connection over a WDM mesh network and there are N hops from S to D. Each connection is assumed to occupy a full wavelength and there are F available wavelengths. Any other connection that travels along at least one of the N hops on any wavelength is called an interfering connection.

Figure 4–6: An N hop connection. Connection links are labeled. The other links are interfering links

Hence, the achievable utilization for a given blocking probability in networks without wavelength converters is [199],

$$TH = \left(1 - \left(1 - P_b^{1/F} \right)^{1/N} \right) \quad (4-23)$$

where P_b is the blocking probability, F is the number of wavelengths carried in one fiber and N is the number of hops along the route. Let TH_{eff} denote the effective achievable utilization for a given blocking probability in networks without wavelength converters, i.e. the utilization under the consideration of the effect of optical impairments on network layer performance, Hence, TH_{eff} is found to be,

$$TH_{eff} = TH (1 - PER) \quad (4-24)$$

4.6. RESULTS AND DISCUSSIONS

The following numerical results focus on the optical layer performance of the LSP (in terms of Q parameter) and the network layer performance of the LSP (in terms of BER and normalized throughput) of the traffic in different scenarios.

Certain assumptions need to be made before the analysis. (1) Gain-flattened EDFAs are used to recover the loss of signal strength. For example, a gain flattening filter, based on thin-film filter technology and metal bonding micro-optics packaging, can provide the inline compensation of the spectral gain profile of EDFAs [205]. This assumption makes the gain of EDFA same for all the wavelengths we are interested in, which greatly simplifies the modeling. (2) Non-blocking Ti:LiNbO₃ based waveguide is used at the OXC, as shown in Figure 4-3 and Figure 4-4. If different optical switch architecture is used at the OXC, the mathematics related to first-order crosstalk calculation may need to be modified to suit the specific architecture. However, the philosophy of 3-step performance analysis can still be applied. (3) Noises generated by the receiver are shot noise and thermal noise only, which are the most significant noises generated by the receiver. (4) Fiber length is not too long, so no inline EDFA is required. Since in our analysis, the longest fiber segment is 100km and the EDFAs at the nodes are sufficient in boosting the power to the level that can be detected by the photodetector correctly, no inline EDFA is required here.

Table 4-2 lists the common parameters and their associated values used in obtaining the numerical results, which are same throughout the entire section 4.6. Most of the parameters are chosen from Table I of [125]. Distance between each hop is 100km. The refractive index of fiber core is 1.46. The multiplexer and demultiplexer have 4dB loss each. The switch loss in dB is calculated by $2\log_2 ML_s + 4L_w$, where $L_s = 1dB$ is the switch element insertion loss, M is the number of input/output fiber pairs, and $L_w = 1dB$ is the waveguide/fiber coupling loss [95]. Moreover, the switch crosstalk ratio is 30dB.

Each tap has a loss of 1dB. The spontaneous emission factor for the EDFA is 1.5. The ratio of electrical bandwidth and optical bandwidth to channel rate are set to be 0.7 and 2.5 respectively. At the receiver, the root mean square (RMS) value of thermal current per square root of electrical bandwidth is $5.3 \times 10^{-12} \text{ Amp} / \sqrt{\text{Hz}}$ and the optical-electrical conversion efficiency is 0.9.

Table 4–2: Common Parameters Setting

Parameters	Value
Distance between nodes (D)	100km
Fiber core refractive index (n_c)	1.46
Multiplexer Loss (L_{mux})	4dB
Demultiplexer Loss (L_{demux})	4dB
Switch element insertion loss (L_s)	1dB
Waveguide/fiber coupling loss (L_w)	1dB
Switch loss (L_{SW})	$2 \log_2 ML_s + 4L_w$
Switch crosstalk ratio (χ_{sw})	30dB
Tap loss (L_{tap})	1dB
Spontaneous emission factor of EDFA (n_{sp})	1.5
Electrical bandwidth (B_e)	0.7* Channel rate
Optical bandwidth (B_o)	2.5 * Channel rate
RMS thermal current / $\sqrt{\text{bandwidth}}(\sqrt{\eta_{th}})$	$5.3 \times 10^{-12} \text{ Amp} / \sqrt{\text{Hz}}$
Optical – electrical conversion efficiency (η)	0.9

We consider two fiber types as fiber links in our numerical analysis. They are Dispersion Shifted Fiber (DSF) and Single Mode Fiber (SMF), which are very different in nature and could produce interesting results. Chapter 2 introduced the characteristics of two fiber systems, and made the qualitative comparison of the pros and cons of them. This section gives the detailed quantitative elaboration of the different systems with different

configurations based on our proposed three-step approach. The parameters pertaining to these fibers are as shown in Table 4–3. These parameters are of standard setting, most of which have been introduced in subsection 2.4.2.1. For the nonlinear index coefficient and third order susceptibility of fiber of the two types of fibers [114], more details can be found in subsection 4.6.1.1 and subsection 4.6.1.2 later on.

Table 4–3: Fiber Parameters

Parameter	Fiber Link Types	
	DSF	SMF
Operation Region	Around 1550nm	Around 1550nm
Fiber loss @ 1550nm (L_f)	0.2dB/km	0.25dB/km
Zero Dispersion Wavelength (λ_0)	1550nm	1310nm
Dispersion @ 1550nm (D_c)	0 ps/(nm.km)	20 ps/(nm.km)
Dispersion Slope @ 1550nm ($dD_c / d\lambda$)	0.07 ps/(nm ² .km)	–
Effective Area (A_{eff})	55 μm^2	80 μm^2
Nonlinear Index Coefficient (n_2)	2.5x10 ⁻²⁰ m ² /W	2.2x10 ⁻²⁰ m ² /W
Third order susceptibility of fiber (χ_{1111})	9.73x10 ⁻²⁰ m ² /W	8.57x10 ⁻²⁰ m ² /W

4.6.1. Optical Layer Performance

Consider the network topology below in Figure 4–7, if three label switched paths (LSPs) are set up and in operation and all the wavelengths carry traffic in the fiber. This can be considered as worst-case scenario, since all the wavelengths of all the fiber pairs are occupied. In this part, we focus on the optical layer performance and we take LSP1 as an example for numerical result. As shown, LSP1 starts from node 7 and follow the path 7 → 8 → 14 → 15 → 21 → 22 → 28 → 29 → 35 → 36. Node 36 is the destination node. In this case, the number of input fiber–output fiber pairs is M=2 for all the intermediate

nodes (i.e., node 8, 14, 15, 21, 22, 28, 29 and 35) of LSP1. Refer to the value of L_s and L_w in Table 4-2, the switch loss per node can be found to be $L_{SW} = 2 \log_2 ML_s + 4L_w = 6dB$.

Figure 4-7: Sample Optical Mesh Network

4.6.1.1. DSF Systems

We consider an all-optical network that utilizes DSF as fiber links. For the detailed setting, as shown in Table 4-3, DSF is operating near to zero-dispersion wavelength – 1550nm, where the fiber loss is 0.2 dB/km. Normally, the dispersion is neglected in the modeling or calculation, but since we want to model the network as close to the real one as possible, we will use 0.07 ps/(nm².km) as the dispersion slope at the zero dispersion wavelength to calculate the dispersion values of each individual wavelengths [23]. The

effective area of DSF is $55 \mu\text{m}^2$, the nonlinear index coefficient is $2.5 \times 10^{-20} \text{m}^2/\text{W}$ [114], and subsequently the third order susceptibility of DSF can be found to be $9.73 \times 10^{-20} \text{m}^2/\text{W}$

$$(\chi_{1111} = \frac{8n_c n_2}{3} \text{ [114]}).$$

As mentioned in section 4.6.1, the number of input fiber–output fiber pairs is $M = 2$ for all the intermediate nodes of LSP1, hence a modified version of Figure 4–2(a) specific to our numerical example (Figure 4–7) is shown in Figure 4–8, where the number of wavelengths carried in each fiber is 8. This is the case of system A configuration, which will be analyzed thoroughly later on. Similarly, for system B and C, where each fiber carries 16 wavelengths, the figure can be modified by simply adding eight more blocks of the 2x2 switches (refer to Figure 4–3(a)).

Figure 4–8: Modeling of Intermediate Nodes in DSF 8 Scenarios

Three 10Gbps DSF systems are investigated and their performances in terms of BER and Q parameter are compared and illustrated. Preamplifier at input port is set to compensate the loss at the fiber segment including the tap losses at the both ends of a fiber segment.

Postamplifier at the output port is set to compensate the loss of the components at the node, including the multiplexer, switch and demultiplexer.

Table 4–4 shows the configurations of these three systems: (A) 8–wavelength system, from 1547nm, channel spacing is 0.8nm; (B) 16–wavelength system, from 1546.6nm, channel spacing is 0.4nm; (C) 16–wavelength system, from 1543.8 nm, channel spacing is 0.8nm. The results are obtained by varying input power of the optical signal from 0.001 watt per channel to 0.006 watts per channel at a step of 0.001 watt. Furthermore, in each case, the change of Q with varying number of hops to destination is also shown. For example, when the number of hops is 4, the destination of LSP 1 is 21; when the number of hops is 8, the destination of LSP 1 is 35.

Table 4–4: Parameter setting for three configurations of DSF 10 Gbps system

Parameters	System A	System B	System C
Wavelengths Per Fiber (F)	8	16	16
Channel Rate (C)	10Gbps	10Gbps	10Gbps
Wavelength Spacing	0.8 nm	0.4 nm	0.8 nm
Pre amplifier gain (G_{in})	22dB	22dB	22dB
Post amplifier gain (G_{out})	14dB	14dB	14dB

Figure 4–9, Figure 4–10 and Figure 4–11 show in great details the Q parameters of system A, B, and C respectively. For each system, the Q parameters of four cases are shown separately, i.e., when the only impairment considered is ASE (case a), when only the linear impairments are considered – ASE and crosstalk (case b), when OXC is assumed to be perfect – ASE and FWM (case c), and when all the three major impairments of a DSF system are considered (case d). To find the practical Q for system design, only case d should be considered. We show the other three cases for the insight of how these impairments affect the optical performance in their unique ways. The results are shown for the middle wavelength at 1549.4nm for all the three systems, i.e. the 4th wavelength of system A, 8th wavelength of both system B and C.

For the results of system A, Figure 4–9(a) shows the case when only ASE noise is considered, where Q improves as the input power of the signal increases. E.g., when the number of hops to destination is 6, Q improves from 14.6 to 35.6 as input power increases from 0.001 to 0.006. Qualitatively, we understand that the stronger the signal is the less impact ASE noise brings to the system. By observing related equations (Equation (4–4), (4–17) and (4–20)), we find the nominator of Q is a function of signal power and the denominator is a function of square root of signal power. This explains the reason to the first finding quantitatively. Another finding is that the Q of the received signal keeps dropping as the number of hops to destination increases for all input powers. E.g., when input power is 0.001, Q drops from 14.6 to 12.65 as number of hops to destination increases from 6 to 8. This is because each hop consists of EDFA component which not only generates its own ASE noise but also amplify the accumulated ASE noise from its previous nodes. Figure 4–9 (b) illustrates the Q when ASE and crosstalk are considered. It shows that the trend and pattern of the curves are similar to those of Figure 4–9(a), because both ASE noise and the first-order homowavelength crosstalk are linear impairments. As expected, Q improves with the increase of input power and degrades with the increase of the number of hops to destination. E.g., when the hop count is 6, Q improves from 7.74 to 8.85 as input power increases from 0.001 to 0.006; when input power is 0.001, Q degrades from 7.74 to 6.7 as the hop count increases from 6 to 8. Figure 4–9 (c) shows the Q when ASE and FWM are captured. Different from Figure 4–9 (a) and Figure 4–9 (b), when hop count is small (up to 6 in this case) Q improves with the increase of input signal power, but as the hop count increases beyond certain value (7, in this case) Q decreases with the increase of input power. E.g., when the hop count is 6, Q improves from 14.46 to 34.8 as input power increases from 0.001 to 0.006. When the hop count is 8, Q drops from 7.15 to 1.4 as input power increases from 0.001 to 0.006. This is due to the nonlinearity of the nature of FWM effect. By observing the equation (4–7), we find the FWM noise term consists of a product of powers of three relevant input signals that increases dramatically as input power increases. It also contains a cube of effective distance term, which increases very fast as distance to the destination increases. When

the hop count is small, the dominant noise is due to ASE and FWM effect is relatively small due to its small coefficient, hence the Q is very similar to the portion in Figure 4–9(a). But as the hop count, corresponding distance to destination, both the cube of distance term and the cube of the signal power term deprive ASE of the dominance. Finally, Figure 4–9(d) shows the case where all the three major impairments are considered in a typical DSF system. We can see that up to 6 hops, the curves behave very much similar to Figure 4–9 (b) where linear impairment are apparent, but when hop count increases beyond 7, the curves behave similar to Figure 4–9 (c) as the nonlinear FWM effect begins to dominate the performance.

Figure 4–9: Q Parameter of System A at 1549.4nm (4th)

Figure 4–10 shows the Q parameters under the four cases for system B, where 16 wavelengths are carried in a fiber, ranging from 1546.6nm and channel spacing is 0.4nm. Compared with Figure 4–9, Figure 4–10 shows the same results in case (a) and case (b), where only linear impairments are considered. This is because both ASE and first-order homowavelength crosstalk is independent of the number of wavelengths a fiber carries and the channel spacing. Differences appear in case (c) and (d), where nonlinear impairments come into the picture. Generally the curves are similar to that of Figure 4–9, however, the dominance of nonlinearity in system B appears earlier than system A, also more severe. E.g., when input power is 0.006, the Q of system B begins to drop dramatically at the 6th hop but the Q of system A only begins to drop apparently when hop count is 7. Also when hop count is 6 and input power is 0.006, the Q of system B is 31 that is lower than the Q of system A – 34.75. This is because system B carries doubled the number of wavelengths as system A does and the channel spacing is only half of that of system A, hence FWM effect becomes more severe in system B.

Figure 4–10: Q of System B at 1549.4nm (8th)

Figure 4–11 shows the Q parameters under the four cases for system C, where 16 wavelengths are carried in a fiber, ranging from 1543.8 nm and channel spacing is 0.8nm. It can be seen that the performance of the Q parameter of system C is very much alike that of system A. This is because although the number of wavelengths is doubled, but the wavelength spacing is kept same – 0.8nm. Hence, in our scenario, the impact of the wavelength spacing is more than the impact of the number of wavelengths.

Figure 4–11: Q of System C at 1549.4nm (8th)

Figure 4–12, Figure 4–13 and Figure 4–14 illustrate the BER performance of the three systems. Same as Figure 4–9 – Figure 4–11, for each system, the BER of four cases are shown separately, i.e., when the only impairment considered is ASE (case a), when only

the linear impairments are considered – ASE and crosstalk (case b), when OXC is assume to be perfect – ASE and FWM (case c), and when all the three major impairments of a DSF system are considered (case d). The results are shown for the middle wavelength at 1549.4nm for all the three systems, i.e. the 4th wavelength of system A, 8th wavelength of both system B and C. BER is another view of the optical layer performance, it also a parameter that network layer concerns. Here, we just want to highlight that the BER perform consistently corresponding to its relative Q parameter of the system. With all the major impairments of the DSF systems, the BER of system A (Figure 4–12(d)) is very much alike that of system C (Figure 4–14(d)), both of which are better than the BER of system B (Figure 4–13 (d)).

Figure 4–12: BER of System A at 1549.4nm (4th)

Figure 4-13: BER of System B at 1549.4nm (8th)

Figure 4-14: BER of System C at 1549.4nm (8th)

As a summary to Figure 4-9 – Figure 4-11 and Figure 4-12 – Figure 4-14, Figure 4-15 and Figure 4-16 illustrate the Q parameters and BER, respectively, of the three systems when all the three major optical impairments (i.e. ASE, FWM and first-order homowavelength crosstalk) are captured.

As expected, Q decreases as the number of hops increases for more use of EDFAs, OXCs and fiber segments. For system A and C, the Q parameters are still higher than 7 when the hop count is 7. It corresponds to BER around 10^{-12} , which is the nowadays rule of thumb acceptable BER of optical transmission system. But for system B, when hop count is 7, the Q of input power 0.003 ($Q = 6.7$, $BER = 10^{-11}$), 0.004 ($Q = 6.1$, $BER = 5 \times 10^{-10}$), 0.005 ($Q = 5.5$, $BER = 2 \times 10^{-8}$) and 0.006 ($Q = 4.95$, $BER = 4 \times 10^{-7}$) have all dropped below 7. Our results consistently demonstrate that as the number of hops from source to destination increases, Q parameter keeps degrading. When the number of hops is small the dominant impacts are from ASE and crosstalk, but beyond certain value of hops (7, in this case), FWM effect begins to show its dominancy.

Figure 4-15: Comparison of Q parameters for System A, B, and C

Figure 4-16: Comparison of BER for System A, B and C

4.6.1.2. SMF + DCF 40 Gbps Systems

We consider an all-optical network that utilizes SMF as fiber link and DCF to manage the dispersion. Since the number of input fiber-output fiber pairs is $M = 2$ for all the intermediate nodes of LSP1, hence a modified version of Figure 4-2(b) to our numerical example (Figure 4-7) is shown in Figure 4-17.

For the detailed setting, as shown in Table 4–3, the SMF+DCF system is still operating near to 1550nm region, where the fiber loss is 0.25dB/km. But unlike DSF, SMF's zero dispersion wavelength is at 1310nm and it has a dispersion value of 20 ps/(nm.km) value around the 1550nm region. Compared with DSF, it has a large core effective area of $80 \mu\text{m}^2$, the nonlinear index coefficient is $2.2 \times 10^{-20} \text{m}^2/\text{W}$, and subsequently the third order susceptibility of SMF can be found to be $8.57 \times 10^{-20} \text{m}^2/\text{W}$ [114]. As we have mentioned in previously, the DCF has to be used to compensate the dispersion at SMF for a transmission link, which has a loss of -0.4dB/km . It has a dispersion of $-100 \text{ps}/(\text{nm.km})$, hence the length of DCF to the length of SMF is 1:5 to compensate fully the dispersion of SMF. The core area of DCF is $25.8 \mu\text{m}^2$ [89]. Preamplifier $G_{in1} = 27 \text{dB}$ is set to compensate the loss at the SMF including the tap losses at the both ends of a SMF segment. Preamplifier $G_{in2} = 10 \text{dB}$ is used to compensate the loss at the DCF including the tap losses at both ends of DCF. Postamplifier $G_{out} = 14 \text{dB}$ is set to compensate the loss of the components at the node, including the multiplexer, switch and demultiplexer.

Figure 4–17: Modeling of Intermediate Nodes in SMF+DCF Scenarios

Two 40Gbps SMF+DCF systems are investigated and their performances in terms of BER and Q parameter are compared and illustrated. Table 4–5 shows the configurations of these three systems: (D) 16–wavelength system, from 1543.8nm, channel spacing is 0.8nm; (E) 16–wavelength system, from 1546.6nm, channel spacing is 0.4nm. Similar to section 4.6.1.1, the results are obtained by varying input power of the optical signal from 0.001 watt to 0.006 watts at a step of 0.001 watt. Furthermore, in each case, the change of Q with varying number of hops to destination is also shown.

Table 4–5: Parameter setting for two configurations of SMF+DCF 40 Gbps systems

Parameters	System D	System E
Wavelengths Per Fiber (F)	16	16
Channel Rate (C)	40Gbps	40Gbps
Wavelength Spacing	0.8 nm	0.4 nm

Figure 4–18 and Figure 4–19 analyze in great details the Q parameters of system D and E respectively. For each system, the Q parameters of four cases are shown separately, i.e., when the only impairment considered is ASE (case a), when only the linear impairments are considered – ASE and crosstalk (case b), when OXC is assume to be perfect – ASE and Fiber Nonlinearity (case c), and when all the three major impairments of a SMF+DCF system are considered (case d). To find the practical Q for system design, only case d should be considered. We show the other three cases for the insight of how these impairments affect the optical performance in their unique ways. The results are shown for the middle wavelength at 1549.4nm for all the three systems, i.e. the 8th wavelength of both system D and E.

For the results of system D, Figure 4–18(a) shows the case when only ASE noise is considered, where Q improves as the input power of the signal increases. E.g., when the number of hops to destination is 6, Q improves from 3.95 to 9.7 as input power increases from 0.001 to 0.006. This is because the stronger the signal is the less impact ASE noise brings to the system. Another finding is that the Q of the received signal keeps dropping

as the number of hops to destination increases for all input powers. E.g., when input power is 0.006, Q drops from 9.7 to 9.05 as number of hops to destination increases from 7 to 8. This is because each hop consists of EDFA component which not only generates its own ASE noise but also amplify the accumulated ASE noise from its previous nodes. Figure 4–18 (b) illustrates the Q when ASE and the first-order homowavelength crosstalk are considered. It shows that the trend and pattern of the curves are similar to those of Figure 4–18 (a), because both ASE noise and the crosstalk are linear impairments. As expected, Q improves with the increase of input power and degrades with the increase of the number of hops to destination. E.g., when the hop count is 6, Q improves from 3.85 to 6.9 as input power increases from 0.001 to 0.006; when input power is 0.006, Q degrades from 6.9 to 6.4 as the hop count increases from 6 to 7. Figure 4–18 (c) shows the Q when ASE and fiber nonlinearity are captured. Different from Figure 4–18 (a) and Figure 4–18 (b), when hop count is small (up to 7 in this case) Q improves with the increase of input signal power, but as the hop count increases beyond certain value (8, in this case) Q decreases with the increase of input power. E.g., when the hop count is 7, Q improves from 4 to 9.7 as input power increases from 0.001 to 0.006. When the hop count is 9, Q drops from 2.85 to 0.85 as input power increases from 0.001 to 0.006. This is due to the nonlinear nature of SPM, XPM and FWM effect that increases dramatically as the increase of input power and effective distance, as shown in equation (4–7). When the hop count is small, the dominant noise is due to ASE and FWM effect is relatively small due to its small coefficient, hence the Q is very similar to the portion in Figure 4–18 (a). But as the hop count, corresponding distance to destination, both the cube of distance term and the cube of the signal power term deprive ASE of the dominancy. Finally, Figure 4–18 (d) shows the case where all the three major impairments are considered in a typical SMF+DCF system. We can see that up to 7 hops, the curves behave very much similar to Figure 4–18 (b) where linear impairment are apparent, but when hop count increases beyond 8, the curves behave similar to Figure 4–18 (c) as the fiber nonlinearity begins to dominate the performance.

Figure 4-18: Q of System D: SMF+DCF 40Gbps 0.8nm Spacing

Figure 4-19 shows the Q parameters under the four cases for system E., where 16 wavelengths are carried in a fiber, ranging from 1546.6nm and channel spacing is 0.4nm. Compared with Figure 4-18, Figure 4-19 shows the same results in case (a) and case (b), where only linear impairments are considered. This is because both ASE and first-order homowavelength crosstalk is independent of the number of wavelengths a fiber carries and the channel spacing. Differences appear in case (c) and (d), where nonlinear impairments come into the picture. Generally the curves are similar to that of Figure 4-18, however, the dominance of nonlinearity in system E appears earlier than system D, also more severe. E.g., refer to case (d) of both figures, when input power is 0.006, the Q of

system E begins to drop dramatically beyond the 7th hop but the Q of system A only begins to drop quickly when hop count is beyond 8. Also when hop count is 8 and input power is 0.006, the Q of system E is 3.15 that is lower than the Q of system D – 5.52. This is because system the channel spacing of system E is only half of that of system D, hence the SPM and XPM-aware FWM nonlinear effect becomes more severe in system E.

Figure 4–19: Q of System E: SMF+DCF 40Gbps 0.4nm Spacing

Figure 4–20 and Figure 4–21 illustrate the BER performance of system D and E, which is not only another view of the optical layer performance but also a parameter that network layer concerns. Same as Figure 4–18 and Figure 4–19, for each system, the BER of four

cases are shown separately, i.e., when the only impairment considered is ASE (case a), when only the linear impairments are considered – ASE and crosstalk (case b), when OXC is assume to be perfect – ASE and Nonlinearity (case c), and when all the three major impairments of a SMF+DCF system are considered (case d). The results are shown for the middle wavelength at 1549.4nm, i.e. 8th wavelength, for both system D and E.

Figure 4-20: BER of System D: SMF+DCF 40Gbps 0.8nm Spacing

Figure 4-21: BER of System E: SMF+DCF 40Gbps 0.4nm Spacing

Our results consistently demonstrate that as the number of hops from source to destination increases, Q parameter and BER keeps degrading. When the number of hops is small the dominant impacts are from ASE and crosstalk, but beyond certain value of hops (8 or 9, in this case), fiber nonlinear effect, in terms of the SPM-XPM-aware FWM, begins to show its dominance.

4.6.1.3. Comparison of Two 40Gbps Systems

In this part, we want to compare two 40Gbps systems in order to give a better picture of how 40Gbps systems are different from 10Gbps systems and how these configurations

impact apparently on the Q and BER performance of the optical network. The first system is a SMF+DCF system with channel rate 40Gbps, which is identical to the system E as investigated in 4.6.1.2. The second system, System F, is a DSF system with channel rate 40Gbps so the analysis of nonlinearity will take into account the XPM and SPM. Table 4–6 lists the values and parameters of the two systems.

Figure 4–22 and Figure 4–23 compare the Q parameters and BER performances of these two systems respectively. Here, we consider the case where all the major impairments are captured for it is meaningful to find the practical Q and BER for system design. The results are shown for the middle wavelength at 1549.4nm for both systems, i.e. the 4th wavelength. It is interesting to observe that before turning point that nonlinearity gains dominancy, system F is actually better in performance. E.g., when hop count is 5, input power is 0.003 watt, the Q of system E is 6.29 which is lower than the Q of system F – 8.11. Correspondingly, the BER of system E is 1.57×10^{-10} which is higher than that of system F – 2.55×10^{-16} . This is because less loss and absence of DCF making less use of EDFA thus generating less ASE noise. However, the turning point of system F (between 6 to 7) appears earlier than that of system E (between system 7 and 8), this is because the small dispersion makes fiber nonlinearity more severe in DSF system than a SMF+ DCF system. Thus actual implementation of the systems depends on the specific requirements of the IP over WDM all-optical network.

Table 4–6: Parameter setting for two 40Gbps Systems

Parameters	System E	System F
Fiber type	SMF+DCF	DSF
Wavelength Per Fiber	16	16
Channel Rate (C)	40Gbps	40Gbps
Wavelength Spacing	0.4 nm	0.4 nm
Pre amplifier gain	$G_{in1} = 27\text{dB}$	$G_{in} = 22\text{dB}$
	$G_{in2} = 10\text{ dB}$	
Post amplifier gain (G_{out})	14dB	14dB

Figure 4-22: Q Comparison of System E and F

Figure 4-23: BER Comparison of System E and F

4.6.2. Effect of Network LSP Setup on Optical Layer Performance

In this chapter, we will focus on how the cross-layer effect happens, i.e. how network layer decision (setting of LSPs or route selection) affects optical performance and how optical impairments affect the effective network performance. We consider the same setup and parameter setting as system A in section 4.6.1.1, i.e. DSF system, 8 wavelengths per fiber with 0.8nm channel spacing. Recall the network topology shown in Figure 4-24, we will look into three scenarios: (a) only LSP 1 is set up, the result will be equivalent to the case where there is no first-order homowavelength crosstalk; (b) LSP 1 and LSP 2 are set up,

in this case, different node will have different J_k (refer to Equation (4-3)); (c) LSP 1,2 and 3 are set up, this is the same as case(d) of system A in section 4.6.1.1.

Figure 4-25 and Figure 4-26 illustrate the Q parameters and BER of the three scenarios respectively. Similar to section 4.6.1, the results are obtained by varying input power of the optical signal from 0.001 watt to 0.006 watts at a step of 0.001 watt. Furthermore, in each case, the change of Q and BER with varying number of hops to destination is also shown.

Figure 4-24: Network Topology

As shown in Figure 4-25, scenario (a) has better Q than scenario (b), which in turn has better Q than scenario (c). E.g., when input power is 0.001 and hop count is 6, the Q of scenario (a) is 14.5, that is higher than the $Q = 9.65$ for scenario (b) and $Q = 7.75$ for

scenario (c). Correspondingly, as shown in Figure 4–26, under the same condition, the BER of scenario (a) is 10^{-48} , which is lower than the $\text{BER} = 2 \times 10^{-48}$ for scenario (b) and $\text{BER} = 5 \times 10^{-15}$ for scenario (c). Furthermore, optical signal can survive longer in scenario (a) than in scenario (b) and scenario (c). E.g., if we assume that $Q=6$ ($\text{BER} = 10^{-9}$) is an acceptable threshold for optical signal, when the hop count is 8, for input power 0.001, $Q=7.15$ ($\text{BER} = 5 \times 10^{-13}$) for scenario (a) is allowed, $Q = 6$ ($\text{BER} = 10^{-9}$) for scenario (b) is just allowed and $Q = 5.3$ ($\text{BER} = 6 \times 10^{-8}$) is not allowed. Thus scenario (a) and (b) survive up to 8th hop, but scenario (c) is only allowed up to 7th hop.

Hence, the network setup has impact on the optical layer performance in terms of Q and BER indeed. If we consider the topology in Figure 4–24, for LSP1, node 7 cannot successfully transmit to node 35 according to current LSP arrangement. Thus, only LSP 2 and LSP 3 can be allowed but LSP 1 is not allowed. But since per-hop distance are all 100km, if we change the LSP 3 path from source 13 $\rightarrow$ node 14 $\rightarrow$ node 20 $\rightarrow$ node 21 $\rightarrow$ node 27 $\rightarrow$ node 28 $\rightarrow$ node 34 $\rightarrow$ node 35 $\rightarrow$ destination 41 to source 13 $\rightarrow$ node 19 $\rightarrow$ node 25 $\rightarrow$ node 31 $\rightarrow$ node 37 $\rightarrow$ node 38 $\rightarrow$ node 39 $\rightarrow$ node 40 $\rightarrow$ destination 41, all the three LSPs are allowed. Hence by understanding the cross-layer effect of the optical performance and network layer route setup, we can do better in system design.

Figure 4–25: Compare LSP Q

Figure 4-26: Compare LSP BER

4.6.3. Adverse Effect of Optical Impairments on Network Performance

To investigate the adverse effect of optical characteristics on the network layer performance, we use the same blocking probability, 0.001, as it was used in [204]. Figure 4-27 shows the effective normalized throughput TH_{eff} for three different system configurations with different packet lengths and input powers. The three systems are exactly the same as those used in section 4.6.1.1, where (A) $F=8$, channel spacing 0.8nm; (B) $F=16$, channel spacing 0.4nm; (C) $F=16$, channel spacing 0.8nm. For all the systems, TH_{eff} after considering the optical impairments is almost the same as the one without considering it when the number of hops to destination is small. However, as the number of hops increases beyond certain value, optical impairments significantly degrade the network performance by means of a drop of TH_{eff} . Also, for small header packet, the situation is less severe than that for large header packet for the same amount of the input signal power. On the other hand, for the same length of packet header, the less the input power, the less degrade the performance does. For the performance comparison, without optical impairments, B and C perform same. Their throughputs are higher than A for more wavelength available for use. However, by considering optical impairments, A

performs best for least degrading when hop count is 8 for 25-byte header, while C is the next and B performs poorest. While A and C perform “perfectly” for 15-byte header when hop count is 8, B begins to degrade apparently.

Let us look closely the point, where the hop count is 8 and the input power is 0.001w. For system A, although $Q = 5.3 < 7$ (Figure 4–15(a)) and $BER = 6 \times 10^{-8} > 10^{-12}$ (Figure 4–16(a)), but for both types of packets, IP packets perform exactly the same as without considering optical impairments. For system B, $Q = 3.25 < 7$ (Figure 4–15(b)) and $BER = 6 \times 10^{-4} > 10^{-12}$ (Figure 4–16(b)), and for both types of packets, IP packets perform not as good as without considering optical impairments. Moreover, the one with smaller IP packet header performs better for it has lower PER with same BER and Q . For system C, although $Q = 5.17 < 7$ (Figure 4–15(c)) and $BER = 1.2 \times 10^{-7} > 10^{-12}$ (Figure 4–16(c)), but for both types of packets, IP packets perform exactly the same as without considering optical impairments.

Figure 4–27: Effective normalized throughput for system A, B, and C

As a summary, we find out if only use $Q=7$ and $BER = 10^{-12}$ as a threshold for allowing setting LSP, the coverage of the network and the throughput of the network will be sacrificed. Hence, we should not use this kind of fix criteria for setting up LSP. Instead we should use dynamic criteria for setting LSP to achieve better utilization of the entire network.

4.7. GUIDELINES TO NEXT GENERATION NETWORK DESIGN

The accelerating growth of data traffic is motivating the research for more efficient, flexible, intelligent optical network architectures. In this direction IP over WDM is becoming accepted as one of the most promising candidates. Currently deployed optical networks are still based on optical connections terminated at each network node by optoelectronic transponders. Because of their high cost, fixed bit rate, and fixed protocol format, optoelectronic transponders limit the network evolution. On the other hand, researchers believe that NGNs should be all-optical and fully transparent in order to get rid of all the electronic bottlenecks. However, due to the expensive components of all – optical OPS network and technological difficulties involved in implementing commercially viable OPS nodes, current focus of all-optical IP over WDM network is still at the wavelength level.

4.7.1. Ultimate Goal: All-Optical Transparent Network

The ultimate ideal structure of an optical network is a transparent network, where optical signals propagate end-to-end without any electrical-domain process at intermediate nodes. Optical transport is evolving from traditional opaque networks [78]-[80] that use all-electronic switching techniques toward all-optical transparent [79] [81] networks. This transparency should provide protocol and bit rate independence, reduced cost for equipments and efficient network provisioning of the resources.

A significant factor that militates against the introduction of all-optical networking is the fact that most optical technologies are analog rather than digital. Many factors may degrade optical signal quality in long-reach optical transmission systems to make the data unrecognizable at the receiver. Contributing factors include optical noise, chromatic dispersion, nonlinear effects, and crosstalk. Hence, the reality shows that the BER

depends on the signal's path through the optical network and there is a maximum transparent reach limit for signals within optical transmission systems [206].

To go beyond these transparent reach limits, signal regeneration is necessary to re-amplify, reshape, and retune (3R) the optical signals. Regeneration is helpful to improve signal quality such that a lightpath can travel farther before it reaches its destination or needs to be regenerated again. Regeneration can, in principle, be accomplished purely in the optical domain. The most suitable candidates for all-optical 3R, that have been reported, are based on either quantum-dot (QD) semiconductor optical amplifiers (SOAs) or cross-phase modulation (XPM) technology in highly nonlinear fibers. But neither of them is able to do multi-wavelength operation. [207]. However, regeneration in the electronic domain, which converts an optical signal into electronic format and then uses the electronic signal to modulate an optical laser, is currently the most reliable technique, which also reflects the current commercial status of regeneration technology. But it is a costly solution particularly for systems supporting a large number of channels at high data rates. Furthermore, for high-speed data transmission, the electronic circuit will become the bottle neck for the network to achieve the ultimate bandwidth it can. Such architecture prevents an OEO regenerator from performing with the same speed as an all-optical regenerator, and is not transparent to the network protocols used.

4.7.2. Intermediate Step: Translucent Network

Without all-optical 3R, fully transparent network is hardly achieved with satisfactory performance and an alternative that fits the current need emerges – translucent optical network [76]-[77] [79] [82]-[83]. Translucent optical networks seek a graceful balance between network design cost and service provisioning performance, which can achieve performance comparable to that of an all-electronic switching network, but requiring far fewer signal regenerators [208].

Translucent optical networks can be broadly classified into three categories: (1) translucent networks that are made up of transparent islands [208]–[209], where these islands are in fact the domains of optical transparency. All nodes within an island are transparent and have no regeneration capability. Regeneration OXC nodes are located only on the boundary of the island. Therefore, within each island, lightpaths sourced from a node can transparently reach any other node without any signal regeneration; whereas between different islands, it is necessary for the lightpaths to pass island boundaries and traverse regeneration OXC nodes on the boundaries. Figure 4–28 illustrates a translucent network made up of three transparent islands, in which nodes 3, 7, 8, 9 and 10 are boundary regeneration nodes to relay lightpaths crossing neighboring transparent islands. (2) Translucent networks with sparsely placed opaque nodes [75] [210], where regeneration capability is distributed strategically around the entire network. In this case, opaque nodes are located among the transparent nodes to regenerate optical signals and convert wavelengths electronically, while transparent ones have the optical switching function only. (3) Translucent networks made up of translucent nodes [208], where each of the nodes contains an optical switching core and electronic switching core, including a bank of regenerating OEO transponders.

Figure 4–28: Example of a translucent network with three transparent islands.

4.7.3. Our choice of NGN architecture

To achieve a good balance of complexity and cost, we consider only the first category in our work since it has lower complexity, in terms of the network management, planning and operation, than the second category and lower cost than the third category for its smaller amount of OEO transponders required. Furthermore, by replacing those boundary OEO regeneration modules with all-optical 3R modules in the future, when the all-optical 3R techniques are mature and able to support WDM operation, the translucent network can seamlessly transform into the ultimate NGN architecture with full transparency.

After we decide that we should adopt the translucent network as the highest level of network architecture, which consists of multiple “islands of transparency”, the problem now becomes two separate topics. The first one is how to do the network planning – how to define the boundary of islands. The second one is the routing and wavelength assignment (RWA) and signaling within each of the island of transparency.

4.7.3.1. *Define Boundary of Island – Larger Islands of Transparency*

So far, all the upper layer operations of IP over WDM network (e.g., the setup of GMPLS label switched path (LSP) from source to destination) assume an error-free physical layer, hence to practically achieve adequate signal quality for every possible route, the islands of transparency have been statically designed during the planning process considering every worst – case situation, e.g., every lightpath between the farthest nodes which traverses the most impaired links. This has determined extremely small islands of transparency and the presence of numerous expensive electronic transponders [211].

The inspiration of defining larger islands of transparency is that the worst-case is rare and non-frequent compared with the normal operation scenarios. Therefore, larger

islands of transparency can be designed, inside which only a few lightpaths passing through the worst – case paths will not satisfy the performance requirement (in terms of BER or Q). Furthermore, the connection between any two stations should be dealt with even under the worst – case scenario, so certain re–routing capability should be included in the operation of the transparent domains. Also, any connection that travels through few domains should also be dealt with. So the question is really how big the size of the island should be in order to achieve a good balance of cost and performance.

In fact, the observation from section 4.6 helps us in deciding the islands of transparency. Recall the study of Q parameters for system C of section 4.6.1.1 as shown in Figure 4–29, where 16 wavelengths are carried in a fiber, ranging from 1543.8 nm and channel spacing is 0.8nm. The distance between neighboring hops is fixed. Figure 4–29(a) shows the recalled result of Figure 4–11(b). This figure illustrates the Q when the major linear impairments, i.e., the ASE and crosstalk are considered. It shows that Q improves as the input power of the signal increases and degrades with the increase of the number of hops to destination. More importantly, when only linear impairments are considered, the drop of the Q is gradual as the hop counts increase for all the values of the signal powers. On the other hand, Figure 4–29 (b) shows the recalled result of Figure 4–11 (d). This figure shows the case where both the major linear impairments (i.e. ASE and first–order homowavelength crosstalk, in this case) and the major nonlinear impairments (i.e. FWM, in this case) are considered in a typical DSF transmission system. It is observed that up to certain number of hops (i.e., 6 hops in this case) the curves behave very much similar to the case where only linear impairments are considered. But when hop count increases beyond certain threshold (i.e., 7 hops), Q values for all the input powers drop abruptly and the performance of higher power becomes worse than the performance of lower power.

Figure 4–29: Recall results in Figure 4–11

So our finding is that when the hop count is small, the dominant noise is due to linear impairments and nonlinear effect is relatively small due to its small coefficient (refer to equation (4–7)). But as the hop count increases, corresponding the increase of distance to destination, both the cube of distance term and the cube of the signal power term in equation (4–7) deprive linear impairments of the dominancy. Or we can say that when the number of hops goes beyond certain threshold, the nonlinearity of the system begins to dominate the performance, which corresponding to the fast fading of Q parameter (or fast rising of BER). Similar trends can also be found in all the other systems investigated in the section 4.6.

Here we do some further investigation on the effect of number of wavelengths used to the Q parameter of system C. The system setting is same as the one shown in section 4.6.1.1, but with different number of wavelengths that are on. The measured wavelength is the central wavelength at 1549.6 nm. Here we consider the worst – case linear impairments of the system. As shown in Figure 4–30, the Q parameter of the measured wavelength degrades as the number of wavelengths that is on increases from 1 to 16. When only one wavelength is on, the performance is same as the one without nonlinear impairments

considered. When four wavelengths are randomly chosen to be on, the Q starts to drop rapidly when number of hops is around 9 ~ 10. When we further increase the number of on wavelengths to 8, Q starts to degrade rapidly when number of hops is around 7~8. When the worst case is considered, where all of the 16 wavelengths are on, the nonlinearity starts to gain dominance when the number of hops is around 6~7.

Figure 4-30: Q performance of System C with different numbers of wavelengths on

Hence, our findings so far can be used to propose a novel method to define island of transparency, i.e. we will use the worst – case linear impairments to define the boundary of transparency. For example, for system C, we can use the rule of thumb number ($Q = 7$, or equivalently $BER = 10^{-12}$) as the criteria. In this case, number of hops can be found to be around 10. So we can use this to define the maximum diameter of the transparent domain.

After we define the larger island of transparency, we can assign all the edge nodes at the boundary to be the translucent nodes, where OEO operation will be taken, as shown in Figure 4–31(a). In the topology, we assume the nodes are equally distributed inside the region. In this case, if d is the diameter (in terms of number of hops), then the number of the OEO boundary nodes is $N_1 = \pi d$ (for system C, $d = 10$), which is proportional to the square root of the number of nodes inside the network. This will not scale in terms of cost saving from eliminating OEO nodes.

We can further save the usage of OEO nodes at the boundary as shown in Figure 4–31(b). If we define d' to be the worst – case diameter when both the linear and nonlinear impairments are considered, we can obtain $d' = 7$. We can assign the OEO nodes at the physical boundary of the islands of transparency with neighboring OEO nodes being separated by d' . Hence we only need to assign $\frac{\pi d}{d'}$ number of OEO nodes at the boundary.

Thus the usage of OEO nodes is decided upon the difference between d and d' , which increases very slowly if not a constant. In this case, if a lightpath from source (S) to destination (D) is found to be not with the satisfactory physical layer quality (as shown by the dotted line with arrow), node S will resort to its nearest boundary node with OEO capability. Since the OEO nodes are separated by the worst-case design distance from each other, in the worst case, they can pass the traffic one to the other until reach the nearest OEO node to node D. That specific OEO node will in turn forward the traffic to the destination node D (as shown in the solid line with arrows).

Figure 4-31: Assign OEO nodes at boundary of island of transparency

4.7.3.2. RWA and Signaling within “Larger Island of Transparency”

After having this new definition of “larger island of transparency”, linear impairments are no longer needed to be evaluated when doing the impairment-aware routing or signaling within the transparent domain. In this way, it can make any impairment-aware signaling or routing more scalable than consider both the linear and nonlinear impairments.

After we build the islands of transparency, we want to look at the routing and wavelength assignment (RWA) and signaling within each of the island of transparency, to which GMPLS is the best candidate [192]. The GMPLS protocol suite implements a distributed control plane which permits the introduction, mainly in the network backbone, of different types of all-optical Network Elements (NE). This determines a reduction of the amount of expensive OEO nodes and the introduction of islands of transparency.

Although GMPLS has been introduced to dynamically set up all-optical end-to-end connections, i.e. lightpaths, within islands of transparency, the end-to-end lightpath provisioning over islands of transparency based on GMPLS protocols assumes that every possible route eligible by the routing protocol is characterized by the adequate signal

quality. Indeed no physical information is included in current version of the GMPLS protocol suite.

Extensions to the GMPLS control plane that provide the dynamic evaluation of lightpath physical impairment would potentially allow to fully exploiting the advantages of the islands of transparency. Indeed larger - size islands of transparency, in which just a major subset of lightpaths complies with physical parameters constraints, could be designed. In this scenario the compliance of the lightpath to be set up with the physical constraints would be guaranteed by the control protocol [212].

We have two choices of the approaches to do impairment-aware GMPLS control in the objective translucent network.

Figure 4-32 (a) shows one choice of the proposed approach, where we first define island of transparency to be large enough using the worst - case linear impairments. By doing that, we can ensure that in most of the operation scenarios, the lightpaths can be set up with satisfactory performance. Then inside the island, we use normal routing algorithm (e.g., OSPF - TE) with various wavelength assign algorithms (e.g., First Fit (FF), Random Fit (RF), Best Fit (BF)) to do the routing. This is because we know that for most of the cases, the route selected will satisfy the system Q requirement. Finally, we use impairment-aware signaling with nonlinear impairments being the only consideration, when doing the GMPLS signaling. If the signal quality is evaluated using measurement instead of calculation, then the measurement will take into account both the linear and nonlinear impairments incurred. In this way, we can avoid adding more information on routing protocol since OSPF - TE itself already suffers from scalability and stability issues in large-scale network [213]-[214]. By adding physical impairments information to the signaling protocol of GMPLS through modifying RSVP-TE, we can ensure the LSP finally will be set up upon required quality.

Figure 4–32 (b) shows another choice of the proposed approach, where we first define the larger islands of transparency. Then inside the island, we use nonlinear impairment-aware RWA algorithm (e.g., nonlinear impairment-aware OSPF – TE (routing) + wavelength assignment or OSPF – TE (routing) + nonlinear impairment-aware wavelength assignment algorithm). This will make sure every LSP found will satisfy the signal quality of the physical layer requirement. Finally, we use normal GMPLS signaling (i.e., RSVP – TE or LDP – TE) to set up the LSP if required resources are available. In this way, we can avoid the long delay if an LSP is rejected after the impairment-aware signaling find out that the LSP does not satisfy the required quality of link in choice (a).

Figure 4–32: Two choices of the proposed approach

By properly designing the islands of transparency, we can ensure that the probability of having LSP rejected after final signaling due to optical impairments is small. Therefore, choice (a) would be a better one in order to achieve scalability, stability and simplicity of the network. Also the computing resources spent for the impairment-aware signaling will be much smaller than the impairment-aware RWA, thus lower cost can be achieved by implementing choice (a).

The third step of choice (a) is the distributed signaling approach, which include the physical link information (in this case all components related to FWM) in the current GMPLS signaling protocols. Thus, in contrast to the distributed routing approach as in step 2 of choice (b), this information is not to be disseminated to the entire network, but

locally stored and requested when the lightpath setup process is passing via a particular network element. For the advantages and characteristics of distributed signaling protocol, please refer to section 2.6.2. In [24], the authors proposed proper extensions (i.e., objects) to GMPLS RSVP-TE Path and PathError messages, in order to convey attenuation parameter as a physical link parameter, estimate the quality of received optical signal, and notify potential path rejection due to physical parameters being out of an acceptable range, respectively. Simply replace the attenuation parameter by the nonlinear parameter (FWM, in this case); the signaling specific to our design is accomplished.

4.8. SUMMARY

IP with GMPLS directly over WDM is recognized as the promising architecture for NGNs, which utilizes all-optical transparent model in its physical layer without OEO conversion for higher efficiency. However, the architecture itself raises the coupling issue of optical and network layers, which just begin drawing the attention of researchers.

Section 4.2 ~ 4.5 aims at highlight the necessity of evaluating effective network performance with cross-layer approach when we insist not to sacrifice any optical transparency, i.e. the entire network of interest is operated without OEO conversion capability. We elaborated a comprehensive model, which captures all the major linear and nonlinear impairments in such system, and a systematic three-step approach that is suitable for analyzing cross-layer effect of the optical layer and network layer performance. Our results show that the optical impairments will apparently degrade the network performance in terms of effective normalized throughput beyond certain threshold, where the fiber nonlinearities gain the dominancy over other impairments. The effects are different with varying system configuration and fiber types. On the other hand, results also show that different route selection of the LSPs will affect the Q parameter of the paths as well. Our results also point out that by setting a fix threshold for Q , e.g., $Q=7$,

for a network as the criteria to set up an LSP will under utilize the network since the network might still perform perfectly with specific input power and/or types of packets.

This model and cross-layer approach can be used by all-optical IP over WDM system designers in the future as a guideline to achieve the desired network performance when the isolation of IP traffic performance in the network layer and optical characteristics in the physical layer is no longer valid and full transparency is not allowed to be sacrificed.

However, current technology development limits the ultimate IP over WDM all-optical network with full transparency, for all-optical 3R on WDM is still not able to be done in near or medium term. Therefore, fully transparent network is hard to achieve. For the intermediate – step development, translucent optical network architecture seeks a graceful balance between network design cost and service provisioning performance.

Furthermore, we find out among the choices of translucent network architecture, a collection of islands of transparency satisfies the cost-saving goal as well as ease of migration towards the fully-transparent network. But previous works on the diameter of one island were decided upon the most-impaired link across the island, which usually resulted in overly-conservative style of network design. This limited the size of the islands to be smaller than necessary and increased the overall amount thus the cost of utilizing OEO equipments at the edge of these islands.

The comprehensive study and performance result of systems with different fiber link types and transmission characteristics in section 4.6 inspired us to present the design for the translucent network of interest. The network is designed in such a way that we can implement normal RWA within a newly defined “larger” island of transparency and use the impairment-aware signaling during the path setup. This is because we find the linear and nonlinear impairments impact the network differently and at different levels. Linear impairments can be considered to define the boundary of the islands of transparency while nonlinear impairments are considered to when doing intra-island signaling.

Chapter 5. Conclusions and Recommendations for Future Work

5.1. CONCLUSIONS

Considering the increasing expectations being placed on backbone networks, the benefits provided by MPLS make it a promising technology. MPLS can enhance backbone network scalability and performance while providing advanced services in a simple package. Therefore with desire to upgrade an existing backbone network or plan a next-generation backbone network, MPLS is a concept to bear in mind. Furthermore, extended from MPLS, GMPLS will be an integral part of deploying the next generation of data networks, which provides the necessary bridges between the IP and photonic layers to allow for interoperable and scalable parallel growth in the IP and photonic dimensions. With GMPLS dynamically bridging the gap between the traditional transport infrastructure and the IP layers, the path is being paved for rapid service deployment and operational efficiencies, as well as increased revenue opportunities. By streamlining support for multiplexing and switching in a hierarchical fashion and combining the flexible intelligence of MPLS traffic engineering, the business value of optical switching GMPLS will prove essential in any solution that aims to enable large volumes of traffic in a cost-efficient manner for service providers.

This project aims at comprehensive study of modeling and performance analysis of IP with MPLS/GMPLS over backbone networks, which can be copper – based or fiber optics–based.

In chapter 3, the proper traffic classification and service mapping of the IntServ, DiffServ and MPLS/GMPLS networks have been proposed, which can be used to facilitate the interoperation of IntServ and DiffServ customer networks over MPLS/GMPLS backbone networks. To adjust to the heterogeneous nature of the network of interest, a four-queue system has been proposed, which can be used at output ports of every modified MPLS/GMPLS LSR/OXC to ensure the desired QoS of the heterogeneous networks. The detailed structures of the queueing system of both electronic form and optical form have been shown. The electronic queues can be used at the electronic MPLS LSR nodes in the electronic – based networks while the optical queues can be used at the GMPLS OPS nodes in the all-optical OPS networks. The design of our optical buffer is realizable by implementing current state-of-the-art development of optics and due to its unique working mechanism, our optical buffer functionally is equivalent to its electronic counterpart. The analytical performance parameters of the proposed queueing system in terms of throughput, loss probability and delay without considering the optical layer impairments have been derived. Furthermore, numerical results, illustrating the performance parameters with combination of different traffic load proportions and various traffic priorities, have been demonstrated.

In chapter 4, we focus on all-optical networks. Firstly, if a network is strictly all-optical, which means no OEO conversion is done at any intermediate node of the network. Due to the coupling effect of network layer and physical layer in such an IP over WDM all-optical network, a novel approach to model the optical impairments at the physical layer of the IP/GMPLS over WDM infrastructure has been proposed considering the linear impairments as well as the nonlinear impairments both along the optical paths and at the optical nodes. Moreover, the performance parameter in terms of BER and Q parameter at the optical layer are derived analytically and end-to-end network performance parameters in terms of throughput and delay have been derived. The numerical results obtained then inspired us to do a high-level design of intermediate-step NGN architecture due to physical layer limitation in developing optical 3R. In the design, we

adopt a translucent network, in the form of a collection of islands of transparency or transparent islands, as the highest level network architecture. Instead of using worst-case physical layer impairments as a criterion to emulate the perfect physical layer transmission medium within a transparent domain, we define a larger island of transparency. The reason is that worst-case design results in a very conservative size of domain, which in turn results in more OEO transponders needed for the same size of backbone network. Higher cost and more electronic bottlenecks thus will be present. We use worst-case linear impairments to define the transparent domain. In this way, we can ensure that in most of the cases, the lightpaths can be set up with satisfactory physical layer performance. Therefore, we can use normal RWA algorithm to do the routing and wavelength assignment, which will not add more scalability and stability issues to the already burdened routing protocols such as OSPF – TE. Inside this “larger” island of transparency, in order to ensure the lightpaths to be set up are of good quality, impairment-aware signaling need to be done.

5.2. RECOMMENDATIONS FOR FUTURE WORK

There are a number of ways in which the present work may be extended.

We have proposed a realizable optical buffer system that can be used at the output port of a network node in an OPS all-optical network. We put our emphasis of the all-optical OPS node on the buffer management and QoS threshold control within the optical output buffer, hence we don't restrain our optical buffer to any specific optical packet switch architecture. However, the architecture of the packet switch needs to be carefully designed in order to fully utilize our proposed optical buffer. Researchers are encouraged to investigate further in this topic. Furthermore, if there are laboratories with the state-of-the-art physical layer development of recirculation-enabled FDLs available, people can actually physically build such a buffer. In this way, we can compare our numerical results with the experimental

data and do some comments and investigations. During the experiment, the number of cycles for an optical packet to recirculate in the FDL is up to the physical layer impairments limit due to the loss incurred, if the recirculation circuit is passive. Even with fiber amplifier as EDFA to compensate the loss, accumulated noise due to ASE effect of EDFA will still be a concern. Hence, physical layer impairments may cause some serious degradation when the node is heavily loaded so that the packets have to be recirculated more cycles before departure. By obtaining data experimentally as well as numerically, we can do some cross-layer design of the optical buffer to optimize its performance in terms of loss probability as well as delay.

Apart from the FDL-based buffer, as we mentioned slow-light based optical buffer can offer more flexibility in designing the function of optical buffer. It is the propagation of an optical pulse or other modulation of an optical carrier at a very low group velocity. The term is usually only applied when the velocity is at least hundreds of times slower than the speed of light in a vacuum. There are many mechanisms which can generate slow light, all of which create narrow spectral regions with high dispersion, i.e. peaks in the dispersion relation. In this way, different wavelengths inside the same fiber will experience slowing down effect differently. Hence we can effectively utilize a single slow light device to realize buffers with different sizes. However, due to the Delay-Bandwidth Product (DBP) of slow light, to achieve delay that is useful for storing an optical packet, the bandwidth useful for data transmission is too low for high data rate application. Hence, people can choose to design buffer virtually first and wait for the breakthroughs in the physical layer characteristics of slow light.

The design of the translucent network also presents a possibility to do some future work. For example, if the inter domain traffic is not that much, the operation mechanism of the impairment-aware signaling and re-routing via nearest OEO nodes are fine. However, if the network characteristics is that the inter domain traffic is heavy, modifications about assigning OEO boundary nodes should be done in order to cater for the specific traffic

requirements. Because in that case, the boundary nodes will be very busy in passing traffic from and to neighboring domains and inter domain traffic is likely traveling more nodes than interdomain traffic, so the probability of resorting to OEO nodes is higher. In this case, if too few OEO boundary nodes are assigned, they may appear as the bottlenecks. But this is specific to different traffic condition and customized designs towards various specific conditions need to be applied.

References

- [1] P. Ferguson and G. Huston, *Quality of Service: Delivering QoS in the Internet and Corporate Network*, Wiley, New York, NY, 1998.
- [2] R. Braden, D. Clark and S. Shenker, “Integrated services in the internet architecture: an overview”, RFC 1633, Jun. 1994.
- [3] R. Braden, L. Zhang, S. Berson, S Herzog, and S. Jamin, “Resource reservation protocol (RSVP) – version 1 functional specification”, RFC 2205, Sept. 1997.
- [4] S. Blakes, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, “An architecture for differentiated services,” RFC 2475, Dec. 1998.
- [5] Bruce Davie and Yakov Rekhter, *MPLS: Technology and Applications*, Morgan Kaufmann Publishers, 2000.
- [6] Xipeng Xiao and L.M. Ni, “Internet QoS: a big picture”, *IEEE Network*, vol. 13, no. 2, pp. 8–18, Mar/Apr. 1999.
- [7] G. P. Agrawal, *Fiber-optic communication systems*, 2nd Edition, John Wiley & Sons, New York, 1997.
- [8] W. Stallings, *Data and Computer Communications*, 8th Edition, Prentice Hall, 2007.
- [9] F. Callegati, M. Casoni and C. Raffaelli, “Packet optical networks for high-speed TCP-IP backbones,” *IEEE Commun. Mag.*, vol. 37, no. 1, pp. 124–129, Jan. 1999.
- [10] V.W.S. Chan, K.L. Hall, E. Modiano, and K.A. Rauschenbach, “Architectures and technologies for high-speed optical data networks”, *J. Lightwave Technol.*, vol.16, no.12, pp. 2146–2168, Dec. 1998.
- [11] E. Modiano, “WDM-based packet networks”, *IEEE Commun. Mag.*, vol. 37, no. 3, pp.130–135, Mar. 1999.
- [12] A. McGuire and P. Bonenfant, “Standards: the blueprints for optical networking,” *IEEE Commun. Mag.*, vol. 36, no. 2, pp. 68–78, Feb. 1998.

References

- [13] C. Y. Metz, *IP Switching, Protocols and Architectures*, McGraw–Hill, Dec. 1998.
- [14] D. Awduche and Y. Rekhter, “Multiprotocol lambda switching: combining MPLS traffic engineering control with optical crossconnects,” *IEEE Commun. Mag.*, vol. 39, no. 3, pp. 111–116, Mar. 2001.
- [15] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella and Y. Rekhter, “Generalized multiprotocol label switching: an overview of routing and management enhancements,” *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 144–150, Jan. 2001.
- [16] A. Banerjee, J. Drake, J.P. Lang, Turner, D. Awduche, K. Kompella and Y. Rekhter, “Generalized multiprotocol label switching: an overview of signaling enhancements and recovery techniques,” *IEEE Commun. Mag.*, vol. 39, no. 7, pp. 144–151, Jul. 2001.
- [17] B. Doshi, “Towards optical core: demand and technology driven future WAN architecture,” *WCC – ICCT 2000*, vol. 1, pp. 618, Aug. 2000.
- [18] S.M. Sam and N. Fisal, “Performance study of IP over SONET,” *APCC 2003*, vol. 2, pp. 467–472, Sept. 2003.
- [19] S. Kasera, *ATM networks: concepts and protocols*, New York: McGraw-Hill, c2007.
- [20] K. Ahmad, *Sourcebook of ATM and IP internetworking*, New York, NY: Wiley-Interscience, c2002.
- [21] B. G. Lee and W. J. Kim, *Integrated broadband networks: TCP/IP, ATM, SDH/SONET, and WDM/Optics*, Boston: Artech House, c2002.
- [22] H. V. Helvoort, *SDH/SONET explained in functional models: modeling the optical transport network*, Chichester : Wiley, c2005.
- [23] Junhua Tang, Chee Kheong Siew and Liren Zhang, “Optical nonlinear effects on the performance of IP traffic over GMPLS-based DWDM networks,” *Computer Communications*, vol. 26, no. 12, pp. 1330–1340, Jul. 2003.
- [24] F. Cugini, N. Andriolli, L. Valcarenghi and P. Castoldi, “Physical Impairment Aware Signaling for Dynamic Lightpath Set Up,” *ECOC 2005*, vol. 4, pp. 979 – 980, Sept. 2005.

References

- [25] Y. Huang, A. Gencata, J. P. Heritage and B. Mukherjee, "Routing and Wavelength Assignment with Quality – of – Signal Constraints in WDM Networks," ECOC 2002, vol. 3, pp. 1–2, Sept. 2002.
- [26] S. Shenker, C. Partridge and R. Guerin, "Specification of guaranteed quality of service," RFC 2212, Sept. 1997.
- [27] J. Wroclawski, "Specification of the controlled-load network element service," RFC 2211, Sept. 1997.
- [28] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski and E. Felstaine, "A framework for integrated services operation over diffserv networks," RFC 2998, Nov. 2000.
- [29] V. Jacobson, K. Nichols and K. Poduri, "An expedited forwarding PHB (per-hop behavior)," RFC 2598, Jun. 1999.
- [30] J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured forwarding PHB Group," RFC 2597, Jun. 1999.
- [31] E. Rosen, A. Viswanathan and R. Callon, "Multiprotocol label switching architecture", RFC 3031, Jan. 2001.
- [32] F.L. Faucheur, L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval and J. Heinanen, "Multi-protocol label switching (MPLS) support of differentiated services," RFC 3270, May 2002.
- [33] R. Coltun, D. Ferguson and J. Moy, "OSPF for IPv6," RFC 2740, Dec. 1999.
- [34] R. Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments," RFC 1195, Dec.1990.
- [35] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1171, Mar. 1995.
- [36] J. Ash, M. Girish, E. Gray, B. Jamoussi and G. Wright, "Applicability Statement for CR-LDP," RFC 3213, Jan. 2002.
- [37] L. Andersson, P. Doolan, N. Feldman, A. Fredette and B. Thomas, "LDP Specification," RFC 3036, Jan. 2001.
- [38] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209, Dec. 2001.

References

- [39] C. Brown, F. Baker and C. Carvalho, "Management Information Base for Frame Relay DTEs," RFC 1315, Apr. 1992.
- [40] K. Tesink, Editor, "Definitions of Managed Objects for ATM Management," RFC 2515, Feb. 1999.
- [41] Multiprotocol Label Switching (MPLS) – Cisco Systems, Website, http://www.cisco.com/en/US/products/ps6557/products_ios_technology_home.html
- [42] Nortel Networks: Customer Solutions – MPLS Technology, Website, <http://www.nortelnetworks.com/corporate/technology/mpls/index.html>
- [43] Lucent Product Cuts Reflect De-Emphasis of ATM, MPLS, Website, http://telephonyonline.com/mag/telecom_lucent_product_cuts/
- [44] MPLS: Juniper Networks IP-Service MPLS, Website, http://www.juniper.net/products/eseries/ip_services_mpls.html
- [45] T. Chahed, G. Hebuterne and C. Fayet, "Mapping of loss and delay between intserv and diffserv," ECUMN 2000, pp. 48-55, Oct 2000.
- [46] J. Harju and P. Kivimaki, "Co-operation and comparison of DiffServ and IntServ: performance measurements," in Proc. IEEE LCN 2000, pp. 177 – 186, Nov. 2000.
- [47] W. Almesberger, S. Giordano, R. Mameli, S. Salsano, F. Salvatore, "A prototype implementation for the IntServ operation over DiffServ networks," in Proc. IEEE GLOBECOM '00, vol.1, pp. 439 – 444, Nov./Dec. 2000.
- [48] G. Yuan, W.D. Wang, Y. Lin and S. D. Cheng, "A QoS management architecture for DiffServ-aware MPLS-based IP network," in Proc. ICCT 2003, vol. 2, pp. 1603 – 1607, Apr. 2003.
- [49] S. H. Lim, M. H. Yaacob, K. K. Phang and T. C. Ling, "Traffic engineering enhancement to QoS-OSPF in DiffServ and MPLS networks," in Proc. IEE Communications, vol. 151, no. 1, pp. 101 – 106, Feb. 2004.
- [50] A. Alwehaibi, M. Kadoch and A. Elhakeem, "Packet loss probability for DiffServ over heterogeneous MPLS multicast networks: a simulation study," Canadian Conference on Electrical and Computer Engineering, vol. 4, pp. 2209 – 2212, May 2004.

References

- [51] L. A. Saidane, P. Minet, S. Martin and I. Korbi, "Deterministic and probabilistic QoS guarantees for real-time traffics in a DiffServ/MPLS architecture," IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, pp. 337 – 340, Sept. 2005.
- [52] F. Tommasi, S. Molendini, and A. Tricco, "Mapping of IntServ/RSVP reservations into MPLS domains", University of Lecce, ITALY.
- [53] N. Rouhanna and E. Horlait, "Differentiated services and integrated services use of MPLS", ISCC 2000, pp. 194–199, Jul 2000.
- [54] Q. Du, "A monotonicity result for a single-server queue subject to a Markov-modulated Poisson process," J. Appl. Prob., vol. 32, no. 4, pp. 1103–1111, 1995.
- [55] Meier-Hellstern K.S. and Fischer W., The Markov-modulated Poisson process (MMPP) cookbook, Performance Evaluation 18, pp. 149-171, 1992.
- [56] H. Heffes and D. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," IEEE J. Select. Areas Commun., vol 4, no 6, pp. 856- 868, Sept. 1986.
- [57] R. Gusella, "Characterizing the Variability of Arrival Processes with Indexes of Dispersion," IEEE J. Select. Areas Commun., vol. 9, no. 2, pp. 203-211, Feb. 1991.
- [58] T. Okuda, H. Akimaru and M. Sakai, "A simplified Performance Evaluation for Packetized Voice Systems," IEICE Trans., vol. E73, no. 6, 1990.
- [59] Å. Arvidsson C. Lind, Using Markovian Models to replicate Real ATM traffics, ATM networks: Performance Modelling and Evaluation, vol. 2, Chapman & Hall, London, England, pp. 39 – 54, 1996.
- [60] J. N. Daigle and D. M. Lucantoni, Queueing Systems Having Phase-Dependent Arrival and Service Rates, Numerical Solution of Markov Chains, William J. Stewart, Ed., Marcel Dekker, pp 161 – 202, 1991.
- [61] K. Sriram and W. Whitt, "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," IEEE J. Select. Areas Commun., vol. 4, no. 6, pp. 833 – 846, Sept. 1986.

References

- [62] W. Whitt, "The queueing network analyzer," *Bell Syst. Tech. J.*, Part 1, vol. 62, no. 9, pp. 2779-2815, Nov. 1983.
- [63] W. Whitt, "Queues with superposition arrival processes in heavy traffic," *Stoch. Proc. Appl.*, vol. 21, pp. 81-91, 1985.
- [64] D. Mitra, "Stochastic theory of a fluid model of producers and consumers coupled by a buffer", *Adv. Appl. Prob.*, vol. 20, no. 3, pp. 646-676, Sept. 1988.
- [65] D. Anick, D. Mitra, and M. M. Sondhi, "Stochastic Theory of a Data – handling System with Multiple Sources," *Bell System Tech. J.* 61, pp. 1871 – 1894, Oct. 1982.
- [66] T. E. Stern and A. I. Elwalid, "Analysis of Separable Markov-Modulated Rate Models for Information-Handling Systems," *Adv. in Appl. Prob.*, vol. 23, no. 1, pp. 105-139, Mar. 1991.
- [67] E. G. Coffman, B. M. Igel'nik and Y. A. Kogan, "Controlled stochastic model of a communication system with multiple sources," *IEEE Trans. Info. Theory*, vol. 37, no. 5, pp. 1379 – 1387, Sept. 1991.
- [68] L. Dittmann and S. B. Jacobsen, "Statistical multiplexing of identical bursty sources in an ATM network," in *Proc. GLOBECOM'88*, vol. 3, pp. 1293 – 1297, Nov/Dec 1988.
- [69] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J.D. Robbins, "Performance models of statistical multiplexing in packet video communications," *IEEE Trans. Commun.*, vol. 36, no. 7, pp. 834-844, Jul. 1988.
- [70] M. Butto, E. Cavallero and A. Tonietti, "Effectiveness of the leaky bucket: policing mechanism in ATM networks," *IEEE J. Select. Areas Commun.*, vol. 9, no. 3, pp. 335 – 342, Apr. 1991.
- [71] I. Norros, J. W. Roberts, A. Simonian and J. T. Virtamo, "The superposition of variable bit rate sources in an ATM multiplexer," *IEEE J. Select. Areas Commun.*, vol. 9, no. 3, pp. 378 – 387, Apr. 1991.
- [72] J. Daigle and J. Langford, "Models for Analysis of Packet Voice Communications Systems," *IEEE J. Select. Areas Commun.*, vol. 4, no. 6, pp. 847 – 855, Sept. 1986.

References

- [73] R. Nagarajan, J. F. Kurose, and D. Towsley, "Approximation techniques for computing packet loss infinite-buffered voice multiplexers," *IEEE J. Select. Areas Commun.*, vol. 9, no. 3, pp. 368 – 377, Apr. 1991.
- [74] H. Kroner, T. H. Theimer and U. Briem, "Queueing models for ATM systems – a comparison," in *Proc. 7th ITC Specialist Seminar*, Morristown, paper 9.1, 1990.
- [75] G. Shen, W. D. Grover, T. H. Cheng and S. K. Bose, "Sparse Placement of electronic switching nodes for low blocking in translucent optical networks," *OSA J. Opt. Net.*, vol. 1, no. 12, pp. 424 – 441, Dec. 2002.
- [76] A. L. Souza Filho, H. Waldman, "Strategies for designing Translucent Wide Area Networks," in *Proc. IMOC 2003*, vol. 2, pp. 931 – 936, Sept. 2003.
- [77] S. Subramaniam, A. Azizoglu and A. Somani, "Connectivity and Sparse Wavelength Conversion in Wavelength routing networks," in *Proc. IEEE INFOCOM 1996*, vol. 1, pp.148–155, Mar. 1996.
- [78] E. Goldstein, J. Nagel, J. Strand and R. Tkach, "National-scale networks likely to be opaque," *Lightwave* 2, pp. 91–95, Feb. 1998.
- [79] B. Ramamurthy, H. Feng, D. Datta, J. P. Heritage and B. Mukereje, "Transparent vs. opaque vs. translucent wavelength-routed optical networks," *OFC/IOOC 1999 Tech. Dig.*, vol. 1, pp.59–61, Feb.1999.
- [80] K. Bala, R. R. Cordell and E. L. Goldstein, "The case for opaque multiwavelength optical networks," *IEEE/LEOS Summer Topical Meeting 1995 Dig.*, pp.58–59, Aug. 1995.
- [81] A. A. M. Saleh, "Transparent optical networking in backbone networks," in *Proc. OFC 2000*, vol. 3, pp. 62 – 64, Mar. 2000.
- [82] A. Morea and J. Poirrier, "A Critical Analysis of the Possible Cost Savings of Translucent Networks," *DRCN 2005*, pp. 311–317, Oct. 2005.
- [83] B. Ramamurthy, S. Yaragorla and X. Yang, "Translucent Optical WDM Networks for the Next-Generation Backbone Networks," in *Proc. IEEE GLOBECOM 2001*, vol. 1, pp. 60 – 64, Nov. 2001.

References

- [84] W. J. Tomlinson and C. Lin, "Optical wavelength-division multiplexer for the 1–1.4 micron spectral region," *Electron. Lett.*, vol. 14, pp. 345–347, May 1978.
- [85] H. Ishio, J. Minowa and K. Nosu, "Review and status of wavelength-division-multiplexing technology and its application," *J. Lightwave Technol.*, vol. 2, no. 4, pp. 448–463, Aug. 1984.
- [86] A. Arora and S. Subramaniam, "Wavelength Conversion Placement in WDM Mesh Optical Networks," *Photonic Network Communications*, vol. 4, no 2, pp. 167 – 177, Nov. 2004.
- [87] I. Kaminow and T. Li, Editor, *Optical Fiber Telecommunications IVB – systems and impairments*, Academic Press, Elsevier Science (USA), 2002.
- [88] J. Hecht, *Understanding fiber optics*, 3rd Ed., Prentice-Hall: New Jersey, 1999.
- [89] S. Nagamatsu, B. Pálsdóttir, T. Hirasawa, A. Fujisaki, R. Takaoka and S. Shikii, "Raman Amplifier with Integrated Dispersion-Compensating Fiber," *Furukawa Review*, no. 24, pp.13–16, 2003.
- [90] Mark Nietubyc, *Raman Amplification in the All-Optical Network – Technology Information*, Telecommunications, http://www.findarticles.com/p/articles/mi_m0TLC/is_6_35/ai_76472139, Jun. 2001.
- [91] R. J. Mears, L. Reekie, I. M. Jauncey and D. N. Payne, "Low-noise Erbium-doped fibre amplifier at 1.54 μm ," *Electron. Lett.*, vol. 23, no. 19, pp.1026–1028, Sept. 1987.
- [92] E. Desurvire, *Erbium-Doped Fiber Amplifiers, Principles and Applications*, Wiley-Interscience, 2002.
- [93] C. Saxtoft and P. Chidgey, "Error rate degradation due to switch crosstalk in large modular switched optical networks," *IEEE Photon. Technol. Lett.*, vol. 5, no. 7, pp. 828–831, Jul. 1993.
- [94] J. Zhou, M.J. O'Mahony, and S.D. Walker, "Analysis of optical crosstalk effects in multi-wavelength switched networks," *IEEE Photon. Technol. Lett.*, vol. 6, no. 2, pp. 302–305, Feb. 1994.
- [95] R. A. Spanke, "Architectures for guided-wave optical space switching systems," *IEEE Commun. Mag.*, vol. 25, no. 5, pp. 42–48, May 1987.

References

- [96] I. Kaminow and T. Li, Editor, Optical Fiber Telecommunications IVA – components, Academic Press, Elsevier Science (USA), 2002.
- [97] K. O. Hill, D. C. Johnson, B. S. Kawasaki, R. I. MacDonald, “CW three-wave mixing in single-mode optical fibers,” *J. Appl. Phys.*, vol. 49, no. 10, pp. 5098–5106, Oct. 1978.
- [98] R. G. Waarts and R. P. Braun, “System limitations due to four-wave mixing in single mode optical fibers,” *Electronics Letters*, vol. 22, no. 16, pp. 873–875, Jul. 1986.
- [99] N. Shibata, R. P. Braun and R. G. Waarts, “Phase-mismatch dependence of efficiency of wave generation through four-wave mixing in a single mode optical fiber,” *IEEE J. Quantum. Electron.*, vol. 23, no. 7, pp. 1205 – 1210, Jul. 1987.
- [100] K. Inoue, “Phase-mismatching characteristic of four-wave mixing in fiber lines with multi stage optical amplifiers,” *Opt. Lett.*, vol. 17, no. 11, pp. 801–803, Jun. 1992.
- [101] W. Zeiler, F. Di Pasquale, P. Bayvel and J. E. Midwinter, “Modeling of four-wave mixing and gain peaking in amplified WDM optical communication systems and networks,” *J. Lightwave Technol.*, vol. 14, no. 9, pp. 1933 – 1942, Sept. 1996.
- [102] M. Eiselt, “Limits on WDM systems due to four-wave mixing: a statistical approach,” *J. Lightwave Technol.*, vol. 17, no. 11, pp. 2261–2267, Nov. 1999.
- [103] K. Inoue, K. Nakanishi, K. Oda and H. Toba, “Crosstalk and power penalty due to fiber four-wave mixing in multi channel transmission,” *J. Lightwave Technol.*, vol. 12, no. 4, pp. 1423 – 1439, Aug. 1994.
- [104] F. Shimizu, “Frequency broadening in liquids by a short light pulse,” *Phys. Rev. Lett.*, vol. 19, no. 19, pp. 1097 – 1100, Nov. 1967.
- [105] T. K. Gustafson, J. P. Taran, H. A. Haus, J. R. Lifshitz and P. L. Kelley, “Self-Modulation, Self-Steepening, and Spectral Development of Light in Small-Scale Trapped Filaments,” *Phys. Rev.*, vol. 177, no. 1, pp. 306 – 313, Jan. 1969.
- [106] C. H. Lin and T. K. Gustafson, “Optical pulsewidth measurement using self-phase modulation,” *IEEE J. Quantum Electron.*, vol. 8, no. 4, pp. 429 – 430, Apr. 1972.
- [107] A. R. Chraplyvy, “Limitations on lightwave communications imposed by optical-fiber nonlinearities,” *J. Lightwave Technol.*, vol. 8, no. 10, pp. 1548 – 1557, Oct. 1990.

References

- [108] R. G. Waarts, A. A. Friesem, E. Lichtman, H. H. Yaffe and R. P. Braun, "Nonlinear effects in coherent multichannel transmission through optical fibers," in *Proc. of IEEE*, vol. 78, no. 8, pp. 1344 – 1368, Aug. 1990.
- [109] K. H. Kim, H. K. Lee, S. Y. Park and E. H. Lee, "Calculation of dispersion and nonlinear effect limited maximum TDM and FDM bit rates of transform-limited pulses in single-mode optical fibers," *J. Lightwave Technol.*, vol. 13, no. 8, pp. 1597 – 1605, Aug. 1995.
- [110] A. V. T. Cartaxo, "Cross-phase modulation in intensity modulation-direct detection WDM systems with multiple optical amplifiers and dispersion compensators," *J. Lightwave Technol.*, vol. 17, no. 2, pp. 178 – 190, Feb. 1999.
- [111] R. Hui, K. R. Demarest and C. T. Allen, "Cross-phase modulation in multispan WDM optical fiber systems," *J. Lightwave Technol.*, vol. 17, no. 6, pp. 1018 – 1026, Jun. 1999.
- [112] M. Shtaif, M. Eiselt, R. W. Tkach, R. H. Stolen and A. H. Gnauck, "Crosstalk in WDM systems caused by cross-phase modulation in erbium-doped fiber amplifiers," *IEEE Photon. Technol. Lett.*, vol. 10, no. 12, pp. 1796 – 1798, Dec. 1998.
- [113] M.W. Maeda, W.B. Sessa, W.I. Way, A. Yi-Yan, L. Curtis, R. Spicer and R.I. Laming, "The effect of four-wave mixing in fibers on optical frequency-division multiplexing systems," *J. Lightwave Technol.*, vol. 8, no. 9, pp. 1402–1408, Sept. 1990.
- [114] G. P. Agrawal, *Nonlinear Fiber optics*, Academic Press, 2001.
- [115] G. P. Agrawal, *Fiber nonlinearities and their applications*, Academic Press, 2000.
- [116] N. Bloembergen, "The stimulated Raman effect," *Am. J. Phys.*, vol. 35, pp. 989, 1967.
- [117] E. P. Ippen and R. H. Stolen, "Stimulated Brillouin scattering in optical fibers," *Appl. Phys. Lett.*, vol. 21, pp 539, 1972.
- [118] E. Lannone, F. Matera, A. Mecozzi and M. Settembre, *Nonlinear Optical Communications Networks*, New York, John Wiley & Sons, 1998.
- [119] Kiyoshi Nosu, *Optical FDM Network Technologies*, Artech House, 1997.

References

- [120] Kyo Inoue, "A simple expression for optical FDM network scale considering fiber four-wave mixing and optical amplifier noise," *IEEE J. Lightwave Technol.*, vol. 13, no. 5, pp. 856 – 861, May 1995.
- [121] Liren Zhang and Junhua Tang, "Characterization and performance study of IP traffic in WDM networks," *Computer Communications*, vol. 24, no. 17, pp. 1702–1713, Nov. 2001.
- [122] X. Liu, C. R. Giles, X. Wei, Y. H. Kao, C. Xie, L. Moller, I. Kang and A. –J. van Wijngaarden, "Experimental demonstration scrambling and FEC," *ECOC 2004 Tech. Dig.*, Post–deadline paper Th 4.1.4, Sept. 2004.
- [123] N.A. Olsson, "Lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 7, no. 7, pp. 1071–1082, Jul. 1989.
- [124] H. Takahashi, K. Oda and H. Toba, "Impact of crosstalk in an arrayed–waveguide multiplexer on NxN optical interconnection," *J. Lightwave Technol.*, vol. 14, no. 6, pp. 1097–1105, Jun. 1996.
- [125] B. Ramanurthy, D. Datta, H. Feng, J. P. Heritage and B. Mukherjee, "Impact of transmission impairments on the teletraffic performance of wavelength–routed optical networks," *J. Lightwave Technol.*, vol. 17, no. 10, pp. 1713–1723, Oct. 1999.
- [126] I. Chlamtac, A. Fumagalli, L. G. Kazovsky, P. Melman, W.H. Nelson, P. Poggiolini, M. Cerisola, A.N.M.M. Choudhury, T. K. Fong, R. T. Hofmeister, Chung-Li Lu, A. Mekittikul, D. J. M. Sabido, Chang-Jin Suh and E. W. M. Wong, "CORD: contention resolution by delay lines," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 1014 – 1029, Jun. 1996.
- [127] R. Langenhorst , M. Eiselt , W. Pieper, G. Grosskopf, R. Ludwig , L. Kuller, E. Dietrich and H. G. Weber, "Fiber loop optical buffer," *J. Lightwave Technol.*, vol. 14, no. 3, pp. 324 – 335, Mar. 1996.
- [128] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen and I. Andonovic, "SLOB: a switch with large optical buffers for packet switching, "," *J. Lightwave Technol.*, vol. 16, no. 10, pp. 1725 – 1736, Oct. 1998.

References

-
- [129] D. K. Hunter, M. C. Chia and I. Andonovic, "Buffering in optical packet switches," *J. Lightwave Technol.*, vol. 16, no. 12, pp. 2081 – 2094, Dec. 1998.
 - [130] H. Harai, N. Wada, F. Kubota and W. Chujo, "Contention resolution using multi-stage fiber delay line buffer in a photonic packet switch," in *Proc. IEEE ICC'02*, vol 5, pp. 2843 – 2847, Apr/May 2002.
 - [131] Y. Liu, M.T. Hill, N. Calabretta, H. de Waardt, G. D. Khoe, H.J.S. Dorren, "All-optical buffering in all-optical packet switched cross connects," *IEEE Photon. Technol. Lett.*, vol. 14, no. 6, pp. 849 – 851, Jun. 2002.
 - [132] Lin Li, S.D. Scott, J.S. Deogun, "A novel fiber delay line buffering architecture for optical packet switching," in *Proc. IEEE GLOBECOM'03*, vol. 5, pp. 2809 – 2813, Dec. 2003.
 - [133] A. Chowdhury, Yong-Kee Yeo; Jianjun Yu; Gee-Kung Chang, "DWDM reconfigurable optical delay buffer for optical packet switched networks," *IEEE Photon. Technol. Lett.*, vol. 18, no. 10, pp. 1176 – 1178, May 2006.
 - [134] R. S. Tucker, P. C. Ku and C. J. Chang – Hasnain, "Slow-light optical buffers: capabilities and fundamental limitations," *J. Lightwave Technol.*, vol. 23, no. 12, pp. 4046 – 4066, Dec. 2005.
 - [135] S. Fu, P. Shum, L. Zhang, C. Wu, and A.M. Liu, "Design of SOA-based dual-loop optical buffer with a 3 /spl times/ 3 collinear coupler: guideline and optimizations," *J. Lightwave Technol.*, vol. 24, no. 7, pp. 2768 – 2778, Jul. 2006.
 - [136] A. S. Acampora and I. A. Shah, "Multihop Lightwave Networks: A Comparison of Store-and-Forward and Hot-Potato Routing," *IEEE Trans. Commun.*, vol. 40, no. 6, pp. 1082–1090, Jun. 1992.
 - [137] B. Mukherjee, *Optical Communication Networks*, McGraw–Hill, 1997.
 - [138] F. Forghieri, A. Bononi, and P. R. Prucnal, "Analysis and Comparison of Hot-Potato and Single-Buffer Deflection Routing in Very High Bit Rate Optical Mesh Networks," *IEEE Trans. Commun.*, vol. 43, no.1, pp. 88–98, Jan. 1995.

References

- [139] X. Wang, H. Morikawa and T. Aoyama, "Photonic burst optical deflection routing protocol for wavelength routing networks," *SPIE Opt. Net. Mag.*, vol. 3, no. 6, pp. 12 – 19, Nov/Dec 2002.
- [140] S. Yao, S. J. B. Yoo, B. Mukherjee and S. Dixit, "All-optical packet-switched networks: a study of contention resolution schemes in an irregular mesh network with variable-sized packets," *SPIE OPTICOMM 2000*, vol.4233, pp. 235 - 246, Oct 2000.
- [141] A. Bononi, G. A. Castañón, and O. K. Tonguz, "Analysis of Hot-Potato Optical Networks with Wavelength Conversion," *J. Lightwave Technol.*, vol. 17, no. 4, pp. 525–534, Apr. 1999.
- [142] L. V. Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 meters per second in an ultracold atomic gas," *Nature*, vol. 397, no. 6720, pp. 594–598, Feb. 1999.
- [143] P. C. Ku, F. Sedgwich, C. Chang-Hasnain, P. Palinginis, T. Li, H. Wang, S. W. Chang, and S. L. Chuang, "Slow light in semiconductor quantum wells," *Opt. Lett.*, vol. 29, no. 19, pp. 2291–2293, 2004.
- [144] A. Turukhin, V. S. Sudarshanam, M. S. Shahriar, J. A. Musser, and P. R. Hemmer, "First observation of ultraslow group velocity of light in a solid," in *Tech. Dig., Quantum Electronics Laser Science Conf. (QELS)*, pp. 6–7, May 2001.
- [145] J. Spring and R. S. Tucker, "Photonic 2×2 packet switch with input buffers," *Electron. Lett.*, vol. 29, no. 3, pp. 284–285, Feb. 1993.
- [146] D. B. Hunter and R. A. Minasian, "Photonic signal processing of microwave signals using and active-fiber Bragg-grating-pair structure," *IEEE Trans. Microw. Theory Tech.*, vol. 45, no. 8, pp. 1463–1466, Aug. 1977.
- [147] J. Capmany, J. Cascon, D. Pastor, and B. Ortega, "Reconfigurable fiber- optic delay line filters incorporating electro optic and electroabsorption modulators," *IEEE Photon. Technol. Lett.*, vol. 11, no. 9, pp. 1174–1176, Sept. 1999.

References

-
- [148] R.S. Tucker, P.C. Ku, and C.J. Chang-Hasnain, "Delay-bandwidth product and storage density in slow-light optical buffers," *Electron. Lett.*, vol. 41, no. 4, pp. 208 – 209, Feb. 2005.
 - [149] Z. Zhu, A.M.C. Dawes, D.J. Gauthier, D.J. L. Zhang, and A.E. Willner, "12-GHz-Bandwidth SBS Slow Light in Optical Fibers," *OFC 2006*, pp. 1 – 3, Mar. 2006.
 - [150] C. Jauregui, H. Ono, P. Petropoulos, and D. J. Richardson, "Four-fold reduction in the speed of light at practical power levels using Brillouin scattering in a 2-m bismuth-oxide fiber," *OFC 2006*, pp. 1 – 3, Mar. 2006.
 - [151] M. J. Karol and M. G. Hluchyj, "Input versus output queueing on a space-division packet switch," *IEEE Trans. Commun.*, vol. 35, no. 12, pp. 1347–1356, Dec. 1987.
 - [152] M. G. Hluchyj and M. J. Karol, "Queueing in high-performance packet switching," *IEEE J. Select. Areas Commun.*, vol. 6, no. 9, pp. 1587–1597, Dec. 1988.
 - [153] H. J. Chao, C. H. Lam and E. Oki, *Broadband packet switching technologies: a practical guide to ATM switches and IP routers*. Wiley-Interscience Publication, 2001.
 - [154] S. T. Chuang, A. Geol, N. McKeown, and B. Prabhakar, "Matching output queueing with a combined input/output-queued switch," *IEEE J. Select. Areas Commun.*, vol. 17, no. 6, pp. 1030 – 1039, Jun. 1999.
 - [155] K. J. Warbrick, P. R. Roorda, and D. Pugh, "Performance and scaling of a recirculating optical buffer," in *Proc. London Communications Symp.* 2000, pp. S10–2, Sept. 2000.
 - [156] Y. N. Singh and M. Naik, "Study of power variation of a buffered packet in optical loop buffer," in *Proc. Photonics 2002*, pp. NET-P11–NET-P11, Dec. 2002.
 - [157] A. A. Als, Z. Ghassemlooy, G. Swirth, P. Ball, and J. Chi, "Performance of the passive recirculating fiber loop buffer within an OTDM transmission link," *Opt. Commun.*, vol. 209, pp. 137–147, 2002.
 - [158] X. Hong, Y. Chen and J. Lin, "Erasable fiber loop optical buffer," *APCC/OECC'99*, vol. 2, pp. 1406 – 1409, Oct. 1999.
 - [159] A. M. Liu, C. Q. Wu, Y. D. Gong, and P. Shum, "Optical buffer configuration based on a 3×3 collinear Fiber," *Electron. Lett.*, vol. 40, no. 16, pp. 1017–1019, Aug. 2004.

References

- [160] F. Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, no. 9, pp. 292 – 294, Sept. 2000.
- [161] J. Yan and G. Kuo, "Performance analysis of WDM optical buffers for asynchronous variable length packets", *HSPR 2005*, pp. 302 – 305, May 2005.
- [162] J. Liu, H. Liu, and T. T. Lee, "Performance modeling of optical buffers supporting variable length packets", *HSPR 2005*, pp. 535 – 538, May 2005.
- [163] R. S.Tucker, "The Role of Optics and Electronics in High-Capacity Routers," *J. Lightwave Technol.*, vol. 24, no. 12, pp. 4655 – 4673, Dec. 2006.
- [164] ITU–T Rec. G.8080/Y.1304, "Architecture for the Automatically Switched Optical Network (ASON)," Nov. 2001.
- [165] E. Mannie, Ed., "Generalized Multi–Protocol Label Switching (GMPLS) Architecture," *IETF RFC 3945*, Oct. 2004.
- [166] Y. Huang, J. P. Heritage and B. Mukherjee, "Connection Provisioning with Transmission Impairment Consideration in Optical WDM Networks with High–Speed Channels," *IEEE/OSA J. Lightwave Technol.*, vol. 23, no. 3, pp. 982–993, Mar. 2005.
- [167] A. Farrel, J. P. Vasseur and J. Ash, "A Path Computation Element (PCE)–Based Architecture," *IETF RFC 4655*, Aug. 2006.
- [168] J. Strand, A. L. Chiu, and R. Tkach, "Issues for Routing in the Optical Layer," *IEEE Commun. Mag.*, vol. 39, no. 21, pp. 81–87, Feb. 2001.
- [169] R. Martinez, C. Pinart, F. Cugini, N. Andriolli, L. Valcarenghi, P. Castoldi, L. Wosinska, J. Comellas and G. Junyent, "Challenges and Requirements for Introducing Impairment–Awareness into the Management and Control Planes of ASON/GMPLS WDM Networks," *IEEE Commun. Mag.*, vol. 44, no. 12, pp. 76 – 85, Dec. 2006.
- [170] E. Rosen, D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li and A. Conta, "MPLS label stack encoding," *RFC 3032*, Jan. 2001.
- [171] J.M. Chung, "Analysis of MPLS traffic engineering," in *Proc. IEEE MWSCAS 2000*, vol. 2, pp. 550–553, Aug. 2000.

References

- [172] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell and J. McManus, "Requirements for traffic engineering over MPLS," RFC 2702, Sept. 1999.
- [173] T. Li, "MPLS and the evolving internet architecture", IEEE Commun. Mag., vol. 37, no. 12, pp. 38–42, Dec. 1999.
- [174] R. Comerford, "State of the Internet: Roundtable 5.0," IEEE Spectrum, vol. 36, no. 10, pp. 42–50, Oct. 1999.
- [175] D. Ferrari and L. Delgrossi, "Charging for QoS," IWQoS 1998, pp. VII – VIII, May 1998.
- [176] J. Wroclawski and A. Charny, "Integrated Service Mapping for Differentiated Services Networks", draft-ietf-issll-ds-map-01.txt, Feb. 2001.
- [177] K. Nichols, S. Blake, F. Baker and D. Black, "Definition of the Differentiated Services field (DS Field) in the IPv4 and IPv6 Headers," RFC 2474, Dec. 1998.
- [178] I. Elwalid and D. Mitra, "Fluid models for the analysis and design of statistical multiplexing with loss priorities on multiple classes of bursty traffic", in Proc. IEEE INFOCOM'92, vol. 1, pp. 415–425, May 1992.
- [179] U. N. Bhat, An Introduction to Queueing Theory: Modeling and Analysis in Applications, Boston: Birkhäuser Boston, 2008.
- [180] R. B. Cooper, Introduction to Queueing Theory, 3rd ed., Washington, DC : CEEPress Books, c1990.
- [181] http://www.eventhelix.com/realtimemantra/CongestionControl/m_m_1_queue.htm
- [182] L. Breuer, An Introduction to Queueing Theory and Matrix-analytic Methods, Dordrecht, The Netherlands : Springer, c2005.
- [183] Ng Chee Hock, Queueing modeling fundamentals, New York: John Wiley & Sons, 1996.
- [184] M. Hirano and N. Watanabe, "Characteristics of a cell multiplexer for bursty ATM traffic", ICC 1989, vol. 1, pp.399–403, Jun. 1989.
- [185] D. E. McDysan, D.L. Spohn, ATM: Theory and Application, McGraw–Hill, New York, 1994.

References

- [186] C. Guillemot et al., "Transparent optical packet switching: The European ACTS KEOPS project approach," *J. Lightwave Technol.*, vol. 16, no. 12, pp. 2117–2134, Dec. 1998.
- [187] D. J. Blumenthal et al. "All-optical label swapping networks and technologies," *J. Lightwave Technol.*, vol. 18, no. 12, pp. 2058–2075, Dec. 2000.
- [189] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 12, pp. 2081 – 2094, Dec. 1998.
- [190] M. G. Hluchyj and M. J. Karol, "Queueing in high-performance packet switching," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 1587 – 1597, Dec. 1988.
- [191] E. Modiano and R. Barry, "Architectural consideration in the design of WDM-based optical access networks," *Computer Networks*, vol. 31, no. 4, pp. 327–341, Feb. 1999.
- [192] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) signaling function description," RFC 3471, Jan. 2003.
- [193] B. Mukherjee, Y. Huang and J.P. Heritage, "Impairment-aware routing in wavelength-routed optical networks," in *Proc. IEEE LEOS 2004*, vol. 1, pp. 428–429, Nov. 2004.
- [194] Xu Rong, Gong Qian and Ye Peida, "A novel IP with MPLS over WDM-based broad-band wavelength switched IP network," *J. Lightwave Technol.*, vol. 19, no. 5, pp. 596–602, May 2001.
- [195] H. Taga, S. Yamamoto, N. Edagawa, Y. Yoshida, S. Akiba and K. Wakabayashi, "Fiber chromatic dispersion equalization at the receiving terminal of IM-DD ultra-long distance optical communication systems," *J. Lightwave Technol.*, vol. 12, no. 6, pp 1042–1046, Jun. 1999.
- [196] H. S. Hinton, "Photonic switching using directional couplers," *IEEE Commun. Mag.*, vol. 25, no. 5, pp. 16–26, May 1987.
- [197] Zwillinger, D. *Handbook of Differential Equations*, 3rd ed. Boston, MA: Academic Press, p. 122, 1997.
- [198] R. G. Gallager, *Information Theory and reliable Communication*, Wiley, 1968.

References

- [199] C. Lee, "Analysis of switching networks," *Bell Syst. Tech. J.*, vol. 34, pp. 1287 - 1315, Nov. 1955.
- [200] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan, "Efficient routing and scheduling algorithms for optical networks," *IBM Res. Rep. RC 18967 (82799)*, Tech. Rep., Jun 1993.
- [201] K. Sivarajan and R. Ramaswami, "Lightwave networks based on dehruijn graphs," *IEEE Trans. Networking*, vol. 2, Feb. 1994.
- [202] M. Kovačević and A. Acampora, "On wavelength translation in all-optical networks," in *Proc. IEEE INFOCOM'95*, vol. 2, pp. 413-422, Apr. 1995.
- [203] A. Birman, "Computing approximate blocking probabilities for a class of all-optical networks," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 852 – 857, Jun. 1996.
- [204] R. A. Barry and P. A. Humblet, "Models of blocking probability in all-optical networks with and without wavelength changers," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 858–867, Jun. 1996.
- [205] http://www.acphotonics.com/products/product_files/GFF_for_EDFA.htm
- [206] C. D. Cantrell, "Transparent Optical Metropolitan–Area Networks," in *Proc. IEEE LEOS'03*, vol. 2, pp. 608 – 609, 2003.
- [207] I. Tomkos and A. Tzanakaki, "Towards Digital Optical Networks," in *Proc. ICTON'05*, vol. 1, pp. 1 – 4, Jul. 2005.
- [208] A. Saleh, "Island of Transparency — An Emerging Reality in Multiwavelength Optical Networking," *Proc. IEEE/LEOS Summer Topical Meeting on Broadband Opt. Net. and Tech.*, pp. 36, Jul. 1998.
- [209] E. Karasan and M. Arisoylu, "Design of Translucent Optical Networks: Partitioning and Restoration," *Photonic Network Commun.*, vol. 8, no. 2, pp. 209–21, Feb. 2004.
- [210] G. Shen and W. D. Grover, "Segment–Based Approaches to Survivable Translucent Network Design Under Various Ultra Long Haul System Reach," *OSA J. Opt. Net.*, vol. 3, no. 1, pp. 1–24, Jan. 2004.

References

- [211] F. Cugini, N. Andriolli, L. Valcarenghi, and P. Castoldi, “A Novel Signaling Approach to Encompass Physical Impairments in GMPLS Networks,” IEEE GLOBECOM Workshop 2004, pp. 369 – 373, Nov/Dec 2004.
- [212] N. Sambo, A. Giorgetti, N. Andriolli, F. Cugini, L. Valcarenghi, and P. Castoldi, “GMPLS signaling feedback for encompassing physical impairments in transparent optical networks,” in Proc. IEEE GLOBECOM’06, pp. 1 – 5, Nov/Dec 2006.
- [213] A. Bosco, A. Botta, G. Conte, P. Iovanna, R. Sabella and S. Salsano, “Internet like control for MPLS based traffic engineering: performance evaluation,” Performance Evaluation, vol. 59, no. 2-3, pp. 121 – 136, Feb. 2005.
- [214] Basu and J. Riecke, “Stability issues in OSPF routing,” ACM SIGCOMM 2001, vol. 31, no 4, pp. 225 – 236, Oct. 2001.

Appendix A: Fluid Flow Model

Consider a physical model in which a buffer receives messages from a finite number of statistically independent and identical information sources that asynchronously alternate between exponentially distributed period in the ‘on’ and ‘off’ states.

A data-handling switch receives messages from many, say N , information sources, which independently and asynchronously alternate between the ‘on’ and ‘off’ state. The on periods as well as the off periods are exponentially distributed for each source. These two distributions, while not necessarily identical, are common to all sources; also, the sources are mutually independent. Without loss of generality, the unit of time is selected to be the average on period; with this unit of time, the average off period is denoted by $1/\lambda$. Again, without loss of generality, the unit of information is chosen to be the amount generated by a source in an average on period. In these units an on source transmits at the uniform rate of 1 unit of information per unit of time. Thus, when r sources are on simultaneously, the instantaneous receiving rate at the switch is r . The switch stores or buffers the incoming information that is excess of the maximum transmission rate, c , of an output channel. (Thus, c is also the ratio of the output channel capacity to an on source’s transmission rate.)

As long as the buffer is not empty, the instantaneous rate of change of the buffer content is $r - c$. Once the buffer is empty, it remains so as long as $r \leq c$. We assume that the buffer is infinite and that the stability condition is satisfied as $\rho = \frac{N\lambda}{c(1+\lambda)} < 1$, where ρ is the traffic intensity.

If at time t the number of on sources equals i , two elementary events can take place during the next interval Δt , i.e., a new source can start or a source can turn off. Since the on and off periods are exponentially distributed, the respective probabilities are $(N-i)\lambda\Delta t$ and $i\Delta t$. Compound events have probabilities $O(\Delta t^2)$. The probability of no change is $[1 - ((N-i)\lambda + i)\Delta t] + O(\Delta t^2)$.

Let $P_i(t, x)$, $0 \leq i \leq N$, $t \geq 0$, $x \geq 0$, be the probability that at time t , i sources are on and the buffer content does not exceed x . Now,

$$\begin{aligned} P_i(t + \Delta t, x) = & (N - (i - 1))\alpha\Delta t P_{i-1}(t + \Delta t, x) + (i + 1)\beta\Delta t P_{i+1}(t + \Delta t, x) \\ & + [1 - ((N - i)\lambda + i)\Delta t]P_i(t, x - (i\lambda - c)\Delta t) + O(\Delta t^2) \end{aligned} \quad (A-1)$$

Passing to the limit $\Delta t \rightarrow 0$:

$$\frac{\partial P_i}{\partial t} + (i - c)\frac{\partial P_i}{\partial x} = (N - (i - 1))\lambda P_{i-1} + (i + 1)P_{i+1} + ((N - i)\lambda + i)P_i \quad (A-2)$$

Since we are interested only in time-independent equilibrium probabilities, we define $F_i(x)$ to be the equilibrium probability that i sources are on and buffer content does not exceed x . Therefore, we set $\frac{\partial P_i}{\partial t} = 0$ and obtain, for $i \in [0, N]$,

$$(i - c)\frac{dF_i}{dx} = (N - (i - 1))\lambda F_{i-1} + (i + 1)F_{i+1} + ((N - i)\lambda + i)F_i \quad (A-3)$$

where the understanding is that $F_i = 0$ if i is not in the stated interval. In the matrix notation,

$$\mathbf{D} \frac{d}{dx} \mathbf{F}(x) = \mathbf{M} \mathbf{F}(x), \quad x \geq 0 \quad (A-4)$$

where $\mathbf{D} = \text{diag}\{-c, 1-c, 2-c, \dots, N-c\}$ is a diagonal matrix and $\mathbf{M}$ is a tri-diagonal matrix in the following format,

$$M = \begin{bmatrix} -N\lambda & 1 & & & & \\ N\lambda & -\{(N-1)\lambda+1\} & 2 & & & \\ & (N-1)\lambda & -\{(N-2)\lambda+2\} & 3 & & \\ & \vdots & \vdots & \vdots & & \\ & & & 2\lambda-\{\lambda+(N-1)\} & N & \\ & & & \lambda & -N & \end{bmatrix}$$

Appendix B: Bit Error Rate (BER)

The performance of a lightwave system is judged by the bit-error rate (BER), defined as the probability of incorrect identification of a bit by the decision circuit of the receiver. For example, a BER of 2×10^{-9} corresponds to two errors per billion bits on average. The BER depends on the signal-to-noise ratio (SNR) of the current generated at the receiver when an optical bit stream is converted into the electric domain. The SNR, in turn, depends on various noise mechanisms associated with the received signal.

The fluctuating electric signal at the receiver is passed to the decision circuit, which samples it periodically at the bit rate to determine individual bits. The sampled value I fluctuates from bit to bit around an average value I_1 or I_0 , depending on whether the bit corresponds to 1 or 0 in the bit stream. The decision circuit compares the sampled value with a threshold value I_D and calls it bit 1 if $I > I_D$ and bit 0 if $I < I_D$. An error occurs if otherwise. Both sources of errors can be included through the error probability, defined as

$$BER = p(1)P(0|1) + p(0)P(1|0) \quad (B-1)$$

where $p(1)$ and $p(0)$ are the probabilities of receiving bits 1 and 0, respectively, $P(0|1)$ is the conditional probability of deciding 0 when a 1 bit is received, and $P(1|0)$ is the conditional probability of deciding 1 when a 0 bit is received. Since 1 and 0 bits are equally likely to occur in a realistic bit stream, $p(1) = p(0) = 1/2$, and the BER becomes

$$BER = \frac{1}{2} [P(0|1) + P(1|0)] \quad (B-2)$$

The quantities $P(0|1)$ and $P(1|0)$ depend on the probability density function $p(I)$ of the sampled value I . The functional form of $p(I)$ depends on the statistics of various noise sources responsible for current fluctuations. It is not easy to find an analytical form of $p(I)$. For this reason, it is common to assume that $I(t)$ is a Gaussian random process. However, both the average and the variance of the Gaussian distribution are taken to be different for 1 and 0 bits such that,

$$p_{\kappa}(I) = \frac{1}{\sigma_{\kappa}\sqrt{2\pi}} \exp\left[-\frac{(I - I_{\kappa})^2}{2\sigma_{\kappa}^2}\right] \quad (B-3)$$

where $m=0$ or 1 , I_1 and I_0 represent average currents for 1 and 0 bits, and σ_1 and σ_0 are the corresponding variances. If I_D is the decision threshold, the probability of an error occurring will be in forms of conditional probabilities, written as

$$P(0|1) = \frac{1}{\sigma_1\sqrt{2\pi}} \int_{-\infty}^{I_D} \exp\left[-\frac{(I - I_1)^2}{2\sigma_1^2}\right] dI = \frac{1}{2} \operatorname{erfc}\left(\frac{I_1 - I_D}{\sigma_1\sqrt{2}}\right) \quad (B-4)$$

$$P(1|0) = \frac{1}{\sigma_0\sqrt{2\pi}} \int_{I_D}^{\infty} \exp\left[-\frac{(I - I_0)^2}{2\sigma_0^2}\right] dI = \frac{1}{2} \operatorname{erfc}\left(\frac{I_D - I_0}{\sigma_0\sqrt{2}}\right) \quad (B-5)$$

where erfc stands for the complementary error function defined as

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-y^2) dy \quad (B-6)$$

Substituting Equations (B-4) and (B-5) in Equation (B-2), the BER is given by

$$\operatorname{BER} = \frac{1}{4} \left[\operatorname{erfc}\left(\frac{I_1 - I_D}{\sigma_1\sqrt{2}}\right) + \operatorname{erfc}\left(\frac{I_D - I_0}{\sigma_0\sqrt{2}}\right) \right] \quad (B-7)$$

Equation (B-7) shows that the BER depends on the decision threshold I_D of the receiver.

In practice, I_D is optimized to minimize the BER. The minimum occurs when I_D is chosen

such that $\frac{I_1 - I_D}{\sigma_1} = \frac{I_D - I_0}{\sigma_0} = Q$ (Besides BER, Q is an alternative parameter to

represent the QoS of a signal), which depends on the noise level and given as,

$$I_D = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1} \quad (B-8)$$

When the noise level is the same for all bits, we can set $\sigma_0 = \sigma_1$ in Equation (B-8). Since

$I_D = \frac{1}{2}(I_1 + I_0)$ in that case, the decision threshold should be in the middle of the average

currents for 0 and 1 bits. This is the situation for most receivers when noise is dominated by the thermal noise.

The BER with the optimum setting of the decision threshold is obtained using Equation (B-8) and is given by the following simple expression:

$$BER = \frac{1}{2} \operatorname{erfc}\left(\frac{Q}{\sqrt{2}}\right) \approx \frac{\exp(-Q^2/2)}{Q\sqrt{2\pi}} \quad (B-9)$$

where the approximate form is reasonably accurate for $Q > 3$. The Q parameter is found from Equation (B – 8) to be

$$Q = \frac{I_1 - I_0}{\sigma_0 + \sigma_1} \quad (B-10)$$

This parameter plays an important role in the system design. The BER improves as Q increases and is about 10^{-9} when $Q=6$. It becomes lower than 10^{-12} for $Q=7$.