Contention resolutions in optical packet switches

Li, Yi

2007

Li, Y. (2007). Contention resolutions in optical packet switches. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/46900>

<https://doi.org/10.32657/10356/46900>

Nanyang Technological University

Downloaded on 26 Apr 2023 21:43:20 SGT

at78352

Contention Resolutions in Optical Packet Switches

Li Yi

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

2007

TK
5105.3
L693
2007

Statement of Originality

I hereby certify that the content of this thesis is the result of original research done by me and has not been submitted for a higher degree to any other university or institute.

Date

Li Yi

*Dedicated
to
My Beloved Parents
And
Brother*

Acknowledgement

I would like to greatly thank my supervisor, Asst. Prof. Gaoxi Xiao, for his continuous support, instructive guidance, and inspiring encouragement through the course of my postgraduate study. During my PhD program in School of Electrical and Electronics Engineering, Nanyang Technological University, his boundless suggestions, dedication to excellence and great patience always help me in the difficult time.

I also wish to express my sincere gratitude to Dr. Hooshang Ghafouri-Shiraz for his hearty instruction and guidance during the first year of my research work. He has given me many suggestions for this research topic and great efforts in providing reviews and comments for my papers through the whole research program.

I am grateful to Dr. Wende Zhong, Dr. Liren Zhang, Dr. Gang Feng, Dr. Yang Qin for the instructive discussions and comments on the research topics.

I thank Dr. Ping Shum, director of Network Technology Research Center, and all the members of this center for the wonderful time together and great help from them.

I would like to extend my appreciation to my fellow students and friends during the four years in NTU for their cooperation and friendship.

Finally, I dedicate this thesis to my family, who has been always giving me love and confidence.

Contents

Acknowledgement iii

Contents.....iv

Abstractvii

List of Abbreviations and Symbols..... viii

List of Figuresxi

List of Tables.....xvi

Chapter 1: Introduction 1

 1.1 Motivations..... 1

 1.2 Contributions of the Thesis4

 1.3 Thesis Outline 7

Chapter 2: Literature Survey.....9

 2.1 Basic Functions of OPS..... 10

 2.1.1 Packet Coding 11

 2.1.2 Packet Synchronization 12

 2.1.3 Packet Routing Control..... 12

 2.1.4 Packet Header Recognition and Replacement 13

 2.2 Major Contention Resolution Technologies 13

 2.2.1 Optical Buffering 13

 2.2.2 Wavelength Conversion..... 17

 2.2.3 Deflection Routing..... 20

 2.3 OPS Node Architectures 21

 2.3.1 Staggering Switch..... 21

 2.3.2 Broadcast-and-Selected Switch 22

 2.3.3 Wavelength Routed Switch 23

 2.3.4 Bufferless Switch..... 25

 2.3.5 Combinational Contention-Resolution Switch 26

 2.4 Queuing Analysis of Packet Switches..... 27

 2.4.1 Queuing Analysis of the Classic Electronic Switches 28

2.4.2	Queuing Analysis in Optical Packet Switches	34
2.5	Scheduling Algorithms	35
2.6	Summary	37
Chapter 3: Contention Resolution in Resource-Sharing OPS		38
3.1	Introduction	39
3.2	Resource-Sharing OPS	40
3.2.1	Resource-Sharing Model and Implementation	40
3.2.2	Scheduling Algorithms	42
3.3	Simulation Platform	43
3.4	Performance Evaluation	46
3.4.1	Evaluations of Scheduling Algorithm 1	47
3.4.2	Evaluations of Scheduling Algorithm 2	53
3.4.3	Comparisons between the Scheduling Algorithms 1 and 2	54
3.5	Conclusion	56
Chapter 4: Fixed Wavelength Conversion in OPS Node		57
4.1	Introduction	58
4.2	Switch Models and Algorithm	59
4.3	Performance Evaluations and Discussions	62
4.3.1	Evaluations under Non-Bursty Traffic Load	63
4.3.2	Evaluations under Bursty Traffic Load	68
4.3.3	Discussions on the Limited Wavelength Conversion	72
4.4	Conclusion	74
Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches		76
5.1	Introduction	77
5.2	MOPS Node Model	78
5.3	Analytical Model of Packet Loss in MOPS	81
A.	Analysis of the Pure MOPS Model	82
B.	Analysis of MOPS with Shared FDLs	83
C.	Analysis of MOPS with SWCs	87
D.	Analysis of MOPS with both FDLs and SWCs	89
5.4	Numerical Results and Discussions	94
5.4.1	Single-Board MOPS	94

5.4.2 Multi-Board MOPS	102
5.5 Conclusion.....	109
Chapter 6: Node Configurations in Multi-fiber OPS Networks	110
6.1 Introduction	111
6.2 Node Configurations	113
6.2.1 Different Connection Schemes: Main Idea.....	114
6.2.2 Different Connection Schemes: Formal Definitions	117
6.3 Evaluations of Different Connection Schemes	120
6.3.1 Effects of Different TSB Schemes	121
6.3.2 Effects of the Different FS3 Schemes.....	123
6.4 Schemes for Improving Switch Performance.....	127
6.5 Conclusion.....	136
Chapter 7: Traffic Load Allocations in Optical Packet Switches	138
7.1 Introduction	139
7.2 Analysis of the Most Fundamental OPS Configuration.....	140
7.3 The Load Balancing (LB) Algorithm.....	147
7.4 Simulation Results and Discussions.....	155
7.4.1 Packet Loss in the Most Fundamental OPS Configuration	155
7.4.2 Packet Loss Evaluation of Other OPS Configurations	158
7.4.3 Load Balancing Algorithm in the Most Fundamental Node Configuration	163
7.4.4 Load Balancing Algorithm in Other OPS Configurations.....	164
7.5 Conclusions	169
Chapter 8: Conclusions and Future Work	170
8.1 Conclusions	170
8.2 Future Work	172
Author's Publications	174
Bibliography	176

Abstract

This project studies the designing and evaluates the performance of optical packet switches (OPS) with a central focus on various contention resolution schemes. In this project, we propose and evaluate several cost-effective contention resolution solutions by exploring the time, wavelength, and space dimensions of wavelength division multiplexing (WDM) networks, respectively. Specifically, node configurations with both the shared fiber delay lines (FDLs) and the tunable wavelength converters (TWCs) are evaluated to reveal the cost-effective manner of installing and utilizing such contention resolution components. We then propose a node configuration with the fixed wavelength conversion (FWC) and show that it performs comparably to the much more expensive configuration with full-range tunable wavelength conversion (FTWC); and it outperforms the conventional limited wavelength conversion (LWC) solution as well. To explore the space dimension in multi-fiber links, we propose the multi-fiber optical packet switch (MOPS) scheme. Analytical and simulation results show that having multiple fibers per link significantly save the TWCs and FDLs needed to achieve satisfactory performance. It is also shown that by carefully configuring the switching-node architecture, multi-board switches can achieve good performance at a much lower cost than that of the ideal, strictly non-blocking switch. Finally, we investigate the impacts of different traffic allocations to the packet-loss performance. We demonstrate that the balanced scheme, which evenly distribute traffic load in time and/or space dimensions as far as possible, leads to better performance. And we propose an algorithm to achieve the best performance subject to limited capacity of traffic load adjustment.

List of Abbreviations and Symbols

Abbreviations

ATM	Asynchronous Transfer Mode
AWG	Arrayed Waveguide Grating
BER	Bit Error Rate
CCF	Critical Cell First
CIOQ	Combined Input-Output Queues
FDL	Fiber Delay Line
FIFO	First-In-First-Out
FSB	FDL Switching Board
FTWC	Full-range Tunable Wavelength Converter
FWC	Fixed Wavelength Converters
FWM	Four-Wave Mixing
GPS	Generalized Process Sharing
HOL	Head-of-Line
IP	Internet Protocol
LB	Load Balancing
LWC	Limited Wavelength Conversion
MOPS	Multi-fiber Optical Packet Switch
OBS	Optical Burst Switching

OCDM	Optical Code-Division Multiplexing
O/E/O	Optical-Electronic-Optical
OPS	Optical Packet Switch/Switching
OQ	Output Queuing
OTN	Optical Transport Network
OXC	Optical Cross-Connects
PIFO	Push-In-First-Out
PL	Packet Loss
PLR	Packet Loss Rate
QoS	Quality of Service
RAM	Random Access Memory
RWA	Routing-and-Wavelength-Assignment
SCM	Subcarrier Multiplexing
SOA	Semiconductor Optical Amplifier
TSB	Transmission-Switching Board
TWC	Tunable Wavelength Converter
VOQ	Virtual Output Queue
WDM	Wavelength Division Multiplexing
WRR	Weighted Round Robin
XGM	Cross-Gain Modulation
XPM	Cross-Phase Modulation

Symbols

N	input/output links to the switching node
-----	--

W	wavelengths per fiber
F	fibers per link
L	number of FDLs
R	traffic load ratio
N_{FDL}	number of FDLs
N_{TWC}	number of tunable wavelength converters
L_b	burst length
a	incoming traffic load per wavelength channel
ρ	probability of arriving at one output link
PL	packet loss
PLR	packet loss rate
N_{TSB}	number of transmission-switching board
N_{FSB}	number of FDL switching board

List of Figures

Figure 2.1: A general structure of optical packet switching node10

Figure 2.2: Generic processing of an optical packet in OPS11

Figure 2.3: Optical buffering by FDLs14

Figure 2.4: Optical buffering for OPS16

Figure 2.5: Multi-stage buffers with large buffer size16

Figure 2.6: An example of contention resolution by wavelength conversion17

Figure 2.7: An example of cross-gain modulation (XGM)18

Figure 2.8: An example of cross-phase modulation (XPM).....19

Figure 2.9: Contention resolutions by wavelength conversion20

Figure 2.10: An example of deflection routing scheme for contention resolution.....21

Figure 2.11: Architecture of a Staggering Switch22

Figure 2.12: The OPS architecture in KEOPS project23

Figure 2.13: A feed-forward AWG-based OPS architecture24

Figure 2.14: Architecture of the bufferless OPS node with shared TWCs.....25

Figure 2.15: Example switch with combinational contention resolutions26

Figure 2.16: A switch with VOQ and shared output queues32

Figure 3.1: Resource-sharing model of an optical packet switch41

Figure 3.2: Inner implementation of optical components connection42

Figure 3.3: Flow chart of the simulation process platform.....45

Figure 3.4: Performance of resource-sharing OPS model with TWCs and FDLs under uniform traffic load with Scheduling Algorithm 1	48
Figure 3.5: Performance of resource-sharing OPS model with TWCs and FDLs under non-uniform traffic load with Scheduling Algorithm 1	49
Figure 3.6: The percentage of lost packets due to lack of TWCs under uniform traffic load with Scheduling Algorithm 1	50
Figure 3.7: The percentage of lost packets due to lack of TWCs under non-uniform traffic load with Scheduling Algorithm 1	50
Figure 3.8: Performance improvement by using shared FDLs under uniform traffic load with Scheduling Algorithm 1	52
Figure 3.9: Performance improvement by using shared FDLs under non-uniform traffic load with Scheduling Algorithm 1	52
Figure 3.10: Performance of resource-sharing OPS model with TWCs and FDLs under uniform traffic load with Scheduling Algorithm 2	53
Figure 3.11: Performance of resource-sharing OPS model with TWCs and FDLs under non-uniform traffic load with Scheduling Algorithm 2	54
Figure 4.1: Conversion ranges of FTWC and FWC	59
Figure 4.2: OPS models with different types of wavelength conversions	60
Figure 4.3: The OPS node with shared FWCs and FDLs	62
Figure 4.4: Performance comparisons between the FWC and the FTWC schemes under uniform and non-bursty traffic with no FDLs	64
Figure 4.5: Performance comparisons between FWC and FTWC schemes under non-uniform and non-bursty traffic with no FDLs ($R = 1.2$)	65

Figure 4.6: Performance comparisons between the FWC scheme and the FTWC scheme under uniform and non-bursty traffic with FDLs	66
Figure 4.7: Performance comparisons between the FWC scheme and the FTWC scheme under non-uniform and non-bursty traffic with FDLs ($R = 1.2$) ..	67
Figure 4.8: Performance of FTWC scheme with different number of FTWCs under uniform and non-bursty traffic	68
Figure 4.9: Performance comparisons between FWC and FTWC schemes with FDLs under bursty traffic with uniform distribution	69
Figure 4.10: Performance comparisons between FWC and FTWC schemes with no FDLs under bursty traffic with non-uniform distribution ($R = 1.2$)	70
Figure 4.11: Performance comparisons between FWC and FTWC schemes with FDLs under bursty traffic with uniform distribution	71
Figure 4.12: Performance comparisons between FWC and FTWC schemes with FDLs under bursty traffic with non-uniform distribution ($R = 1.2$)	71
Figure 4.13: Performance comparisons between FWC and LWC schemes under non-bursty traffic, uniform traffic pattern	73
Figure 4.14: Performance comparisons between FWC and input LWC schemes under bursty traffic load ($L_b=10$) with the non-uniform traffic pattern ($R=1.2$)..	75
Figure 5.1: General node architecture of single-board MOPS.....	80
Figure 5.2: Node architecture of multi-board MOPS.....	81
Figure 5.3: The “equivalent” two-stage switch configuration of which the packet-loss performance closely resembles that in MOPS with both FDLs and SWCs	90
Figure 5.4: PLRs of single-board pure MOPS under uniform traffic load	95

Figure 5.5: PLRs of single-board pure MOPS under non-uniform, geometrically-distributed traffic load where $R = 1.2$	96
Figure 5.6: PLRs of single-board MOPS with FDLs under uniform traffic load	97
Figure 5.7: PLRs of single-board MOPS with FDLs under non-uniform traffic load where $R = 1.2$	97
Figure 5.8: PLRs of MOPS with SWCs under uniform traffic load.....	98
Figure 5.9: PLRs of MOPS with different number of SWCs	99
Figure 5.10: PRLs of MOPS with SWCs under non-uniform, geometrically-distributed traffic load where $R = 1.2$	100
Figure 5.11: PLRs of MOPS with both FDLs and SWCs under uniform traffic load..	101
Figure 5.12: PLRs of multi-board MOPS with different switching-board configurations	103
Figure 5.13: Average buffer size per fiber vs. number of fibers per link	104
Figure 5.14: Performance of MOPS with FDLs under very high traffic load	105
Figure 5.15: Performance of MOPS with a fixed total number of FDLs	106
Figure 5.16: PLRs of multi-board MOPS with fixed switching-board sizes	108
Figure 6.1: A schematic representation of the ideal single-board switch.....	113
Figure 6.2: A schematic representation of the multi-board switch.....	116
Figure 6.3: Examples of the two different TSB schemes	118
Figure 6.4: Comparisons between the two different TSB schemes (without FDLs)...	122
Figure 6.5: Comparisons between the two different TSB connection schemes (with FDLs).....	124
Figure 6.6: Comparisons between the different multi-board node configurations where FDLs are shared and not shared among different TSBs respectively	125

Figure 6.7: Comparisons between the different FSB schemes	126
Figure 6.8: Comparisons between the multi-board switches and the ideal single-board switch	129
Figure 6.9: Effects of increasing the number of fibers per link, where $M^T = M^F = 8$ and $M_2^T = 2$	130
Figure 6.10: Effects of increasing TSB size, where $M^F = 8$ and $M_2^T = 2$	131
Figure 6.11: Effects of increasing FSB size, where $M^T = 8$ and $M_2^T = 2$	132
Figure 6.12: Effects of increasing M_2^T , where $M^T = M^F = 8$	134
Figure 6.13: Trade-off between the packet-loss performance and the switch complexity ($M^T \times M^T$) with increasing values of M_2^T	135
Figure 7.1: The most fundamental OPS configuration	141
Figure 7.2: Performance of the most fundamental OPS under various load distributions in time and space dimensions	157
Figure 7.3: Typical OPS configuration with SWCs and FDLs	158
Figure 7.4: Packet-loss performance in the OPS with FDLs but no SWC	160
Figure 7.5: Packet-loss performance in the OPS with SWCs but no FDL	161
Figure 7.6: Packet-loss performance in the OPS with both SWCs and FDLs	162
Figure 7.7: Performance of the LB algorithm in the most fundamental OPS configuration	165
Figure 7.8: Performance of the LB algorithm in OPS with only FDLs ($N_{FDL} = 2$) ...	166
Figure 7.9: Performance of the LB algorithm in OPS with only SWCs. ($N_{SWC} = 2$) ...	167
Figure 7.10: Performance of the LB algorithm in OPS with both FDLs and SWCs ($N_{FDL} = 2, N_{SWC} = 2$)	168

List of Tables

Table 3.1: Comparison of PLR between the two scheduling algorithms55

Table 5.1: The computation time needed by the two analytical models respectively..100

Chapter 1

Introduction

With high-capacity optical fibers encircling the world in terrestrial and undersea installations, optical networks emerge as a major platform for data communications. Unlike the early-day systems containing only a few spans of fibers, a light source and a photo-detector, today's optical networks are enjoying powerful and flexible functions enabled by various passive and active devices [1]. To meet the need of handling exponentially growing Internet traffic and services, future optical networks are expected to become increasingly "optical", where network functions, notably routing and switching of end-to-end communications paths, and eventually routing and switching of data packets, are executed in optical domain as much as possible [2]-[4].

1.1 Motivations

The most significant development of optical communications in the past decade is probably the wide deployments of wavelength division multiplexing (WDM) networks [5], [6]. By supporting a number of parallel transmissions on different wavelengths in

Chapter 1: Introduction

the same fiber, WDM technology provides an effective solution for utilizing optical bandwidth. Currently most WDM networks are providing circuit-switched connections (termed as *lightpath*) between source-destination nodes. By doing so, data signals can remain in optical domain until they either reach the destination, or reach the point where regeneration is necessary (e.g., to maintain a satisfactory bit error rate (BER)). With the development of optical Internet, however, it is expected that more flexible and efficient traffic handling methods will be needed, while the most promising candidates include *optical packet switching* (OPS) and *optical burst switching* (OBS).

Optical packet switching can be viewed as a straightforward extension of electronic packet switching into optical domain: traffic load is handled in the unit of data packet; being stored and forwarded at each intermediate node. However, compared to electronic networks, OPS networks have much larger bandwidth (and therefore much higher arrival rates of data packets which require high-speed processing), much smaller buffer size, and costly optical-electronic-optical (O/E/O) conversions in foreseeable near future. The flexibility and bandwidth efficiency of OPS network can hardly be fully explored until these shortages are largely overcome.

A “trade-off” solution with less flexibility yet (arguably) simpler control than OPS is OBS, which handles traffic load in the unit of *burst*. Roughly speaking, a burst is a “jumbo package” containing a large number of data packets (yet not too large to be worthy of establishing a lightpath as that in circuit-switched network) [7]-[14]. By handling the package rather than each individual packet in it, processing operations on intermediate nodes typically become easier. In OBS networks, bandwidth is still reserved, however, only in a one-way process. In other words, a burst may be sent out

Chapter 1: Introduction

without waiting for the acknowledgment that the reservation has been successful. Another character of OBS is that a burst can cut through an intermediate node (or a switch) without being processed when an out-of-band control signal has reserved the bandwidth for it in advance.

Both OPS and OBS have been hot topics for quite a few years. While OBS is now in early-stage implementations in labs, OPS, despite of all the challenges it is still facing, is regarded as “tomorrow’s technology”. In fact, people are expecting that OPS will fit in a realistic network scenario where circuit and packet switching are integrated into the same optical transport network (OTN), such that the bandwidth efficiency of OPS can stand side by side with the control efficiency of optical circuit switching [15]. For example, OPS may be applied at network edge, functioning as an interface between electronic and optical domains, while optical circuit switching are carrying large-chunk transmissions in backbone networks [16]. To make these two technologies work together effectively, OPS may need an aggregation mechanism that can accept various data (e.g., IP packets and ATM cells, etc.) from a number of sources and map them into optical packets. It also needs to map optical packets into appropriate wavelengths for transport over OTN. To achieve that, OPS needs a protocol capable of discovering the optical cross-connects (OXC) network topology, such that the aggregation function collaborates smoothly with the quality of service (QoS) provisioning in OTN [17]-[21].

A crucial problem that has to be solved in any packet switch is *packet contention*, where there are two or more packets destined for the same output port simultaneously. Due to high-speed requirement and constrained by today’s optical technology,

Chapter 1: Introduction

contention resolution in OPS is much more difficult than that in an electronic switch.

Existing solutions can be divided into three different classes as follows:

1. Optical buffering solution [22]-[25], where some contending packets are temporarily buffered.
2. Wavelength conversion solution [26]-[30], where contending packets are converted to other wavelengths in the output link, when applicable.
3. Deflection routing solution [31], [32], where contending packets are switched to another output link other than the originally destined one. It leaves to the downstream nodes to properly switch the deflected packets to the destination.

Further details of these existing solutions and their respective pros and cons will be discussed later in Chapter 2.

We are motivated to study contention resolutions in OPS networks. The main purposes include: (i) evaluating the efficiency of combining some existing solutions; (ii) achieving better understanding and more accurate analysis of packet contention in OPS network; and (iii) proposing some new cost-effective contention resolution schemes.

1.2 Contributions of the Thesis

The main contributions of our project can be summarized as follows:

1. Evaluating the efficiency of different combinations of existing solutions. Specifically, we tested the effectiveness of using buffering components and wavelength conversion components respectively, under different traffic loads. In addition to that, two different scheduling algorithms were investigated where

Chapter 1: Introduction

tunable wavelength converters (TWCs) and fiber delay lines (FDLs) are utilized in different priorities. Backed by extensive simulation results, we proposed some guidelines for achieving better cost-effectiveness in OPS design. Early-stage work of this part has been published in Papers 1 and 2 as listed in the Author's Publications on page 174, followed by a more comprehensive study reported in Paper 3. In this thesis, the study is presented in Chapter 3.

2. Improving the cost-effectiveness of the wavelength conversion solutions. We investigated the performance of OPS with fixed wavelength converters (FWC) instead of the highly expensive full-range wavelength converters. We find that nearly the same performance can be achieved in these two different cases. In further studies, we compared FWCs with limited wavelength converters (LWCs). It is shown that FWC with a full input-wavelength range and a single output wavelength outperforms LWC with halved input and output wavelength ranges. These results have been published in Paper 4 as listed in the Author's Publications. A more comprehensive report can be found in Chapter 4 of this thesis.
3. Studying contention resolutions in multi-fiber WDM networks. Our contributions are two-fold:
 - a. For the first time to the best of our knowledge, we provided a detailed analysis of various contention resolution schemes in multi-fiber optical packet-switched (MOPS) networks. We show that it helps to greatly improve network performance and lower network cost by adopting multi-fiber solutions rather than the traditional single-fiber solutions. In addition, by carefully evaluating the efficiencies of different contention resolution components under different

Chapter 1: Introduction

traffic loads, we proposed some guidelines for achieving better cost-effectiveness in MOPS networks. The study was firstly reported in Paper 5, while further details were later published in Paper 7¹. In this thesis, this part of study is revealed in Chapter 5.

- b. We studied the effects of using multi-board switch instead of the strictly non-blocking single-board switch. Several different node configurations were proposed and compared with each other. Extensive simulations have been conducted to evaluate how the packet-loss performance of switch is affected by switching-board size, buffer volume, and different buffer-space sharing methods, respectively. We show that different from that in circuit-switched networks, multi-board OPS cannot perform comparably to the ideal, single-board OPS by slightly increasing the number of fibers per link. Among all the possible alternative solutions we have tested, the most efficient one turns out to be increasing the buffer size rather than enlarging the switching-board size or enhancing buffer sharing between different switching boards. Detailed results have been published in Paper 8 and can be found in Chapter 6 of this thesis.
4. Evaluating the effects of different traffic load allocations in OPS networks. By proving that the packet loss in a simple OPS model is a convex function of traffic load, we show that it generally helps to lower packet loss by evenly allocating traffic load as far as possible, and a strictly balanced allocation achieves the best performance. For more general node configurations equipped with FDLs and/or TWCs, extensive simulation results show that the conclusion remains valid. In

¹ A contention resolution scheme for time-critical applications in MOPS is described in Paper 6, which is not included in this thesis because it has a different objective of eliminating buffering delay at a rather high cost, rather than considering anything about cost-effectiveness.

Chapter 1: Introduction

In addition, we proposed a Load Balancing algorithm which minimizes packet loss subject to a given capability of shifting unbalanced traffic load, either in time or in space dimensions. These results have been included in Paper 9 which is currently under review. In this thesis, they are presented in Chapter 7.

1.3 Thesis Outline

The rest of the dissertation is organized as follows. A brief survey of optical packet switching technologies and node architectures is presented in Chapter 2, with a special focus on contention resolution techniques. In Chapter 3, we investigate the efficiency of different contention resolution devices, specifically wavelength converter and FDL, in a typical resource-sharing OPS node. Guidelines have been proposed for achieving more efficient utilizations of these components. Chapter 4 demonstrates that contention resolution by the fixed wavelength conversion can achieve nearly the same performance as that by full-range wavelength conversion though the former one is much cheaper. Multi-fiber OPS is analyzed in Chapter 5, where analytical models are developed for studying the simplest node configuration as well as those with shared FDLs and shared wavelength converters respectively. Chapter 6 evaluates the performance of different configurations of multi-board switch, while favorable schemes of inter-board connections are revealed by extensive simulations. Effects of traffic load allocations on packet-loss performance are studied in Chapter 7. By proving the convexity of packet-loss function in a simple multi-channel OPS model and conducting extensive simulations on some other more general node configurations, the optimality of uniform traffic allocation is proved. Furthermore, a Load Balancing algorithm is proposed to achieve best performance subject to a given capacity of shifting

Chapter 1: Introduction

unbalanced traffic load. Finally, Chapter 8 concludes the whole dissertation and briefly discusses some possible future research directions.

Chapter 2

Literature Survey

In this chapter, we present a brief survey of the current OPS technologies and recent research progress, focusing on contention resolution methods, OPS node architectures, classic queuing analysis and scheduling algorithms. Due to space limitations, physical-level details of various implementation and measurement technologies (e.g., measurements and controls of crosstalk and bit error rate, etc.) are largely omitted.

Theoretically speaking, OPS can handle fixed-length, synchronous data cells as well as variable-length, asynchronous data packets. Although directly handling asynchronous traffic helps to achieve better flexibility and capacity efficiency [33]-[36], it remains as a challenge to today's optical technologies. Specifically, with no synchronization, it is difficult to efficiently identify and read the packet header without converting the whole packet into electronic domain. Therefore, most existing results study the case with synchronous traffic [18], [23], [30], [37]-[47]. Throughout this chapter, unless otherwise specified, we assume that optical switches are handling synchronous data packets.

2.1 Basic Functions of OPS

The basic functions of an OPS generally include packet routing, space switching, optical buffering and optional header processing and translation, etc., typically handled by three sub-blocks namely input interface, switching matrix, and output interface, as shown in Figure 2.1. The input interface aligns the incoming packets for synchronization and reads packet header information before the packets enter the switching matrix. The switching matrix routes the packets to their output ports. Contention resolutions are also handled in this part. In output interface, new headers are inserted to outgoing packets. Regenerations and buffering of data packets can also be handled in this part if necessary.

Figure 2.1: A general structure of optical packet switching node.

An ideal optical packet switch should preserve end-to-end transparency (in terms of data rate, wavelength and format) throughout the routing process [48] by keeping data signals in optical domain from source to destination. The generic process of an optical packet is shown in Figure 2.2: the packet header is converted into electronic domain for processing, while data payload remains in optical domain. To make a practical optical

Chapter 2: Literature Survey

switch, the fundamental functions should at least include packet coding, packet synchronization, packet routing control, packet header recognition and replacement, and packet contention resolutions [9], [17]. We briefly introduce the first several functions in this section while detailed discussions on contention resolutions will be presented in the next section.

Figure 2.2: Generic processing of an optical packet in OPS [49].

2.1.1 Packet Coding

Packet coding defines how data payload and routing information are encoded in each data packet. There are basically three categories of packet coding methods: bit serial, bit parallel, and out-of-band signaling [9]. Both bit serial and bit parallel are the so-called *in-band signaling* methods, where data payload and routing information are transmitted on the same wavelength(s). The difference between them is that, in bit serial coding, data payload and routing information are transmitted in a serial sequence on a single wavelength, whereas in bit parallel coding, multiple bits are parallel transmitted on separate wavelengths [9]. In-band signaling can be implemented by using optical code-division multiplexing (OCDM), optical pulse interval, or mixed-rate techniques [17], [50], [51]. Specifically, an OCDM packet denotes its output port by its

Chapter 2: Literature Survey

code format, while the latter two techniques include the routing information in the packet headers. *Out-of-band* signaling puts the routing information on a different channel other than the data transmission wavelength. Therefore it is relatively easier for the switching nodes to retrieve routing information [9]. The main schemes include sub-carrier multiplexing (SCM), which carries routing information in a separate electrical sub-carrier [17], and dual-wavelength coding, which transmits packet headers on a separate optical wavelength [9].

2.1.2 Packet Synchronization

Synchronization is the process of aligning two optical pulses in time scale. In OPS networks, it refers to either the alignment of an incoming pulse stream and a locally available clock pulse stream, or the relative alignment of two incoming pulse streams. A general implementation of packet synchronization consists of a packet start recognizer, which identifies the packet start, and a programmable delay module, which adjusts the arrival time of the packets [52], [53]. Since packets enter a node from different links, they may be totally out of phase with each other. The input synchronization stage aligns packets before they enter the switch fabric, typically by using the fiber delay lines [38].

2.1.3 Packet Routing Control

Based on the routing information carried by the optical packets, either in-band or out-of-band, the switching matrix in each node directs the packets to their destined output ports. Optical packets are kept in optical domain as far as possible, while header processing, switch configuration and buffering control are executed in electronic

Chapter 2: Literature Survey

domain. The control unit has to efficiently supervise the operations of various optical components, e.g., splitters, wavelength converters and optical gates etc. In addition, high-speed calculations and control operations are needed for contention resolutions.

2.1.4 Packet Header Recognition and Replacement

The packet-header reading circuit recognizes the stream with a special bit pattern at the beginning of each packet. Detailed discussions on packet header recognition by different optical components, e.g., semiconductor optical amplifier (SOA) and fiber Bragg gratings, etc, require lengthy introductions to the physics background. Therefore they are omitted in this chapter. Interested readers could refer to [51], [54]-[56]. At the output port side, routing information may need to be updated where the old packet header is replaced with a new one. Wavelength converter, among the other solutions, can block the old header while inserting the new one.

2.2 Major Contention Resolution Technologies

Based on today's technologies, there are mainly three different methods for contention resolutions: optical buffering, wavelength conversion, and deflection routing. Hereafter we discuss them separately.

2.2.1 Optical Buffering

Optical buffers can temporarily hold some contending packets until a certain later time slot [22], [24], [57]-[59]. Since high-speed optical random access memories (RAMs) are not available, the only practical option at this moment is to use fiber delay lines (FDLs) [22], [24]. Specifically, contending packets are temporarily held by going

Chapter 2: Literature Survey

through a span of fibers attached to the switching node. The time that a packet can be held is decided by the length of FDLs and the times that the packet loops around the FDL, as illustrated in Figure 2.3. The case where each packet can go through an FDL at most once is called *traveling type* [60]; and the case where packets can loop around FDLs multiple times is called *recirculating type* [60]. In either case, the buffering time would be limited, severely constrained by optical attenuation and dispersion, etc. The number of FDLs is another important designing parameter: having more FDLs helps to hold more contending packets, yet a large number of FDLs increase the optical hardware volume, enlarge the switch size, and increase the noise level as well.

Figure 2.3: Optical buffering by FDLs.

Optical switches with buffers can be classified essentially according to the position of the buffers [61] as output buffered switch, input buffered switch and recirculation buffered switch [62] etc., as shown in Figure 2.4. Output buffered switches have optical

Chapter 2: Literature Survey

buffers installed on the output side. It is assumed that the switching matrix operates at a very high speed, such that multiple packets destined for the same output port could be transmitted in each time slot. Then all but one of these packets would be buffered on the output side (as long as buffer size permits). Theoretically speaking, output buffer achieves excellent performance; however, to construct a high-speed switch with large bundles of output buffers is highly difficult. Input buffered switch has buffers installed on the input side therefore high-speed switching matrix is no longer needed. Usually, its performance is much worse than output buffered switch, mainly because of the head-of-line (HOL) blocking [61], [63]. Specifically, a packet not at the head of buffering line could be blocked even if its destined output port is available. To fix the problem, various solutions have been proposed. The most well-known solutions include the virtual output queue (VOQ) where each input has separate logic buffers for different outputs [61], [64] or a combination of input-output queues [61], [63], [65], of which satisfactory performance can be achieved by a switching matrix operating faster than data transmission speed but not as fast as that in an output-buffered switch. In recirculation buffered switch, a number of FDLs are installed to loop from the output side of the switch back to its input side. Contending packets are therefore circulated back to the input side, and be handled altogether with newly arrived packets.

To overcome the problem that only a limited number of FDLs could be installed in each block of switching matrix, solutions have been proposed to cascade multiple small switches to form a bigger one (e.g., [36], [39], [57]), with a buffer depth of up to several thousands slots. A typical example is illustrated in Figure 2.5. The main drawbacks of such solutions, however, include the additional delays caused by FDLs, the difficulties in integrating a large-scale practical switching system, the complicated

Chapter 2: Literature Survey

control algorithm, and the degraded signal quality after going through multiple stages of switches, etc.

Figure 2.4: Optical buffering for OPS.

Figure 2.5: Multi-stage buffers with large buffer size [39].

Chapter 2: Literature Survey

2.2.2 Wavelength Conversion

Wavelength conversion is the technology of converting optical signals from one wavelength to another, which has been extensively studied in recent years [66]-[71]. By using wavelength conversion, contentions could be resolved by transferring the contending packets to other wavelengths in the output link, if applicable. A simple example is illustrated in Figure 2.6.

Figure 2.6: An example of contention resolution by wavelength conversion.

Wavelength conversion could be realized through optical-to-electronic-to-optical (O/E/O) conversion. In other words, the optical signal is converted into electronic domain which in turn drives a laser operating on a different wavelength. Such operations are straightforward yet highly expensive, especially in high-speed systems: high-speed transmitters/receivers have to be installed, together with the in-between high-speed electronic processing capabilities, for wavelength conversion purpose. To achieve all-optical wavelength conversion which keeps the signal in optical domain throughout the conversion procedure, most of the current solutions adopt nonlinear modulations of optical power gain or refractive index in semiconductor optical amplifiers (SOAs) [72], [73]. Specifically, there are three mainstream solutions, namely

Chapter 2: Literature Survey

cross-gain modulation (XGM) [72], [74], [75] cross-phase modulation (XPM) [76]-[79] and four-wave mixing (FWM) [72], [80]-[82] solutions, respectively.

XGM employs the interactions between two optical signals. Specifically, the gain in an SOA saturates when the optical power level increases to a certain level. Therefore, the amplifier gain can be modulated by the input signal, which in turn can be encoded (in inversion) on another wavelength traveling through the amplifier. An example is illustrated in Figure 2.7.

Figure 2.7: An example of cross-gain modulation (XGM).

XPM works in a somewhat similar way, only that it makes use of the change of refractive index in SOA controlled by optical modulation signals. As shown in Figure 2.8, input optical signal passes through one of the arms and modulates the phase of that arm. The interferometric nature of the device converts this phase modulation to an amplitude modulation of probe signal at another wavelength. Wavelength conversion at a speed up to 168 Gb/s has been achieved in a Symmetric-Mach-Zehnder (SMZ)-type switch [76].

Chapter 2: Literature Survey

Figure 2.8: An example of cross-phase modulation (XPM).

FWM technique makes use of the nonlinear interferences between optical signals on different wavelengths. The outcome is the generation of a signal at a frequency other than the interacting wavelengths. The intensity of the newly generated signal is proportional to the product of the intensities of interacting wavelengths. The phase and frequency of the generated conjugate wave is a linear combination of the interacting waves. FWM process preserves the signal format and amplitude information. It is so far the only kind of strictly transparent wavelength conversion. Wavelength conversion at 2.5 Gb/s over a wavelength range of 80nm, and 40 Gb/s over 24.6nm has been demonstrated in [72].

Despite of the enormous research efforts in the past years, wavelength conversion remains as difficult and expensive. Consequently it is still a very hot research topic. Important recent developments include the so-called tunable wavelength converters (TWCs) where output wavelengths of some (or all) input wavelengths can be dynamically changed upon request [83]-[86]. Wavelength conversion can resolve packet contentions by itself, as shown in Figure 2.9; or it can be jointly utilized with FDLs in the same optical switch. Such combined solutions generally lead to much

Chapter 2: Literature Survey

better efficiency in contention resolutions. Further details will be discussed later in Chapter 3.

Figure 2.9: Contention resolutions by wavelength conversion.

2.2.3 Deflection Routing

Deflection routing is to deflect the contending packets to outgoing links other than the originally destined ones [87]. An example is illustrated in Figure 2.10, where three contending packets are deflected to three different links, going through three different routes before they finally reach the destination node 4. By utilizing the idle capacities on the less-loaded links, deflection routing helps to reduce buffer depth and the number of optical gates. The main drawbacks include that some packets may have to go through rather long routes before they can finally reach the destination with a degraded signal quality; and network traffic control and management may also become more complicated. In addition, the effectiveness of this technique critically depends on network topology [32].

Figure 2.10: An example of deflection routing scheme for contention resolution.

2.3 OPS Node Architectures

We introduce several typical OPS node architectures, which exploited different contention resolution solutions at different stages through the developments of OPS technologies.

2.3.1 Staggering Switch

Staggering switch is a well-known early-stage design of OPS node configurations, in which contention resolutions are resolved by a set of FDLs of different lengths [23]. An example is illustrated in Figure 2.11, where FDLs are installed between two non-blocking switching stages.

In a staggering switch, each optical packet is split upon its arrival: a small faction of optical power is converted into electronic signal and forwarded to the control unit. Control unit controls the routing of the packet based on the information contained in the packet header. To resolve packet contentions, data packets can be distributed to different FDLs. Packets destined for the same output can be delivered into different

Chapter 2: Literature Survey

FDLs with different lengths, such that these packets would reach the switching stage at different time. Routing calculations of the control unit also have to make sure that there is no contention between newly arrived packets and the packets buffered in the previous time slots. Therefore, packet-scheduling algorithm based on routing information of all the delayed packets has to be carefully developed.

Figure 2.11: Architecture of a Staggering Switch [23].

2.3.2 Broadcast-and-Selected Switch

Another node configuration heavily depending on FDLs for contention resolutions is proposed in European ACTS KEOPS (Keys to Optical Packet Switching) project [42]. The original node architecture was proposed for single-transmission fiber systems rather than WDM networks. In this switch, wavelength converters are adopted for internal wavelength encoding, while optical passive couplers are installed side-by-side with FDLs for handling contention resolutions. Specifically, packets arriving at the same time are converted to different wavelengths. Signals on different wavelengths are then multiplexed and broadcast (via passive couplers) to the stage equipped with FDLs,

Chapter 2: Literature Survey

as illustrated in Figure 2.12. Similar to that in the staggering switch, different FDLs have different lengths. Buffered packets are then broadcast again to all the outgoing links and be selected by one (in unicast transmissions) or multiple (in multicast transmissions) of them. The selections are performed by optical gates. Finally, the selected signal is demultiplexed where data on a certain wavelength is chosen (again, by optical gates) to be transmitted to the outgoing link.

This project is also known for taking some other initiatives, e.g., in signal coding format. Further details can be found in [37], [42], [88]. The node architecture design can be extended to be adopted in WDM networks. Though the node architecture becomes more complicated, the main idea of contention resolutions nevertheless remains largely unchanged, where the related discussions can be found in [42].

Figure 2.12: The OPS architecture in KEOPS project [37].

2.3.3 Wavelength Routed Switch

To lower the power loss and the interferences between different transmissions, arrayed waveguide grating (AWG) is becoming a popular technology adopted in core

Chapter 2: Literature Survey

components of optical switches [38]. A typical AWG-based OPS architecture is proposed in WASPNET (Wavelength Switched Packet NETwork) project [40].

Figure 2.13: A feed-forward AWG-based OPS architecture [15].

Two different node architectures were proposed in this project. The first one is the so-called *feed-forward* AWG-based OPS architecture, as depicted in Figure 2.13. Specifically, incoming packets are converted to appropriate wavelengths, which facilitate the switching operations in the first AWG: in an AWG, the wavelength of each incoming packet determines its output port. Each output of the first AWG is connected through a tunable wavelength converter to a number of FDLs of different lengths installed in parallel between a demultiplexer and a multiplexer. Which FDL the signal will actually go through is, again, decided by the wavelength of the signal. After a proper delay, the signal would then go through another tunable wavelength converter and another AWG switch, before it finally reaches the proper outgoing link.

The other proposed node architecture is called *feedback* AWG-based OPS. The main idea remains largely the same as the feed-forward solution. The major difference lies in that, instead of having two AWG switches, the packets are looped back to go

Chapter 2: Literature Survey

through a single, larger-sized AWG switch. Further details can be found in [41]. Both of these two architectures can be extended to be applied in WDM networks [41].

2.3.4 Bufferless Switch

Figure 2.14: Architecture of the bufferless OPS node with shared TWCs [30].

With the developments of WDM technologies, wavelength conversion is playing an increasingly important role in OPS networks. In [89], a solution is proposed to achieve contention resolutions purely in wavelength dimension, literally without any optical buffer. The main problem, however, is that it needs a large number of TWCs, one for each outgoing wavelength channel. To save wavelength converter resources, a limited number of shared TWCs can be deployed, as illustrated in Figure 2.14 [30]. Packets

Chapter 2: Literature Survey

that do not need wavelength conversion are directly routed to output lines (e.g., the packet arriving on wavelength λ_i of input link 1); while the others are forwarded to the pool of TWCs (as long as there are still available TWCs).

This node architecture eliminates the bundles of FDLs. However, solely depending on shared TWCs in contention resolutions leads to at least two major drawbacks:

- the switching matrix is enlarged by a factor at least equaling to the number of TWCs; and
- additional attenuation and crosstalk are added to the signals, especially if they go through the switching matrix twice.

2.3.5 Combinational Contention-Resolution Switch

Figure 2.15: Example switch with combinational contention resolutions.

Recent developments have proposed to integrate FDLs and TWCs in the same node to achieve better performance [46], [90]-[92]. Meanwhile, deflection routing can also

Chapter 2: Literature Survey

be adopted, if necessary. An example node configuration with the combinational contention-resolution solution is illustrated in Figure 2.15. With the time-dimension resolution provided by FDLs, the space-dimension resolution by deflection routing, and the wavelength-dimension resolution by TWCs, packet contentions can be resolved more efficiently. In the next chapter, the performance of combinational solutions will be carefully evaluated.

To summarize, various contention resolution schemes have been implemented and evaluated. Early-stage developments depended mostly on optical buffers. The main problem is that the buffer size and the buffering time are both quite limited. Wavelength conversion proposes a different solution, where the main drawback, however, is that the conversion technology itself is still immature. Integrated utilizations of optical buffers and wavelength conversion are very likely to be the direction of the future, especially if wavelength conversion technology could be significantly improved. Deflection routing is not as popular as the other two solutions at least for the time being, since it brings nontrivial additional complexities to traffic control and may not be equally efficient in different network topologies. Overall, contention resolution technologies are still facing some technical bottlenecks, despite of all the significant progresses that have been made in recent years.

2.4 Queuing Analysis of Packet Switches

Queuing analyses in packet-switched networks generally aim at developing analytical models for calculating the packet loss rate (PLR) or the average packet delay. Accurate analytical models help simplify the calculations of PLR, especially under

Chapter 2: Literature Survey

light traffic load. And they help achieve better understandings of various factors affecting switch performance. In classic electronic packet switches, packets queue up in buffers, managed by a control unit running a scheduling algorithm. Optical packet switches basically emulate the functions of the electrical packet switches, though enhanced with some new features, e.g., those imposed by wavelength converters and/or circulate FDLs. In this section, following a review of the queuing analyses of electronic packet switches (e.g., [61]-[63]), we briefly discuss how to handle the special features of OPS.

2.4.1 Queuing Analysis of the Classic Electronic Switches

The classic electronic packet switches can be classified according to the positions of their buffers. Queuing analyses for different classes have some nontrivial differences. Hereafter we would discuss them separately. Unless otherwise specified, we assume that the switches operate in a synchronous mode with slotted time, and the buffers are first-in-first-out (FIFO).

Output Queuing

When buffers are implemented at output ports of a switch, packets destined for the same output port would queue up due to the limited output throughput. Consider a single output port. Defining the number of packets arriving at this output port in a given time slot as A , under uniform traffic load where every packet has an equal probability $1/N$ of destining for each output port, we have [61], [63]:

$$a_k = \Pr[A = k] = \binom{N}{k} \cdot \left(\frac{p}{N}\right)^k \cdot \left(1 - \frac{p}{N}\right)^{N-k}, \quad k = 0, 1, 2, \dots, N, \quad (2.1)$$

Chapter 2: Literature Survey

where p denotes the normalized traffic load to each output port. When $N \rightarrow \infty$, Eq. (2.1) becomes

$$a_k = \Pr[A = k] = \frac{p^k e^{-p}}{k!}, k = 0, 1, 2, \dots \quad (2.2)$$

Let Q_m denote the number of packets in the destined queue at the end of m -th time slot, A_m denote the number of packet arrivals during the m -th time slot, and b denote the buffer size. We have:

$$Q_m = \min\{\max(0, Q_{m-1} + A_m - 1), b\}. \quad (2.3)$$

For $N = \infty$ and $b = \infty$, the size of Q_m can be modeled as an $M/D/1$ queue. For finite N and b , Q_m can be modeled by a finite-state, discrete-time Markov chain with state transition probabilities $P_{ij} = \Pr[Q_m = j | Q_{m-1} = i]$ given by [63]:

$$P_{ij} = \begin{cases} a_0 + a_1 & i = 0, j = 0 \\ a_0 & 1 \leq i \leq b, j = i - 1 \\ a_{j-i+1} & 1 \leq j \leq b - 1, 0 \leq i \leq j \\ \sum_{m=j-i+1}^N a_m & j = b, 0 \leq i \leq j \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

Based on this very fundamental model, the steady-state queuing length can be obtained recursively from a number of Markov chain balance equations [61], [63]. Consequently, the buffer size required for a certain PLR can be calculated. To provide some benchmark values, under traffic load of $p = 0.8$, a buffer size of $b = 28$ is sufficient to keep the packet loss below 10^{-6} for any switch size N , whereas it takes a buffer depth of 110 to ensure packet loss of 10^{-11} under a traffic load of 0.9 [63], [64].

Chapter 2: Literature Survey

Output queuing can achieve the best throughput-delay performance since packets are delayed only by limited outgoing link capacity and never by limited switch capacity. For such case, the mean waiting time $\bar{W}$ can be easily calculated with Little's result [61]. For example, when $N \rightarrow \infty$ and $b \rightarrow \infty$, we have from the M/D/1 model that [61]:

$$\bar{W} = \frac{p}{2 - (1 - p)} \quad (2.5)$$

Input Queuing

To evaluate the performance of the input queuing, typically it is assumed that only the packets at the head of the buffers (i.e., the head-of-line, HOL, packets) can contend for their destined output ports. When multiple HOL packets are destined for the same output port, only one of them can be sent out. It is well-known that input-queued switches are impractical because of head-of-line blocking [61], [63]. In fact, it is proven that the maximum throughput is only about 58.6% [63].

With the assumptions of Bernoulli arrival and uniform destination distribution of each incoming packet, the average waiting time of each packet in an FIFO input queue can be obtained as [61]:

$$\bar{W} = \frac{p\overline{S(S-1)}}{2(1-p\overline{S})} + \bar{S} - 1 \quad (2.6)$$

where S is the average service time obtained from M/D/1 model [61]. The model reveals that, when the normalized traffic load approaches 0.586, the average waiting time of each packet will be exponentially increased to approach infinity.

Chapter 2: Literature Survey

Virtual Output Queuing & Combined Input-Output Queuing

As mentioned above, switches with input queues cannot perform as efficiently as those with output queues because of the HOL blocking problem. Such problem can be fixed by adopting proper buffering strategy [64], where the most successful ones include the virtual output queue (VOQ) [61]. In a switch with VOQ, as shown in Figure 2.16, every input is equipped with N logic buffers, each of which containing packets destined for a different output port as an independent FIFO queue (though they may physically share the same memory). By doing so, the HOL blocking is eliminated. It is shown VOQ can achieve 100% throughput under uniform and non-uniform traffic load [93]. The major challenge is that such a switch requires fast and intelligent arbitration mechanisms to arbitrate a total of N^2 queues in each time slot.

For the VOQ model in [93], the performance of average packet delay degrades differently under low and heavy loads. Under a fixed low load, the queuing delay converges to a constant value independent of N . Under a fixed heavy load, however, the queues will behave similarly to that of an M/D/1 queue with an arrival rate λ/N and a deterministic service time of N . Therefore, under heavy traffic with Bernoulli arrivals, the delay can be approximately as:

$$\overline{W} = \frac{\lambda N}{2(1 - \lambda)} \quad (2.7)$$

which is *proportional* to N .

Chapter 2: Literature Survey

Figure 2.16: A switch with VOQ and shared output queues.

With speedup operations, buffers are needed on both the input and the output sides [65], which leads to the combined input-output queues (CIOQ) [65], [94], as shown in Figure 2.16. The output buffers in CIOQ not only enable distributed scheduling but also assure fairness across all VOQs [65]. With the speedup factor $S = 4$ or 5 , a CIOQ can easily achieve about 99% throughputs when arrivals are independently, identically distributed (i.i.d) at each input, and uniformly destined for all the outputs. With well designed packet scheduling algorithms, a CIOQ switch can behave identically to an output queuing switch with a speedup of two [65]. Specifically, important theoretical results for a CIOQ switch to emulate an output queuing (OQ) switch include:

- 1) An $N \times N$ CIOQ switch with a speedup of $2 - 1/N$ can exactly emulate a $N \times N$ FIFO-OQ switch;
- 2) Regardless of the incoming traffic pattern, a CIOQ switch that uses the Critical Cell First (CCF) [65] algorithm with a speedup of two exactly mimics a FIFO-OQ switch, as well as a OQ switch adheres to a push-in-first-out (PIFO) queuing policy;

Chapter 2: Literature Survey

- 3) A minimum speedup of $2 - 1/N$ is necessary for an $N \times N$ CIOQ switch operating with *any* algorithm.

Since CIOQ switches typically aim to emulate the OQ switches, the blocking analysis remains the same as that for output queue once the objective is achieved.

Shared-Buffer Queuing

The shared-buffer queue means that all the contending packets are stored in a common physical memory, shared by all inputs and/or outputs with no dedicated separations of the memory space to each or them. Such implementations ensure that fewer buffers are needed to achieve a given PLR. Such advantage is highly favorable especially in optical networks, since large-sized optical buffer remains as highly complicated and large volume.

A typical switch model with completely shared buffers has been studied in [63], where a total of (Nb) buffers are shared by all the input and output ports. These buffers increase the switch size from $N \times N$ to $N(b+1) \times N(b+1)$, as shown in [63]. Let Q_m^i denote the number of packets destined for output i in the buffer at the end of the m th time slot. Then $\sum_{i=1}^N Q_m^i$ is the total number of packets in the shared buffer at the end of the m th time slot. If the buffer size is infinite, we have

$$Q_m^i = \max\{0, Q_{m-1}^i + A_m^i - 1\} \quad (2.8)$$

where A_m^i is the number of packets addressed to output i that arrive during the m th time slot. This equation can also approximately analyze the case with finite buffer size when

Chapter 2: Literature Survey

PLR is low enough (e.g., lower than 10^{-6}). Fortunately, in real-world applications, typically we can expect a well-engineered switch to have a low PLR. With the assumption of having infinite buffer size, the steady-state number of packets held in the buffer can be modeled as N M/D/1 queues. The PLR can then be obtained as:

$$PLR = \Pr[\sum_{i=1}^N Q^i \geq Nb] \quad (2.9)$$

2.4.2 Queuing Analysis in Optical Packet Switches

Basically, optical packet switches share similar node architectures and functions as electronic switches, only that the signals remain in optical domain. Therefore, optical packet switches, with or without FDL buffers, can be analyzed by making minimum modifications to the models for the classic electronic switches [24], [25], [36], [39], [60], [62]. The new challenges in queuing analysis mainly come from some new features in OPS, especially the introduction of wavelength conversion [30], [38]. Since optical packets cannot be buffered for too long time, packet loss becomes the most important performance indicator while packet delay is usually no longer a crucial issue. For OPS with shared FDLs, especially with shared wavelength converters, the analytical models can be complicated [30], [93], [94]. The existing analytical results mainly reflect OPS with shared buffering by adopting the classic queuing models; while for shared wavelength conversion it needs novel convolution method to model the switch [30]. However, a simple analytical model for OPS configured with both shared FDLs and shared wavelength converters is still not available. Since in later chapters we would discuss in details existing results as well as our new, simpler models, further discussions on the existing models are omitted here.

2.5 Scheduling Algorithms

Scheduling algorithm is of critical importance for high-speed electronic switches, where various algorithms have been proposed [65], [93]-[100]. Although these scheduling algorithms may not be directly applicable in OPS due to the immaturity of optical technologies, they still provide some guidelines for the future developments of OPS scheduling.

For output-queued switches, packet scheduling is very simple since each packet can be transferred to the destined output port directly. Input-queued or VOQ systems with a certain internal speedup factor form the area where scheduling algorithms are extensively studied. It is shown that adopting proper scheduling algorithm in such switches can make them perform almost as good as the output-queuing switches. For such systems, there are basically two classes of scheduling methods: (i) *central scheduling*, which makes decisions centrally using all available information. Usually the scheduler would run a certain kind of maximal matching algorithm for achieving high throughput. Typical methods include Two-Dimensional Round-Robin (2DRR) [101], Time-Slot Assignment (TSA) and Incremental Time-Slot Assignment (I-TSA) [102] etc.; (ii) *distributed scheduling*, where each input and/or output port makes its own decisions with minimum or no information exchange. The distributed scheduling can be further classified into two categories: shared memory scheduling, with different input/output ports sharing some common information [96], [97]; and message-passing scheduling, where such common information is not available. The well-known methods belonging to the latter class include PIM [103], RRM [64], and *i*SLIP [64], [93], [95], etc. In most cases, such methods include several iterations of message exchanges

Chapter 2: Literature Survey

between an input port and its destined output ports, until finally a set transmission requests is accepted and confirmed while the others, if any, are rejected. In some best existing results, e.g., the *iSLIP* method, 100% throughput has been achieved. In shared memory scheduling, with more information being available to each port, better performance can be achieved.

Continuous growth in the demands for diversified quality-of-service (QoS) imposes new challenges to the scheduling algorithm. With the highly distributed nature of large-size switches, which typically have multiple contention points in both ingress and egress ports as well as in the switching fabric, distributed scheduling has been studied to provide differentiated QoS guarantees to individual end-to-end flows [99]. The distributed multilayered scheduler (DMS) [99] has a hierarchical scheduling structure combining several well-known packet scheduling techniques. Specifically, they include the weighted round robin (WRR) method [64] for enforcing accurate bandwidth guarantees, the generalized process sharing (GPS) [104] fluid service for providing the minimum bandwidth guarantees and achieving fairness in distribution of service and deterministic end-to-end delay bound, etc. DMS has been shown to achieve comparable performance to the ideal output-buffered switch in supporting guaranteed-delay, guaranteed-bandwidth and best-effort traffic classes. .

For wide-area optical networks, relatively less attention has been paid to develop proper scheduling algorithm that can support diverse traffic types and satisfy different QoS requests. Scheduling algorithms in wide-area networks have to be very fast and take into account physical constraints including propagation delay and the optical

Chapter 2: Literature Survey

components' tuning/switching time, etc. It is presented in [105] that a hierarchical scheduling framework is capable of supporting mixed types of traffic and QoS requests.

Last but not least, still another important topic of scheduling in OPS is the effective management of multiple shared FDLs [106], [107]. Meanwhile, scheduling algorithms for new OPS configurations are under intensive studies.

2.6 Summary

In this chapter, to provide a general background for the later discussions on contention resolutions in OPS, we proposed a brief survey on various OPS technologies including basic node architectures and functional units, queuing analysis and scheduling algorithms, etc. Three mainstream contention resolution methods and a few typical implementations in the past years have been discussed. In the following chapters, studies will be conducted with the main objective of achieving efficient contention resolution in OPS network

Chapter 3

Contention Resolution in Resource-Sharing OPS

Contention resolution is of critical importance in optical packet switching systems. As presented in the previous chapters, the most popular solutions include optical buffering and wavelength conversion, where the main constraints include the limited capacity of the fiber delay lines (FDLs) and the high cost of tunable wavelength converters (TWCs). To achieve higher efficiency, integrated configurations with both FDLs and TWCs have been proposed and extensively studied. Generally speaking, such configurations lead to better performance, while their costs, if not carefully controlled, can be high. In this chapter, we study the integrated contention resolution schemes, with the objective of proposing some guidelines for achieving better cost-effectiveness. Specifically, through extensive simulations, we evaluate how much the packet-loss performance of an OPS node can be improved by an increasing number of TWCs and FDLs, respectively. In addition, we also evaluate the different effects of utilizing FDLs and TWCs with different priorities.

3.1 Introduction

Current OPS systems generally adopt FDLs to emulate the traditional electronic buffers. Wavelength conversion can help to enhance the efficiency of optical buffers in two different ways. First, it exploits the WDM capacity in the FDLs [42], [90]. Specifically, wavelength channels inside the FDLs, called *internal wavelengths* [90], can be exploited such that the buffer capacity is increased. More significantly, wavelength conversion can exploit the wavelength dimension in the output WDM fibers where multiple packets can be sent to different wavelengths of the same outgoing link simultaneously [91], [92]. Our study in this chapter focuses on the second point: we assume WDM is not adopted in FDLs therefore each FDL can contain only a single packet each time.

In an OPS node, TWCs and FDLs can be dedicated to each input or output port, or they can be shared among some or all of the ports. Installing dedicated contention resolution components generally leads to good performance yet rather high cost. For example, in an output-buffered switch with FDLs dedicated to each output port, under uniform Bernoulli traffic load it requires a buffer depth of 55 packets *per output* to achieve a packet loss rate (PLR) of 10^{-11} under a normalized traffic load of 0.8 [63]. All the output ports combined together, the required number of FDLs becomes prohibitively large. To efficiently utilize contention resolution components, resource-sharing schemes have been investigated and demonstrated to be achieving good performance [29], [30], [108]-[112]. The reason is obvious: while dedicated components can be left idle most of time, much higher utilizations can be achieved in resource-sharing schemes. The main disadvantage of sharing contention resolution

Chapter 3: Contention Resolution in Resource-Sharing OPS

components, however, is the higher complexity in component control and coordination.

We are motivated to study the feasibility of reducing the number of contention resolution components without significantly sacrificing network performance. Specifically, we evaluate through extensive numerical simulations the performance of a typical resource-sharing scheme with different numbers of TWCs and FDLs respectively. Moreover, to evaluate the effects of utilizing different contention resolution components in different sequences, we study two different schemes as follows: 1) packet contentions are resolved by TWCs as far as possible before FDLs are used. We say that in such a scheme TWCs are utilized with a higher priority; 2) packet contentions are resolved by FDLs as far as possible before TWCs are used. We say that in such a scheme FDLs are utilized with a higher priority. Based on these evaluation results, some guidelines about how to achieve better cost-effectiveness in resource-sharing OPS will then be proposed.

3.2 Resource-Sharing OPS

3.2.1 Resource-Sharing Model and Implementation

The simple resource-sharing model that we would study in this chapter is shown in Figure 3.1. We assume that it has N input/output fibers, where each fiber contains W wavelength channels $\{\lambda_1, \dots, \lambda_W\}$. Incoming packets from each input fiber are demultiplexed and then, controlled by the control unit, scheduled to be sent to their destined output fibers. Upon a packet contention, a TWC or an FDL may be used. Selecting which kind of contention resolution resource should be used is controlled by the scheduling algorithm, depending on the availability of the contention resolution

Chapter 3: Contention Resolution in Resource-Sharing OPS

resources. Specifically, in the scheme where TWCs are utilized with a higher priority, a contending packet will be switched to an available FDL if and only if all the TWCs have been used up or there is no free wavelength in the destined output link. If all the FDLs have been exhausted, the contending packet would be dropped and lost. In the scheme where FDLs are utilized with a higher priority, a TWC would be used if and only if all the FDLs have been occupied and there are still free wavelengths in the destined output link. If all the TWCs have been exhausted, the contending packet would be dropped. Throughout this chapter, we assume that the packet switching system is time slotted and synchronized.

Figure 3.1: Resource-sharing model of an optical packet switch.

Figure 3.2 demonstrates a possible inner implementation of optical components in such a node, showing in details how TWCs and FDLs are shared among all the input/output ports. After being demultiplexed from the incoming WDM signal, each incoming packet is split by an optical coupler and then selected by a set of optical gates, which are connected to different output fibers or the shared contention resolution

Chapter 3: Contention Resolution in Resource-Sharing OPS

resources. The control unit drives all the optical components such as TWCs and optical gates etc. according to the scheduling algorithm. Going-through packets are multiplexed at the output ports and forwarded into the output WDM fibers.

Figure 3.2: Inner implementation of optical components connection.

3.2.2 Scheduling Algorithms

The packet scheduling algorithms in different OPS systems [41], [113]-[115] may be quite different with different complexities. Detailed design of the packet scheduling algorithm generally depends on OPS node configuration. In this chapter, since the main focus is to achieve better understandings of the significance of TWCs and FDLs rather than proposing a complex node configuration or an efficient scheduling algorithm, we consider two simple methods in a simple node configuration, where the main difference only lies in the sequences of utilizing the two different types of contention resolution components.

Chapter 3: Contention Resolution in Resource-Sharing OPS

We define the two methods in detail as follows:

Scheduling Algorithm 1:

- Randomly select an incoming packet and search for an available wavelength in the destined output fiber. If such is available with no packet contention, switch the packet directly.
- If a contention happens, select one available TWC to convert the wavelength of the packet to a free wavelength of the destined output fiber, if applicable.
- If there are no more available converters or the wavelengths in the destined output fiber are all occupied, buffer the packet in an FDL for one time slot. If all the FDLs have been occupied, drop the packet.

Scheduling Algorithm 2:

- Randomly select an incoming packet and search for an available wavelength in the destined output fiber. If such is available with no packet contention, switch the packet directly.
- Where a packet contention happens, select one available FDL to buffer the contending packet for one time slot, if applicable.
- If there are no more available FDLs, select an available TWC to convert the packet to a free wavelength in the destined output fiber, if applicable. If no wavelength converters are available or all the wavelengths in the output fiber have been occupied, drop the packet.

3.3 Simulation Platform

To evaluate the performance of different OPS models, we refer to a simulation process platform. For different OPS architectures and various traffic patterns, we can

Chapter 3: Contention Resolution in Resource-Sharing OPS

configure this platform to evaluate each of the specific conditions. Statistics of packet-loss performance can be drawn from the simulation results.

The flow chart of the simulation platform is demonstrated in Figure 3.3. It is assumed that all the incoming channels at the input interface are synchronized such that time-slotted operations can be conducted. In each time slot, a packet is generated at each input wavelength channel with the probability equaling to the normalized traffic load on that channel. The destinations of the generated packets are calculated according to the traffic pattern. For example, a bursty traffic can be simulated by using an on-off source model [27] while the packets are continuously generated until the status of the source changes from “on” to “off”. Meanwhile, the destination of the generated traffic can be randomly calculated according to the given output distribution.

The newly generated packets, together with those packets buffered in previous time slots (if any), will then go through the stage of contention judgment. Those packets with no contentions are directly sent out to their destined output ports, while the others move to the next stage to be handled by the contention resolution resources such as buffers and/or wavelength converters. Some packets may have to be dropped when the contention-resolution resources are exhausted.

Statistics of packet-loss performance can be obtained by simulating the above procedure for a large number of time slots. Specifically, under high traffic load when PLR is higher than 10^{-4} , simulating the packet switching operations for 10^6 time slots is typically sufficient to achieve a reliable result. When traffic load is low, however, very few packets would be dropped. Under such case, we may have to simulate 10^7 or even more time slots in the simulation.

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.3: Flow chart of the simulation process platform.

Chapter 3: Contention Resolution in Resource-Sharing OPS

There are different OPS architectures and contention resolution schemes that need to be simulated. For example, the OPS node may have only FDLs or only TWCs or both of them; and there may be different priorities in utilizing FDLs and TWCs in contention resolutions as that has been shown in Section 3.2.2. For these different cases, detailed design of the function block of “contention resolution” as shown in Figure 3.3 would be different, while the overall flowchart structure and all the other function blocks can largely remain unchanged.

3.4 Performance Evaluation

In most existing OPS models, to simplify the analysis and evaluations, the traffic load through different output ports is assumed to be following a uniform distribution. In our simulations, similarly, we first consider the case where the arriving packets conform to the Bernoulli process with a normalized arrival rate α , destined to each one of the N output ports with a uniformly-distributed probability of $\rho_i = 1/N$, $i = 1, 2, \dots, N$. Considering the fact that the arriving and outgoing traffic load at a switching node is mostly unbalanced due to the dynamic fluctuations in the networks, we will also evaluate the performance of the switch under non-uniform traffic load. Specifically, we would simulate a non-uniform traffic pattern where $\rho_i / \rho_{i-1} = R$, ($i = 2, 3, \dots, N$) and $\sum_{i=1}^N \rho_i = 1$, ($0 \leq \rho_i \leq 1$, $i = 1, 2, \dots, N$). In other words, the average traffic load destined for different output ports distributes as a geometric sequence.

Denote the number of the shared TWCs as N_{TWC} and the number of the shared FDLs as L . The performance metric adopted in our evaluations is the *Packet Loss Rate*

Chapter 3: Contention Resolution in Resource-Sharing OPS

(*PLR*) defined as follows:

$$PLR = \sum_{j=1}^N P_l(j) / \sum_{k=1}^N P_a(k) \quad (3.1)$$

where $P_l(j)$ represents the number of lost packets in the j -th output link and $P_a(k)$ the number of arriving packets from the k -th input link, respectively. From the definitions of the scheduling algorithms, we see that packet losses may contain three different parts: (1) those caused by insufficient output wavelengths (denoted as L_w); (2) those caused by exhaustions of TWCs (denoted as L_c) and (3) those due to the limited number of FDLs (denoted as L_f). In this chapter, we investigate the methods of improving *PLR* performance by relaxing the latter two factors. The following two sections present the evaluation results for the two different scheduling algorithms respectively.

3.4.1 Evaluations of Scheduling Algorithm 1

Figure 3.4 shows the results of *PLR* versus the number of TWCs under the uniform traffic load. We assume that the switch is connected to $N = 16$ input/output fibers with $W = 4$ wavelengths per fiber. Each input wavelength channel has a packet arrival rate of $\alpha = 0.8$. When we increase the number of TWCs from $N_{TWC} = 1$ to 5, *PLR* decreases quickly. After that, however, the improvements to *PLR* become saturated. We observe that having six TWCs and sixteen TWCs perform nearly the same. In fact, even if we further increase the number of TWCs to 64 (which is equivalent to the dedicated TWC case where each input port is equipped with a wavelength converter) no significant further improvement could be made. On the other hand, as we could see in Figure 3.4, increasing the number of FDLs steadily leads to better *PLR* performance, regardless of the number TWCs used.

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.4: Performance of resource-sharing OPS model with TWCs and FDLs under uniform traffic load with Scheduling Algorithm 1.

Figure 3.5 shows the results of PLR versus the number of TWCs under non-uniform traffic load. We set the traffic load ratio $R = 1.2$ and leave all the other parameters unchanged. Generally, $PLRs$ under non-uniform traffic become higher than those under uniform traffic. This is because that under un-balanced loads, some output ports may be heavily loaded and subject to a large number of packet contentions, which cannot all be handed by limited contention resolution resources (Later in Chapter 7, a comprehensive study will be conducted to verify this observations by analytical as well as extensive simulation results). However, despite of the increased PLR , the main conclusion remains valid. That is, with an increasing number of TWCs, the improvements to the PLR performance soon become saturated. Specifically, six TWCs still manage to perform nearly the same as sixteen TWCs. Meanwhile, an increasing number of FDLs still steadily improve the PLR performance.

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.5: Performance of resource-sharing OPS model with TWCs and FDLs under non-uniform traffic load with Scheduling Algorithm 1.

To explain the saturation of improvements to the PLR performance with an increased number of TWCs, Figure 3.6 shows the percentage among all the lost packets those due to exhaustions of TWCs (denoted as L_c) where we set $N = 16$, $L = 0$ and $\alpha = 0.8$. We see that the percentage that L_c accounts for drops sharply when N_{TWC} increases from 1 to 6. When $N_{TWC} \geq 8$, less than 1% of the packet losses are caused by the exhaustion of TWCs. In other words, under such case, most packet losses happen when all the wavelength channels in the destined output fiber have been occupied, for which case adding more TWCs would not really help. If more FDLs are available in such case, on the other hand, contending packets can be buffered temporarily instead of being dropped, which leads to better PLR performance,

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.6: The percentage of lost packets due to lack of TWCs under uniform traffic load with Scheduling Algorithm 1.

Figure 3.7: The percentage of lost packets due to lack of TWCs under non-uniform traffic load with Scheduling Algorithm 1.

Chapter 3: Contention Resolution in Resource-Sharing OPS

The percentage of L_c accounts for under non-uniform traffic is shown in Figure 3.7, where similar phenomenon as that in Figure 3.6 can be observed: the percentages of packet losses caused by exhaustions of TWCs drop quickly and soon approach virtually zero. After that the packet losses mainly come from insufficient output wavelengths. Therefore, while a limited number TWCs manage to improve the PLR performance significantly, installing more than enough TWCs does not really help.

The performance improvements by increasing the numbers of shared FDLs are shown in Figure 3.8 and Figure 3.9. In these evaluation simulations, we fix the number of TWCs as 8. Under uniform traffic load, as shown in Figure 3.8, where $W = 4, 5, 6$, PLR can all drop to below 10^{-6} with a small number of shared FDLs ($L = 4, 3, 2$). Under non-uniform traffic load, as shown in Figure 3.9, though $PLRs$ become higher, the effects of FDLs in helping lower PLR remain as significant. What is most significant is the steady decrease of PLR with an increasing number of FDLs (with or without shared TWCs). In fact, in all the simulations we have conducted so far, the effects of FDLs in helping lower PLR have never got exhausted. More FDLs always lead to better performance until PLR reaches zero. However, we have to point out that installing more FDLs has its own drawbacks, among which the most serious one may be that it increases the switch size. Larger switch size leads to higher power loss, worse signal distortions and crosstalk, and equivalently seriously, higher cost as well. Therefore, we have to carefully keep a good balance between switch performance and hardware cost. Better understandings of how FDLs and TWCs help to improve packet-loss performance obviously help to achieve such a balance.

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.8: Performance improvement by using shared FDLs under uniform traffic load with Scheduling Algorithm 1.

Figure 3.9: Performance improvement by using shared FDLs under non-uniform traffic load with Scheduling Algorithm 1.

Chapter 3: Contention Resolution in Resource-Sharing OPS

3.4.2 Evaluations of Scheduling Algorithm 2

In Scheduling Algorithm 2, if and only if all the FDLs have been used up, would a contending packet search for an available wavelength converter. In the following evaluations, we would still use the same set of parameters as those in Figure 3.4 and 3.5, where $N = 16$, $W = 4$, $\alpha = 0.8$. Under the uniform traffic case, as shown in Figure 3.10, we still observe the saturated improvements in PLR s when more than enough TWCs are installed. Similar observation applies to the case under non-uniform traffic load, as shown in Figure 3.11. In both cases, the conclusion holds that six TWCs perform nearly the same as sixteen TWCs, whereas an increased number of FDLs is favorable and steadily leads to lower PLR . In short, the main conclusions about the effects of TWCs and FDLs to the PLR performance remain unchanged as those for Scheduling Algorithm 1.

Figure 3.10: Performance of resource-sharing OPS model with TWCs and FDLs under uniform traffic load with Scheduling Algorithm 2.

Chapter 3: Contention Resolution in Resource-Sharing OPS

Figure 3.11: Performance of resource-sharing OPS model with TWCs and FDLs under non-uniform traffic load with Scheduling Algorithm 2.

3.4.3 Comparisons between the Scheduling Algorithms 1 and 2

It is then of interest to compare the performance of the two different scheduling algorithms. Intuitively, we expect that Scheduling Algorithm 1 would perform better: by giving utilizing TWCs a higher priority: fewer packets will be buffered than those in Scheduling Algorithm 2, while buffered packets would contribute as part of the traffic load that has to be handled in the next time slot. However, it is not obvious how significant the difference in the PLR performance caused by the two scheduling algorithms would be. We make comparisons in Table 3.1 by summarizing the simulation results for the case where $N = 16$, $W = 4$, $\alpha = 0.8$, and $N_{TWC} = 8$. Here S1 represents the Scheduling Algorithm 1 and S2 the Scheduling Algorithm 2. It verifies our intuitive observation that S1 should outperform S2. However, the difference has

Chapter 3: Contention Resolution in Resource-Sharing OPS

never been significant. In other words, the buffered packets, though add to the next time-slot traffic, do not significantly degrade the switch performance.

PLR	S1	S2	PLR	S1	S2
$L = 0$	1.8202E-3	1.8202E-3	$L = 0$	3.3563E-3	3.3563E-3
$L = 1$	2.3183E-4	2.3653E-4	$L = 1$	1.0458E-3	1.0735E-3
$L = 2$	2.5976E-5	2.6432E-5	$L = 2$	2.5468E-4	2.5834E-4
$L = 3$	2.4411E-6	2.5335E-6	$L = 3$	3.4726E-5	3.5435E-5
$L = 4$	2.8909E-7	2.9853E-7	$L = 4$	3.9851E-6	4.0724E-6

(a) Uniform Traffic

(b) Non-uniform traffic

Table 3.1: Comparison of PLR between the two scheduling algorithms.

Since Scheduling Algorithm 1 steadily outperforms Scheduling Algorithm 2 (though not quite significant), unless otherwise specified, we will use the Scheduling Algorithm 1 in the later chapters when TWC and FDL are both implemented.

Lastly, from the *PLR* performance comparisons, people may get the impression that shared TWCs are inferior to shared FDLs since shared TWCs anyway also make switch size larger, while they are not as efficient in lowering *PLR* as FDLs. However, such a conclusion may not be fair after all: TWCs have one advantage that FDLs can never compete; that is, in TWC solutions, contending packets are not delayed (buffered) at the switching node. As a result, the power losses and the distortions imposed on the optical signals may become lower. In networks with long routes, this advantage may be of significant importance. Therefore, integrated schemes, which adopt a combination of a small number of TWCs and a moderate number of FDLs, to our point of view, remain as most attractive solutions.

3.5 Conclusion

In this chapter, we evaluated the packet loss performance of resource-sharing OPS models which exploit the integrated contention resolutions with shared TWCs and FDLs. We investigated two different scheduling algorithms where TWCs and FDLs are utilized in different sequences of priority. For each of the scheduling algorithms, the relationship between *PLR* performance and the number of TWCs and FDLs was studied. Extensive simulation results showed that the improvements achieved by an increasing number of TWCs can easily get exhausted. On the other hand, installing more FDLs steadily leads to a lower *PLR*. Such conclusions hold for different switches equipped with different number of contention resolution components, under uniform or non-uniform traffic load. Comparing the performance of the two scheduling algorithms, we demonstrated that utilizing TWCs at a higher priority leads to a slightly better performance. Finally, we briefly discussed why the integrated schemes utilizing both TWCs and FDLs remain as an attractive candidate among various contention resolution solutions.

Chapter 4

Fixed Wavelength Conversion in OPS Node

The wavelength channels in WDM links create a wavelength dimension for data-packet transmission. Wavelength conversion plays an important role in exploiting this dimension, especially in contention resolutions. However, implementing a large number of full-range tunable wavelength converters (FTWCs) in an OPS network remains hardly feasible, mainly due to the immature all-optical technologies and the prohibitively high cost. To achieve better cost-effectiveness, we propose a cheaper scheme for OPS networks where fixed wavelength converters (FWCs) instead of FTWCs are used in contention resolutions. By comparing its performance against that of the OPS with FTWCs under various traffic loads, we show that the FWC scheme performs nearly the same as the FTWC scheme. Therefore, the proposed FWC scheme provides a more cost-effective solution.

4.1 Introduction

As presented in Chapter 2, wavelength conversion enables efficient utilizations of the wavelength dimension in WDM systems such that more efficient contention resolutions could be achieved. The main problem, however, is the immature wavelength conversion technology [73], [80]-[82], [116], [117]. It would take time before the FTWCs with reasonable price could be commercially available. Therefore, alternative solutions have to be developed to achieve good performance, comparable to that of OPS with FTWCs if applicable.

Basically, there are two different types of wavelength converters: *tunable* wavelength converters (TWCs) which can convert incoming optical signal to many different output wavelengths upon request; and *fixed* wavelength converters (FWCs) which convert incoming signal to only a single, pre-defined wavelength. Among the TWCs, the most powerful one is the FTWC which covers all the output wavelengths, as shown in Figure 4.1. Though FTWC with fast tuning speed has the best flexibility, it is still under development and is expected to be quite expensive in any foreseeable near future. To date, the majority of wavelength conversion and regeneration relies on the optical modulation of power gains and/or refractive indexes of semiconductor optical amplifiers (SOAs). Such schemes, though generally not good enough for constructing efficient FTWCs, can steadily support FWC operation [74], [75], [77]-[79], [117].

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.1: Conversion ranges of FTWC and FWC.

As we discussed in Chapter 3, to lower the cost of constructing an optical packet switch, a certain number of wavelength converters could be shared among all the input/output ports. In this chapter, we study the packet loss performance where shared FWCs, instead of shared FTWCs, are installed in OPS nodes. Extensive simulation results show that the much cheaper FWC solution manages to achieve comparable performance as that of the classic FTWC solution under various traffic loads. Consequently, better cost-effectiveness is achieved.

4.2 Switch Models and Algorithm

The switch models with FWCs and FTWCs are illustrated in Figure 4.2 (a) and 4.2 (b) respectively, both of which have N input/output links. In each link, each fiber carries W wavelength channels $\{\lambda_1, \dots, \lambda_W\}$, and the incoming WDM channels are demultiplexed before entering the switching matrix. When several incoming packets on the same wavelength are destined for the same output fiber, the contending packets can be converted into free wavelengths (if any) by wavelength converters (if available).

Chapter 4: Fixed Wavelength Conversion in OPS Node

(a) FWC model

(b) FTWC model

Figure 4.2: OPS models with different types of wavelength conversions.

Chapter 4: Fixed Wavelength Conversion in OPS Node

In FWC model, the incoming packets can reach every FWC, which respectively has a specific output wavelength. As for FTWC model, the incoming packets can reach every FTWC, which can be tuned to any of the output wavelengths. Since the FWCs and FTWCs are shared among different input/output ports, high utilization of wavelength conversion resources can be achieved.

When packet contentions cannot be totally resolved by wavelength conversion, contending packets can be switched to FDLs for buffering when applicable. A node configuration with both shared FWCs and shared FDLs, which is the case we would evaluate in this chapter, is shown in Figure 4.3. A similar configuration for full-range conversion model can be obtained by replacing FWCs with FTWCs. As usual, we assume that all the packets are of a fix length and synchronized when they enter the switches. The simple switching algorithm is defined as follows:

Switching Algorithm of FWC Scheme

1. An incoming packet is directly switched to its destined output fiber on its original wavelength, if no contention occurs.
2. When there is a packet contention, an available FWC will be used to convert the contending packet to another free wavelength of the destined output fiber, if applicable.
3. If there is no available wavelength converter or no free wavelength in the destined output fiber, then the contending packets will be switched to FDLs, as long as there is still free capacity in FDLs; otherwise, the contending packets will be dropped.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.3: The OPS node with shared FWCs and FDLs.

4.3 Performance Evaluations and Discussions

To evaluate the performance of the proposed fixed wavelength conversion scheme, we compare its packet loss rate (*PLR*) to that of the corresponding node configuration equipped with FTWCs. Configuration parameters of the OPS node include: the number of input/output ports (N), the number of wavelengths per fiber (W), the size of FDL buffers (N_{FDL}), and the total number of FWCs or FTWCs (N_{TWC}). In this chapter, we still assume that the packet arrival pattern in each wavelength channel follows a Bernoulli process with an average normalized traffic load of α . We study different cases with different levels of traffic burstiness and different traffic patterns.

Chapter 4: Fixed Wavelength Conversion in OPS Node

The burstiness of dynamic traffic load measures the traffic distributions in time domain. Specifically, in different time durations, the number of arrival packets may be quite different, forming peaks and bottoms. An even distribution of traffic load in time domain is called non-bursty traffic, where the packets are arriving smoothly; otherwise, we call it as bursty traffic. For each of them, we further study different cases with uniform and the non-uniform traffic patterns, respectively.

4.3.1 Evaluations under Non-Bursty Traffic Load

Figure 4.4 compares the performance of the FWC scheme and the FTWC scheme under non-bursty, uniformly-distributed traffic load. Comparison results under non-uniform traffic load are presented in Figure 4.5. We assume that the non-uniform traffic follows the geometric distribution where $\rho_i / \rho_{i-1} = R = 1.2$, ($i = 2, 3, \dots, N$). As that has been defined in Chapter 3, ρ_i denotes the probability that a packet is destined for output port i .

From Figure 4.4, we see that in a small-sized switch with $N = 8$, $W = 4$, and $N_{TWC} = 4$, FWC scheme performs nearly the same as FTWC scheme under different traffic load. In a larger switch where $N = 16$, $W = 16$, and $N_{TWC} = 16$, the variance between the *PLRs* of the two schemes can be observed under low traffic load (e.g., where $\alpha = 0.2$). However, since the *PLR* of the FWC scheme is already as low as 10^{-6} , this difference is not critical.

In Figure 4.5, although the non-uniform traffic load ($R = 1.2$) brings certain deterioration in *PLR* because of the more intensive contentions at highly-loaded ports, we see that the FWC scheme can still achieve nearly the same performance as the FTWC scheme.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.4: Performance comparisons between the FWC and the FTWC schemes under uniform and non-bursty traffic with no FDLs.

Another observation worth mentioning is that, with the same normalized traffic load, large-sized switches perform much better than small-sized ones under light traffic load. The difference, however, reduces to virtually zero under heavy traffic load. The better performance of larger switches under light traffic load comes from the more redundant wavelength resources that can be used. Under heavy traffic load, however, the increased free wavelengths are competed by more contending packets. Therefore, the *PLR* cannot be significantly lowered.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.5: Performance comparisons between FWC and FTWC schemes under non-uniform and non-bursty traffic with no FDLs ($R = 1.2$).

For the case where there are FDLs in the switching node, we observe in Figure 4.6 that the two schemes remain to perform comparably to each other under uniform traffic load. When the traffic load is heavy, they actually perform nearly the same. This is because that under heavy traffic load, packet contentions are mainly caused by insufficient free wavelength channels in the output fibers. Therefore even the most powerful wavelength conversion cannot make significant improvement to the performance of the switch.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.6: Performance comparisons between the FWC scheme and the FTWC scheme under uniform and non-bursty traffic with FDLs.

Evaluation results under non-uniform traffic (again, where $R = 1.2$) can be found in Figure 4.7, where FWC and FTWC schemes still perform comparably to each other. We see that the performance variance of the two schemes with FDL buffers is a bit larger than that of the case without FDL, though the variance is not outstanding. The increased variances demonstrate that FDLs help to slightly better exploit the benefit of full-range wavelength conversion, which can be explained as follows: when we allow random access to optical buffers by contending packets, the non-uniformity of traffic distributions is generally increased (since heavier-loaded ports, statistically speaking, would have more packets to be buffered), while increased non-uniformity makes the better flexibility of FTWC schemes a more visible advantage in contention resolutions.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.7: Performance comparisons between the FWC scheme and the FTWC scheme under non-uniform and non-bursty traffic with FDLs ($R = 1.2$).

In above discussions, we have let the number of wavelength converters be equal to that of different wavelengths. Such is necessary in FWC scheme. It is of research interest to see whether FTWC can actually achieve nearly the same performance when $N_{TWC} < W$ (in which case the benefit of having more FTWCs gets saturated *before* the number of FTWCs reaches W). We evaluate the example cases under uniform traffic load $a = 0.5$ where $N = 8$, $W = 16$, and $N_{FDL} = 0, 2, 4$ respectively. Figure 4.8 shows that in all the different cases, having more FTWCs steadily lead to better performances. Therefore, indeed having W FWCs achieves comparable performance as that of having the same number of FTWCs.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.8: Performance of FTWC scheme with different number of $FTWCs$ under uniform and non-bursty traffic.

4.3.2 Evaluations under Bursty Traffic Load

Burstiness exists in most real-world traffic. In an optical switch, a long IP packet can be divided and contained in multiple fixed-time slots, while a short packet can be contained in just one time slot. It has been shown that under bursty traffic, the probability of having packet contentions generally becomes higher [37]. As we have shown in the last section, under non-bursty traffic, different traffic patterns do not significantly affect the variance between FWC and FTWC schemes. Now we study the performance of optical switches under bursty traffic, with both uniform and non-uniform traffic patterns.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.9: Performance comparisons between FWC and FTWC schemes with no FDLs under bursty traffic with uniform distribution.

In Figure 4.9, we compare the performance of the two schemes with no FDLs under bursty traffic with uniform traffic pattern. We assume that each input port can be modeled as a Bernoulli-distributed burst with an average length of $L_b = 5$. It is shown that FTWC scheme outperforms FWC scheme where both of them have the same number of 16 wavelength converters. However, FWC scheme with 32 FWCs does outperform FTWC scheme with 16 FTWCs. Therefore, the more cost-effective solution may be to use a larger number of mature, cheaper FWCs instead of a smaller number of immature, highly expensive FTWCs. When traffic load is non-uniformly distributed, as shown in Figure 4.10, though $PLRs$ become higher than those under uniform traffic load, FWC scheme with 32 FWCs still outperforms FTWC scheme with 16 FTWCs.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.10: Performance comparisons between FWC and FTWC schemes with no FDLs under bursty traffic with non-uniform distribution ($R = 1.2$).

Figure 4.11 compares the performance of FWC and FTWC schemes under uniform traffic where FDLs are implemented. We assume that the buffer size is $N_{FDL} = 4$. We see that though the $PLRs$ are lowered by FDLs, the conclusion holds that FWC scheme with 32 FWCs outperforms FTWC scheme with 16 FTWCs. This conclusion remains valid under non-uniform traffic pattern, as shown in Figure 4.12, despite of the increased overall $PLRs$. Therefore, a moderately larger number of FWCs can be implemented, as a more cost-effective solution compared to FTWC scheme.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.11: Performance comparisons between FWC and FTWC schemes with FDLs under bursty traffic with uniform distribution.

Figure 4.12: Performance comparisons between FWC and FTWC schemes with FDLs under bursty traffic with non-uniform distribution ($R = 1.2$).

4.3.3 Discussions on the Limited Wavelength Conversion

From previous discussions, we have seen that FWCs with the full-range input wavelengths and a single output wavelength perform comparably to FTWCs. A more general case is the limited wavelength conversion (LWC), where each wavelength converter can only handle a limited range of input and output wavelengths [67], [116]. We have shown that a limited output wavelength range (even as small as having only a single wavelength) does not significantly degrade switch performance. In this section, we study the effect of having a limited input wavelength range.

In Figures 4.13 and 4.14, we compare the performance of LWC and FWC under uniform and non-uniform traffic patterns respectively. Letting $N = 8$, $W = 16$, $N_{TWC} = 16$ and $N_{FDL} = 0$, we simulate two different LWC cases. In the first case (named as LWC-1), the input wavelength range contains half of all the different wavelengths (i.e., there are eight different input wavelengths) while the output range remains as a single wavelength as that of an FWC. Specifically, the first converter handles input wavelengths $\{\lambda_2, \lambda_3, \dots, \lambda_8, \lambda_9\}$ while its output wavelength is $\{\lambda_1\}$; the second converter receives from input wavelengths $\{\lambda_3, \lambda_4, \dots, \lambda_9, \lambda_{10}\}$ and outputs on $\{\lambda_2\}$, etc. In the second case (named as LWC-2), each wavelength converter has the same input wavelength range as that of LWC-1 but a much larger output wavelength range. Specifically, we let the output range cover all the eight wavelengths outside the input wavelength range. For example, the first converter has an input range of $\{\lambda_1, \lambda_2, \dots, \lambda_7, \lambda_8\}$ and an output range of $\{\lambda_9, \lambda_{10}, \dots, \lambda_{15}, \lambda_{16}\}$.

Chapter 4: Fixed Wavelength Conversion in OPS Node

Figure 4.13: Performance comparisons between FWC and LWC under non-bursty traffic with the uniform traffic pattern.

Figure 4.14: Performance comparisons between FWC and LWC under bursty traffic load ($L_b=10$) with the non-uniform traffic pattern ($R=1.2$).

Chapter 4: Fixed Wavelength Conversion in OPS Node

From Figure 4.13, we see that LWC-1 performs much worse than FWC, especially under light traffic load ($a < 0.4$). Similar observation can be made in Figure 4.14 under bursty and non-uniform traffic load. The observation is that reducing the input wavelength range of a LWC solution leads to serious degradation of packet-loss performance, which can be explained as follows: though FWC has only a single output wavelength, the W FWCs combined together still provide satisfactory flexibility in contention resolutions: a contending packet can reach all the converters and therefore has access to all the output wavelengths. When the input wavelength range is reduced, however, each contending packet can only access a portion of all the wavelength converters and hence loses the chance of utilizing the other wavelength converters though there may be still idle capacities within their output wavelength ranges. Such factor also exists in LWC-2: being equipped with a larger output wavelength range, LWC-2 does significantly outperform LWC-1; however, its PLR remains to be higher than that of the FWC scheme, especially under light traffic load.

To summarize, we see that the FWC scheme outperforms a typical LWC scheme where both the input and output wavelength ranges are halved.

4.4 Conclusion

Tunable wavelength converters with large tunable range and fast tuning speed are still under developments and will unlikely to become mature and cheap in any foreseeable near future. To achieve better cost-effectiveness of contention resolution in OPS, we proposed a new scheme using only the fixed output wavelength converters. Simulation results show that under either bursty or non-bursty traffic load with uniform

Chapter 4: Fixed Wavelength Conversion in OPS Node

or non-uniform traffic patterns, with or without FDLs, the new scheme can achieve comparable performance to that of the scheme utilizing the same number of FTWCs. By comparing the FWC scheme with the scheme adopting LWCs, we showed that the FWC actually performs better.

Chapter 5

Performance Evaluation of Multi-fiber Optical Packet Switches

Multi-fiber WDM networks are becoming the major telecommunication platforms for transmitting exponentially increasing data traffic. While today's networks are mainly providing circuit-switched connections, optical packet-switching technologies have been investigated for years, aiming at achieving more efficient utilizations of network resources. In this chapter, we evaluate, for the first time to the best of our knowledge, the packet-loss performance of multi-fiber optical packet switches (MOPS). Our main contributions are three-fold. Firstly, we propose simple and accurate analytical models for analyzing packet loss performance of (i) the most fundamental MOPS configuration, (ii) MOPS equipped with fiber delay lines (FDLs) and (iii) with shared wavelength converters (SWCs). Secondly, we show that MOPS networks *cannot* achieve the same performance as the ones with full-range tunable wavelength conversion (FTWC), which is quite different from the well-known conclusion in circuit-switched networks. However, MOPS does significantly outperform the classic

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

single-fiber switches. By introducing a small number of FDLs or SWCs, it outperforms the highly expensive FTWC solution as well. Finally, we take the hardware constraints into consideration by evaluating the performance of MOPS with multiple fixed-sized switching boards, which leads to some useful insights for developing cost-effective MOPS configurations in the future.

5.1 Introduction

Recently, *multi-fiber* networks, in which each link contains multiple WDM fibers [2], [118]-[124] are attracting increasing research interests. This is because, in practical network deployments, generally a large number of fibers contained in a cable are laid underground [123]. More significantly, it has been shown that, circuit-switched multi-fiber networks with a moderate number of wavelengths in each fiber achieve better cost-effectiveness than single-fiber networks with high-density wavelengths [124]. It is well-known that a multi-fiber WDM network with no wavelength conversion performs equivalently well as a single-fiber network with unlimited wavelength conversion (known as *full wavelength conversion*) [2]. All these results, however, works on circuit-switched networks, with the assumption that end-to-end lightpaths are set up based on global link-state information. This assumption may not hold in packet-switched networks, where every node has to make quick, *local* decisions upon the arrivals of data packets.

In this chapter, we evaluate the performance of MOPS. Simple yet accurate analytical models are proposed for the most fundamental MOPS configuration as well

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

as MOPS with FDLs and/or SWCs, respectively. Analytical and extensive simulation results show that, under both uniform and non-uniform traffic load,

- Similarly to that in circuit-switched networks, multi-fiber networks with a moderate number of wavelength channels in each fiber significantly outperform single-fiber networks with high-density wavelengths, and
- Differently from that in circuit-switched networks, multi-fiber OPS networks *cannot* achieve nearly the same performance as single-fiber networks with FTWC. However, equipped with a few FDLs and/or SWCs, MOPS can easily outperform the single-fiber switch with FTWC. Therefore, better cost-effectiveness could be achieved.

Finally, to take into account the hardware constraint that large-sized optical switching boards are complex and expensive [21], [38], [125], we study a simple node configuration containing multiple fixed-sized switching boards. By demonstrating the relationship between the packet loss rate (PLR) and the size of switching boards as well as the number of fibers and FDLs connected to each board respectively, we obtain some insightful observations which are useful for the future developments of cost-effective MOPS networks.

5.2 MOPS Node Model

A generic MOPS node architecture is shown in Figure 5.1, where all the incoming and outgoing wavelength channels are connected to a single switching board. To resolve packet contentions, a certain number of shared FDLs and/or SWCs can be installed. In this chapter, we assume that each SWC has full conversion range. In other

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

words, it can convert an optical packet from any incoming wavelength to any outgoing wavelength. Both SWCs and FDLs are shared among all the input/output ports and are accessible by all the contending packets. It is well-known that such resource-sharing configurations lower system cost while maintaining high efficiency of contention resolutions [24], [30], [126].

When there are many fibers in each link, the single switching board in Figure 5.1 can be extremely large. Specifically, we assume there are N input/output links with F fibers per link and W wavelength channels per fiber, therefore the switching board size would be as large as $N \cdot F \cdot W + N_{FDL} + N_{SWC}$, where N_{FDL} and N_{SWC} denote the numbers of FDLs and SWCs, respectively. Based on today's technology, manufacturing large-sized, high-speed optical switching matrix on a single board is still difficult and expensive [20], [21]. To solve this problem, we have proposed a simple multi-board configuration as shown in Figure 5.2, where every node contains multiple fixed-sized switching boards, each of which operating on a single wavelength. For simplicity, we assume that each switching board is connected to the same number of T channels of each input/output link ($T \leq F$). Note that there could be multiple boards operating on the same wavelength (if $T < F$). Since different wavelengths would not be connected to the same board, SWCs are no longer necessary while FDLs can still be installed. The multi-board configuration helps lower the complexity and the cost of the switch. However, the switch is no longer strictly non-blocking. We will study how PLR is affected by the switching-board size and, for a given switching-board size, how PLR is affected by the number of incoming/outgoing fibers and FDLs connected to each board respectively.

Figure 5.1: General node architecture of single-board MOPS.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Figure 5.2: Node architecture of the multi-board MOPS.

5.3 Analytical Model of Packet Loss in MOPS

For the MOPS configuration as shown in Figure 5.1, we assume that there are N input/output links with F fibers per link, where each fiber carries W wavelengths denoted as $\lambda_1, \lambda_2, \dots, \lambda_W$, respectively. Since there are multiple λ_i 's, $i = 1, 2, \dots, W$, in each link, there is a “space dimension” for contention resolutions. Specifically, F packets destined for the same output link on the same wavelength can be accepted without wavelength conversion or FDL. We initiate our analysis by considering MOPS without FDL or SWC (referred to as *pure MOPS* hereafter) where contention

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

resolutions solely rely on exploiting the space dimension and then study MOPS with FDLs, SWCs and both of them, respectively.

A. Analysis of the Pure MOPS Model

We first consider the single-board, pure MOPS configuration by removing all the FDLs and SWCs from Figure 5.1. Since contention resolutions on different wavelengths are independent operations in absence of wavelength conversion, we start by analyzing PLR on a single wavelength w ($1 \leq w \leq W$). Here, we assume that MOPS is operating in a time-slotted style where incoming packets are synchronized before entering the switching board(s).

Denote the probability that in each time slot there is a packet arriving at each input port as α_w ; and this packet is destined for the i -th output link with a probability $\rho_{w,i}$

(Apparently, we have $\sum_{i=1}^N \rho_{w,i} = 1$). The probability that there are k packets destined for

the same output port in the i -th link on wavelength w can be calculated as:

$$P_{w,i}(k) = \binom{NF}{k} \cdot (\rho_{w,i} \alpha_w)^k \cdot (1 - \rho_{w,i} \alpha_w)^{NF-k}. \quad (5.1)$$

For the special case where $F = 1$ and under uniform traffic load, by letting $\rho_{w,i} = 1/N$, $i = 1, 2, \dots, N$, in Eq. (5.1), we have the well-known result that [30], [41], [42], [62], [63]:

$$P_{w,i}(k) = \binom{N}{k} \cdot \left(\frac{\alpha_w}{N}\right)^k \cdot \left(1 - \frac{\alpha_w}{N}\right)^{N-k}. \quad (5.2)$$

From Eq. (5.1), the packet loss rate of the pure MOPS can be calculated as:

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

$$PLR = \frac{\sum_{i=1}^N \sum_{w=1}^W \sum_{k=F+1}^{NF} P_{w,i}(k) \cdot (k - F)}{N \cdot F \cdot W \cdot \alpha}. \quad (5.3)$$

For the multi-board pure MOPS model where FDLs are removed from Figure 5.2, the calculations of PLR are quite similar. Specifically, since there are T channels from each input/output link connected to the same board on the same wavelength, the modifications required include:

- F in Eqs. (5.1) and (5.3) should be replaced by T in order to calculate the PLR on each switching board;
- Eq. (5.3) should be properly revised to calculate the overall PLR, not on all the different wavelengths, but on all the different switching boards.

Detailed discussions are straightforward and therefore omitted.

B. Analysis of MOPS with Shared FDLs

PLR analysis of OPS with shared contention resolution devices (FDLs or SWCs) under general non-uniform traffic load is notoriously difficult [24], [30], [41]. The main reason is that in such switches, different output ports have different probabilities of having contentions and different probabilities that such contentions can be finally resolved by FDLs or SWCs. The relations between these probabilities and the traffic load destined for different output links, however, are highly complicated even under the simplest assumption we could reasonably make (e.g., that contending packets will randomly compete for the available resolution resources). To avoid the complicated analysis as those in existing literature, we restrict ourselves to analyze MOPS with shared FDLs and/or SWCs under *uniform* traffic load. Switch performance under general non-uniform traffics will be evaluated through extensive simulations.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

It is shown in [63] that the total packet loss in a number of identical and independent distribution (i.i.d.) channels where each channel has statistically identical arrival traffic with no correlation in between can be evaluated by convolution method. Similar to that, for pure MOPS where $F > 1$, with the assumption that every output channel is identical and independent, the probability $P_w^F(x)$ of having totally x packets destined for the same output link on wavelength w (Note that now each link has F channels on the same wavelength.) can be obtained as:

$$P_w^F(x) = \bigotimes_{i=1}^F P_w(k) \Big|_{k=x}, \quad (5.4)$$

where $P_w(k)$ is defined in Eq. (5.2), and $\bigotimes_{i=1}^F$ represents that the calculations of convolution need to be executed F times.

For MOPS with shared FDLs, we make the assumption that contending packets on different wavelengths have the same probability of being buffered. Under such case, PLR on different wavelengths would be the same, which equals to the overall PLR of the switch. Therefore in this subsection wavelength subscript is dropped in the analytical models.

For simplicity purposes, we assume that all FDLs are of the same length that can hold optical packets for one time slot. The analytical model is developed as follows. Denote $U(x)$ as the probability of having x contending packets that *cannot* be handled by F fibers in the same output link. We have:

$$U(x) = \begin{cases} \sum_{k=1}^F P^F(k), & x = 0, \\ P^F(x + F), & x = 1, 2, \dots, N \cdot F - F. \end{cases} \quad (5.5)$$

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

In this equation, $U(0)$ calculates the probability that no packet is lost when there are fewer than F packets destined for the same output link on the same wavelength; and $U(x)$, $x = 1, 2, \dots, N \cdot F - F$, calculates the probability that x packets have to be either buffered or dropped. To calculate the probability of having x contending packets in the entire N output links that have to be either buffered or dropped, we have:

$$C(x) = \bigotimes_{i=1}^{NW} U(k) \Big|_{k=x} \quad (5.6)$$

Assume we have N_{FDL} FDLs. When the number of contending packets $x \leq N_{FDL}$, all the packets can be buffered; otherwise, $x - N_{FDL}$ packets will be lost. Therefore, similar to Eq. (5.3), the PLR of MOPS with shared FDLs can be calculated as:

$$PLR = \frac{\sum_{j=N_{FDL}+1}^{NFW-FW} C(j) \cdot (j - N_{FDL})}{N \cdot F \cdot W \cdot \alpha}. \quad (5.7)$$

Closer observations show that Eq. (5.7) has *not* considered the packets buffered in FDLs in the last time slot. These packets have to be handled altogether with newly arrived packets, which slightly increase the traffic load arriving at the input interface of the switch. To take this part of “additional loads” into account, the average arrival rate of data packets destined for each output port can be approximately adjusted as:

$$\alpha' = \alpha + \frac{\sum_{i=1}^{NFW-FW} C(i) \cdot \min(i, N_{FDL})}{N \cdot F \cdot W} \quad (5.8)$$

where α represents the probability of having a *newly arrived* packet destined for each output port. In other words, the effects of the buffered traffic are approximated as increasing the average arrival rate of the traffic in each time slot. The above analytical

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

model requests iterative adjustments of α' . In all our experiences, however, calculating α' once can already lead to highly accurate analytical results.

Remark 1: Note that in the above analysis, the negative correlation between the numbers of packets destined for different output ports (Specifically, the effect that more packets destined for some output ports leads to fewer packets destined for the others.) has been neglected, which may affect the accuracy of the analysis. In this study, we nevertheless choose to neglect it, for two reasons: (1) the neglect significantly simplifies the analytical models and their extensions (e.g., the extension for analyzing MOPS with both FDLs and SWCs that later will be reported in Section III-D); (2) the side effects of the neglect are generally not significant, especially in large-sized switches or under low traffic load [63]. As later we will see in Section IV, the simplified analytical model appears to be quite accurate under most cases. Comprehensive studies on the analytical models with the correlation effect taken into account can be found in [63], [127].

Remark 2: In the above analysis, we assumed that each FDL is of the same length that can buffer a packet for one time slot. The analytical model can be extended to analyzing MOPS with variable-depth buffers where different FDLs have different lengths. Specifically, we only need to revise Eq. (5.8) to reflect the fact that some packets buffered in the current time slot will not contribute to increasing α' in the next time slot, but several time slots later instead. The detailed discussions, however, will be rather lengthy, since we need to define the distribution of FDL lengths as well as the scheme for choosing an FDL when there are multiple FDLs (of different lengths) available. Such discussions are therefore omitted.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

C. Analysis of MOPS with SWCs

For MOPS with SWCs, we still conduct analysis under uniform traffic load. Existing analytical models for SWC configuration have satisfactory accuracy yet tend to be quite complicated (e.g., [30]). We propose in this section a simpler, yet still accurate analytical model. The main idea is as follows. In MOPS with SWCs, packet loss can be viewed as two separate parts: those caused by exhausted output-link capacity; and those caused by insufficient number of SWCs. Specifically, when there are K packets destined for the same output link and $K \leq FW$, packet losses (if any) could be viewed as caused by insufficient number of SWCs. If $K > FW$, $K - FW$ packets are dropped because of exhausted output link capacity, while the other packet losses (if any) can still be viewed as caused by insufficient SWCs. Therefore, PLR of MOPS with SWCs can be calculated as a combination of these two parts.

We first calculate the PLR caused by exhausted output-link capacity, which equals to the probability of having more than FW packets destined for the same output link. Since the probability of having x packets destined for the same output link (Note that here each output link contains FW wavelength channels and calculation is based on Eq. (5.4)) equals to

$$P_L(x) = \bigotimes_{i=1}^{FW} P(k) \Big|_{k=x}, \quad (5.9)$$

the packet loss rate caused by exhausted output link capacity can be calculated as:

$$PLR_1 = \frac{\sum_{j=FW+1}^{NFW} P_L(j) \cdot (j - F \cdot W)}{F \cdot W \cdot \alpha}. \quad (5.10)$$

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

To obtain the PLR caused by insufficient number of SWCs, packet loss caused by exhausted output link capacity should be pre-excluded. Therefore, the average arrival rate of data packets to each output port can be adjusted to:

$$\alpha_1 = \alpha \cdot (1 - PLR_1). \quad (5.11)$$

Since each SWC, like an FDL, can resolve one packet contention in each time slot, the total lost packets due to insufficient number of SWCs therefore can be calculated in nearly the same way as that for calculating packet loss due to insufficient FDLs. Specifically, Eqs. (5.2), (5.4), (5.5) and (5.6) can be adopted directly based on α_1 , while Eq. (5.7) needs to be slightly modified as follows:

$$PLR_2 = \frac{\sum_{j=N_{SWC}+1}^{NFW-FW} C(j) \cdot (j - N_{SWC})}{N \cdot F \cdot W \cdot \alpha}, \quad (5.12)$$

where N_{SWC} denotes the total number of SWCs. Since those packet losses due to limited link capacity, reflected by PLR_1 in Eq. (5.10), have been pre-excluded from the calculations of PLR_2 in Eq. (5.12), the two parts of packet losses have no overlap with each other. Therefore, the gross PLR can be obtained by adding up the two parts. In other words:

$$PLR = PLR_1 + PLR_2. \quad (5.13)$$

Remark: An important assumption we adopted in analyzing MOPS with FDLs is the uncorrelated traffic pattern, where traffic arrivals in different time slots are independent from each other. For bufferless switches such as MOPS with SWCs but no FDLs, such an assumption is not necessary since packet losses in such switches are solely decided by traffic arrivals in the current time slot.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

D. Analysis of MOPS with both FDLs and SWCs

When FDLs and SWCs are both implemented in MOPS as shown in Figure 5.1, packet contentions can be resolved in three different dimensions including (i) space dimension of multi-fiber links, (ii) wavelength dimension of SWCs, and (iii) time dimension of FDLs. To develop simple analytical models, we separate the effects of these three dimensions as much as possible. Specifically, we develop in Figure 5.3 an “equivalent” model which closely resembles the packet-loss performance of the general MOPS, but with the effects of space and wavelength dimensions being separated into two different stages. Next we describe the operations in the two-stage model in details (Note that the “equivalent” model is proposed solely to resemble the packet-loss performance of MOPS. The fact that this model is far too expensive and difficult to be implemented in the real world does not affect fulfilling its objective.).

In the first stage, we assume that there are unlimited full-range wavelength converters (FTWC) before *and* after the switching board, shown as “WC” in Figure 5.3. Since the full-range wavelength conversion before the switching board can resolve all the wavelength conflicts, packet losses in this stage, if any, are all caused by the exhausted outgoing link and FDL capacities. The wavelength conversion after the first switching board is responsible for converting all the going-through packets back to their *original* wavelengths, such that wavelength constrain can be taken into account after the packet losses caused by insufficient capacities, which anyway cannot be resolved by SWCs as mentioned, have been excluded. In the second stage, SWCs will be applied to resolve packet contentions caused by wavelength conflicts.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Figure 5.3: The “equivalent” two-stage switch configuration of which the packet-loss performance closely resembles that in MOPS with both FDLs and SWCs.

The missing part in this description is how to use FDLs in the two stages respectively. We let FDLs help resolve packet contentions in the first stage as far as possible before handling packet contentions in the second stage. Specifically, if the number of FDLs is larger than the number of packet contentions in the first stage, we let the number of FDLs installed in the first stage be *equal* to the number of contentions in this stage; otherwise, we let all the FDLs be installed in the first stage. The leftover FDLs, if any, are installed in the second stage. Finally, we assume that the packets buffered in *both* the stages are circulated back to join the newly arriving packets in the next time slot.

The reasonability of introducing the two-stage model lies in the fact that the packet contentions caused by insufficient link capacities cannot be resolved by SWCs. FDLs can resolve those packet contentions caused by insufficient capacities as well as those caused by wavelength conflicts. No matter which type of contention an FDL resolves,

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

the effect is basically the same: a packet is buffered until the next time slot. Therefore we assign resolving the first type of contentions a higher priority. As later we will see, this helps simplify the analysis.

We now calculate the PLR in the first stage. By adopting $P_L(x)$ calculated by Eq. (5.9) in Eq. (5.5), we obtain the number of contentions caused by insufficient capacities as:

$$U_1(x) = \begin{cases} \sum_{k=1}^{FW} P_L(k), & x = 0, \\ P_L(x + F \cdot W), & x = 1, 2, \dots, NFW - FW. \end{cases} \quad (5.14)$$

Combining the N output links, we have the probability of having a total of x contending packets as

$$C_1(x) = \bigotimes_{i=1}^N U_1(k) \Big|_{k=x}. \quad (5.15)$$

Similarly to that in Eq. (5.7), the PLR in the first stage therefore can be calculated as:

$$PLR_f = \frac{\sum_{j=N_{FDL}+1}^{NFW-FW} C_1(j) \cdot (j - N_{FDL})}{N \cdot F \cdot W \cdot \alpha}. \quad (5.16)$$

To take into account the additional traffic coming from the buffered packets, similar to that in Eq. (5.8), the arrival rate of data packets needs to be adjusted as:

$$\alpha_f = \alpha + \frac{\sum_{j=1}^{NFW-FW} C_1(j) \cdot \min(j, N_{FDL})}{N \cdot F \cdot W}. \quad (5.17)$$

We then proceed to analyze the second stage. We assume that the contending packets will firstly exploit wavelength conversion and then look for available FDLs. Therefore, if the number of contending packets in the second stage is larger than N_{SWC} ,

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

some packets need to be buffered if such is feasible. To take this part of buffered traffic into consideration, Eq. (5.17) needs to be further modified to

$$\alpha'' = \alpha_f + \frac{\sum_{j=0}^{N_{FDL}} C_1(j) \cdot \left[\sum_{i=N_{SWC}+1}^{NFW-FW} C(i) \cdot \min(i - N_{SWC}, N_{FDL} - j) \right]}{N \cdot F \cdot W}. \quad (5.18)$$

In this equation, parameter j ($0 \leq j \leq N_{FDL}$) denotes the number of FDLs occupied in the first stage (Therefore $N_{FDL} - j$ FDLs are available in the second stage); and parameter i ($0 \leq i \leq NFW - FW$) denotes the number of the contending packets in the second stage. When $i > N_{SWC}$, $i - N_{SWC}$ packets would look for FDLs. The number of packets that can be actually buffered is therefore $\min(i - N_{SWC}, N_{FDL} - j)$. Furthermore, $C(j)$ in this equation is calculated by replacing α in the calculations of $C_1(j)$ by $\alpha_f(1 - PLR_f)$ to better approximate the traffic arriving the second stage.

In the second and all the following iterations, the modified arrival rate α'' will be adopted in Eq. (5.15) to get more accurate $C_1(j)$, and consequently more accurate PLR_f in (5.16). Meanwhile, the average arrival rate of the traffic reaching the second stage would be calculated as:

$$\alpha_2 = \alpha'' \cdot (1 - PLR_f). \quad (5.19)$$

In the second and all the following iterations, α_2 , instead of $\alpha_f(1 - PLR_f)$, will be used to calculate $C(i)$ more accurately.

Finally, the packet loss rate in the second stage can be calculated as:

$$PLR_s = \sum_{j=0}^{N_{FDL}} C_1(j) \cdot F_1(N_{SWC}, N_{FDL} - j) + \sum_{j=N_{FDL}+1}^{NFW-FW} C_1(j) \cdot F_2(N_{SWC}), \quad (5.20)$$

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

where $F_1(N_{SWC}, N_{FDL} - j)$ and $F_2(N_{SWC})$ denote the PLRs with $(N_{FDL} - j)$ FDLs and no FDLs respectively. In both cases, there are N_{SWC} shared wavelength converters. We have

$$F_1(N_{SWC}, N_{FDL} - j) = \frac{\sum_{i=N_{SWC}+N_{FDL}-j+1}^{NFW-FW} C(i) \cdot [i - N_{SWC} - (N_{FDL} - j)]}{N \cdot F \cdot W \cdot \alpha}, \quad (5.21)$$

and

$$F_2(N_{SWC}) = \frac{\sum_{i=N_{SWC}+1}^{NFW-FW} C(i) \cdot (i - N_{SWC})}{N \cdot F \cdot W \cdot \alpha}. \quad (5.22)$$

Since PLR_f and PLR_s occur in two different stages with no overlapping, we have

$$PLR = PLR_f + PLR_s \quad (5.23)$$

Remark: The values of α_1 and α_2 need to be iteratively updated. In all our experiences, the results converge in no more than three iterations. As later we will see in Section IV, accuracy of the analytical models appears to be reasonable.

Finally, we summarize our main contributions in analytical model developments as follows: For the pure MOPS, we extended the existing analysis for the classic single-fiber switches to MOPS, under both uniform and non-uniform traffic load. For MOPS with FDLs, SWCs or both of them under uniform traffic load, by carefully applying the convolution method, we developed some new analytical models. These new models are rather simple yet, as later we will see, still highly accurate. Analyzing resource sharing MOPS configurations under non-uniform traffic, however, remains as a big challenge.

5.4 Numerical Results and Discussions

In this section, we evaluate the packet loss performance of single- and multi-board MOPS, under both uniform and non-uniform traffic load. For non-uniform traffic load, we test the case with geometrically-distributed traffic pattern.

In single-board MOPS, the high accuracy of the proposed analytical models is verified by comparing to the simulation results. Since the same models apply to multi-board MOPS, to save computation time especially when under light traffic load, the multi-board MOPS is evaluated by analytical solutions.

5.4.1 Single-Board MOPS

Performance of the single-board pure MOPS under uniform traffic load is demonstrated in Figure 5.4. We let $N = 4$ and $F \times W$ be fixed at 32, where F varies from 1 to 16 and W from 32 to 2 accordingly. The worst performance happens in single-fiber switch ($F = 1$) where the space dimension for contention resolutions is not available. With an increasing number of fibers per link and a decreasing number of wavelength channels per fiber (while keeping the product of them as a constant), the performance improves significantly. However, *unlike* that in circuit-switched networks, even when $F = 16$ and $W = 2$, pure MOPS still cannot perform nearly the same as OPS with FTWC. This is because, in OPS networks we cannot make end-to-end wavelength assignment decisions based on global link-state information. As a result, the space dimension provided by multi-fiber links cannot be fully exploited: we may have exhausted capacities and consequently have packet losses on some wavelengths while on other wavelengths idle capacities are still available. The high accuracy of the

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

analytical model for the pure MOPS (Eq. (5.3)) is also verified in Figure 5.4, where analytical and numerical results perfectly match each other.

Figure 5.4: PLRs of single-board pure MOPS under uniform traffic load.

As shown above, in pure MOPS, PLRs on different wavelengths are independent of each other, with the overall PLR a simple linear combination of them. Therefore in the evaluations under non-uniform traffic load, we simulate only on a single wavelength. We present in Figure 5.5 the results for the case where traffic load destined for different output links is distributed as a geometrical sequence such that $\rho_i / \rho_{i-1} = R = 1.2$, $i = 2, 3, \dots, N$. In this and all the other cases that we have tested, the conclusion is that MOPS achieve much better performance than single-fiber pure OPS, but not as good as OPS with FTWC.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Figure 5.5: PLRs of single-board pure MOPS under non-uniform, geometrically-distributed traffic load where $R = 1.2$.

Single-board MOPS with FDLs under uniform traffic load is evaluated in Figure 5.6, where the switch configuration is set as $N = 4$, $F = 4$ and $W = 8$. We see that with an increasing number of FDLs (N_{FDL}), PLR drops quickly. When $N_{FDL} = 12$, MOPS steadily outperforms the switch with FTWC throughout the region of $\alpha < 0.6$, i.e., as long as the normalized traffic load per channel is lower than 0.6. Meanwhile, the high accuracy of the analytical model in Eq. (5.7) is also verified. Simulation results under geometrically-distributed non-uniform traffic load where $R = 1.2$ are presented in Figure 5.7. We see that MOPS with 12 FDLs achieves lower PLR than OPS with FTWC throughout the region where the normalized average traffic load per channel is $\alpha < 0.5$.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Figure 5.6: PLRs of single-board MOPS with FDLs under uniform traffic load.

Figure 5.7: PLRs of single-board MOPS with FDLs under non-uniform traffic load where $R = 1.2$.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Simulation and analytical results of MOPS with SWCs under uniform traffic load are presented in Figure 5.8, where we still let $N = 4$, $F = 4$ and $W = 8$. We see that when the number of SWCs is increased from one to four, PLR is lowered significantly. The improvements, however, become less significant when the number of SWCs is increased from eight to sixteen. Such an observation is verified in Figure 5.9, where we plot PLRs of MOPS with different number of SWCs. We see that the improvement in PLR performance becomes less significant with an increasing number of SWCs, until finally MOPS performs the same as OPS with FTWC. More significantly, we see that having more fibers per link with fewer wavelength channels per fiber lowers the number of SWCs required to achieve the same performance as the FTWC scheme. For example, when $F = 1$ and $W = 32$, 28 SWCs are needed to achieve the same performance as OPS with FTWC. This number can be reduced to 12 in a multi-fiber network where $F = 4$ and $W = 8$.

Figure 5.8: PLRs of MOPS with SWCs under uniform traffic load.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Figure 5.9: PLRs of MOPS with different number of SWCs.

Under non-uniform traffic load, as we see in Figure 5.10, the conclusion holds that having more fibers per link helps improve packet loss performance and save SWCs. Finally, we point out that the satisfactory accuracy of the analytical model in Eqs. (5.9) to (5.13) is verified in Figure 5.8.

The performance of MOPS with SWCs can also be evaluated by the analytical model proposed in [30]. The comparisons between the computation time needed by the two models respectively are shown in Table 5.1, where Model 1 refers to the model in [30] and Model 2 the new model. The computation time of each model is averaged over 10^3 simulations on a computer with 1.5GHz CPU and 760MB memory. We see that our model is typically one-order faster than the existing one. More significantly, the Markovian model proposed in [30] would be difficult to be extended to MOPS with

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

FDLs or MOPS with both FDLs and SWCs. By using the proposed model, as we have seen in Section 5.3, such extensions are rather simple and straightforward.

Figure 5.10: PRLs of MOPS with SWCs under non-uniform, geometrically-distributed traffic load where $R = 1.2$.

	$N_{SWC} = 1$	$N_{SWC} = 4$	$N_{SWC} = 8$	$N_{SWC} = 16$
Model 1	629ms	629ms	630ms	630ms
Model 2	55 ms	55 ms	56 ms	56 ms

Table 5.1: The computation time needed by the two analytical models respectively.

The performance of MOPS with SWCs can also be evaluated by the analytical model proposed in [30]. The comparisons between the computation time needed by the two models respectively are shown in Table 5.1, where Model 1 refers to the model in [30] and Model 2 the new model. The computation time of each model is averaged over

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

10^3 simulations on a computer with 1.5GHz CPU and 760MB memory. We see that our model is typically one-order faster than the existing one. More significantly, the Markovian model proposed in [30] would be difficult to be extended to MOPS with FDLs or MOPS with both FDLs and SWCs. By using the proposed model, as we have seen in Section 5.3, such extensions are rather simple and straightforward.

Figure 5.11: PLRs of MOPS with both FDLs and SWCs under uniform traffic load.

For the most general MOPS model with both FDLs and SWCs, we show in Figure 5.11 both the simulation and the analytical results under uniform traffic load. The node configuration remains as $N=4, F=4$ and $W=8$, while the numbers of FDLs and SWCs increase from 2 to 8 respectively. We see that reasonable accuracy has been achieved by the proposed analytical model.

5.4.2 Multi-Board MOPS

The node configuration of multi-board MOPS is proposed in Figure 5.2. The main objectives of our studies include evaluating the impacts of switching-board size on PLR; and for a given board size, the effects of having different numbers of fibers and FDLs connected to the switching board respectively. We assume that each board is connected to T channels of each input/output link on the same wavelength. Since the overall packet loss is simply a linear combination of those on different switching boards, we conduct calculations on only a single switching board.

Firstly, we evaluate PLRs with different switching-board sizes. We present here only the analytical results (the accuracy of which has been verified in Section IV-A) under uniform traffic load. The conclusions presented here also hold in all the non-uniform cases we have tested. Figure 5.12 presents the results under light ($\alpha = 0.1$) and moderate traffic load ($\alpha = 0.5$), respectively. We see that under light traffic load (Figure 12(a)), larger values of T steadily lead to lower PLRs, while improvements achieved by larger values of N_{FDL} are even more significant. Under heavier traffic load, however, the improvements, either by larger values of T or larger values of N_{FDL} , become drastically less significant. In fact, as we can see in Figure 5.12(b), when T increases from 1 to 2 in a switching board with $N_{FDL} = 6$, PLR performance becomes even slightly worse. This is because that under light traffic load, there are generally fewer than T packets destined for the same output link in a switching board with T fibers per link, especially when T is of a large value. Therefore, a larger value of T lowers PLR. For the occasional cases where there are more than T packets destined for the same output link, a limited number of FDLs can generally resolve the contentions.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

(a) Light traffic load ($\alpha = 0.1$).

(b) Moderate traffic load ($\alpha = 0.5$)

Figure 5.12: PLRs of multi-board MOPS with different switching-board configurations.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

Under heavier traffic load, on the other hand, with T fibers per link there is a good chance that there are more than T packets destined for the same link. Therefore, the improvements made by larger T or larger N_{FDL} become less significantly. In some cases where we increase the value of T without installing more FDLs, there are actually more contending packets competing for the limited FDL resources, which leads to a higher PLR.

Figure 5.13: Average buffer size per fiber vs. number of fibers per link.

A supportive observation may be made in Figure 5.13: when we increase the number of fibers per board, the average buffer size per fiber (N_{FDL}/T) decreases quickly at the beginning, from $T = 1$ to $T = 4$. Then the changing speed becomes significantly slower with further increased values of F . This helps to explain the fluctuations in some PLR curves.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

(a) $\alpha = 0.7$

(b) $\alpha = 0.9$

Figure 5.14: Performance of MOPS with FDLs under very high traffic load.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

As we have seen in Figure 5.12, the performance improvement by increasing the number of fibers per board under heavy traffic load ($\alpha = 0.5$) is not as significant as that under light load ($\alpha = 0.1$). Further evaluations under extremely high loads ($\alpha = 0.7$, $\alpha = 0.9$) can be seen in Figure 5.14 (a) and 5.14 (b), where in both cases very limited benefits have been achieved by multi-fiber links. In fact, in such cases, even an increasing number of FDLs make no significant differences. This can be easily understood: since the benefits of MOPS mainly come from the free channels over multiple fibers, most of the packet contentions cannot be resolved where these channels have already be heavily occupied.

Figure 5.15: Performance of MOPS with a fixed total number of FDLs.

Now we consider the case where the total number of FDLs in the switching node is fixed. Under such case, a larger value of T reduces the number of switching boards, which consequently increases the number of FDLs on each board. We assume that $N =$

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

4, $F = 8$ and the total number of FDLs on each wavelength is 8. For $T = 1, 2, 4, 8$, we have that the number of switching boards is 8, 4, 2, 1, with $N_{FDL} = 1, 2, 4, 8$ on each board respectively. Simulate results are presented in Figure 5.15. We see that a larger board size with more FDLs attached to each board leads to much better performance. However, the large switching board leads to much higher cost as well.

In Figures 5.12-15, we have assumed that the switching-board size can be increased when necessary. In real-world implementations, however, we may have very limited freedom in choosing the switching-board size from only several options, e.g., 8×8 , 16×16 , etc. For a given switching-board size, an important issue is to decide how many fibers and FDLs should be connected to each board, i.e., the values of T and N_{FDL} , respectively. The switching board size S is given by $S = N \cdot T + N_{FDL}$. As we have seen in Figure 5.12, higher values of T and N_{FDL} generally lead to lower PLR. When both values cannot be set high simultaneously due to the given switching-board size, we expect that having more FDLs and fewer fibers per board would lead to lower PLR. This is because, more FDLs provide more contention-resolution capacities while fewer fibers carrying lower traffic load to each board. However, connecting fewer fibers to each board increases the number of switching boards to support a given multi-fiber network and the control burdens may be increased as well.

To better understand how to keep a balance between cost and performance, we plot in Figure 5.16 the performance of two fixed-sized switching boards where S is set as 64 and 128, respectively. Numerical results confirm that larger N_{FDL} with smaller T leads to lower PLR. Moreover, the performance gains achieved by having larger N_{FDL} are

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

(a) Switching board size: 64

(b) Switching board size: 128

Figure 5.16: PLRs of multi-board MOPS with fixed switching-board sizes.

Chapter 5: Performance Evaluation of Multi-fiber Optical Packet Switches

more significant under lower traffic load, regardless of the switching-board size. On the other hand, even under heavy traffic load like $\alpha = 0.7$, 12 to 16 FDLs connected to each board are sufficient to achieve a PLR lower than 10^{-12} . Therefore, for most applications, the number of FDLs per board does not need to be very large, while for some special applications, extremely low PLR (lower than 10^{-20}) could be achieved when necessary. More studies in order to figure out the efficient board configurations in MOPS and the trade-off between the performance and cost will be presented in the next chapter.

5.5 Conclusion

In this chapter, we evaluated the performance of several typical MOPS configurations, with or without shared contention-resolution resources respectively. Simple yet highly accurate analytical models were proposed and extensive simulations were conducted as well. Analytical and simulation results showed that, by exploiting the space dimension of multi-fiber links, MOPS significantly outperforms the classic single-fiber OPS. With a limited number of FDLs and SWCs, MOPS even outperforms OPS with FTWC. For MOPS with a fixed number of FDLs, a larger switching-board size leads to much better performance at a higher cost. For MOPS containing multiple fixed-sized switching boards, we have shown how their performance can be affected by the switching-board size, as well as the number of fibers and FDLs connected to each switching board respectively. These evaluations help to achieve some insights highly helpful for the future developments of cost-effective MOPS networks.

Chapter 6

Node Configurations in Multi-fiber OPS Networks

With the development of optical packet switching (OPS) technologies, multi-fiber OPS networks will finally play an important role in the future data transmissions. In such networks, instead of constructing some extremely expensive node configurations with strictly non-blocking switching function, a more practical solution may adopt multi-board architecture containing a number of small-sized switching boards. In this chapter, we evaluate the performance of several different multi-board switches, based on the following two main objectives: (i) to better understand the effects of different connection schemes between switching boards and optical buffers; and (ii) to investigate possible multi-board schemes for achieving comparable performance to that of the ideal, strictly non-blocking switches.

6.1 Introduction

With the deployments of OPS networks, we expect that a significant part of them will be multi-fiber systems, where each link contains multiple optical fibers, as that in today's circuit-switched core networks [5], [6], [15]. Contention resolution in multi-fiber OPS (MOPS) networks is more complicated than that in single-fiber networks, where a main challenge is imposed by the limited switching-board size. Specifically, in the foreseeable near future, manufacturing large-sized, high-speed optical packet-switching boards will be difficult and expensive [15], [20]. Hence in multi-fiber networks where each link contains a large number of fibers, we may not be able to connect all the incoming/outgoing channels (even only those on the same wavelength) to a single switching board, but to multiple smaller-sized switching boards instead.

The reduced switching-board size has nontrivial impacts to the packet-loss performance of MOPS networks. Specifically, with multiple channels on the same wavelength in each link, a "space dimension" emerges [121], [128], [129], providing an efficient solution for contention resolutions (i.e., multiple packets having the same wavelength and destined output link can be carried out immediately). When the switching-board size is limited, the space dimension is restricted. Consequently, there may be idle output ports in a certain switching board while another board still has many contending packets.

Various schemes have been proposed to solve this problem [57], [107], [120], [122], [130]. The main idea is typically to interconnect a large number of small-sized switching boards, generally into a multistage architecture, such that strictly or

Chapter 6: Node Configurations in Multi-fiber OPS Networks

conditionally non-blocking switching function could be fulfilled [57], [120]. The major drawback, however, is that the multistage solutions generally lead to complicated control and degraded signal quality [57], [107]. Such drawbacks can be serious in OPS networks where all the control operations have to be conducted at very high speed to handle the data traffic at packet granularity [5], [122], [130].

In circuit-switched networks, it has been shown that with proper designs, a single stage of small-sized switching boards can achieve rather good performance [118]-[120]. In fact, by slightly increasing the number of fibers per link, a single-stage design can even outperform the ideal, strictly non-blocking switch [120]. In packet-switched networks, to the best of our knowledge, there is no existing study on whether similar conclusion still holds or not.

We are motivated to investigate the performance of various *single-stage* node configurations in MOPS networks. Specifically, we study the effects of different (i) switching-board sizes and the numbers of fibers/FDLs connected to each board; (ii) schemes for connecting the incoming/outgoing fibers to the switching boards; and (iii) schemes for sharing FDLs between different switching boards. We find that the multi-board switches *cannot* achieve the same performance as the strictly non-blocking switch by slightly increasing the number of fibers per link. To approach the ideal performance, solutions with higher cost, such as increasing the switching-board size or the number of FDLs connected to each board or enhancing the sharing of FDLs between different switching boards, have to be considered. And under various load, different solutions show diverse effects.

6.2 Node Configurations

The simplest single-board MOPS node configuration is represented in Figure 6.1. As mentioned before, with multiple fibers per link, the space dimension greatly facilitates contention resolutions in the switch. Meanwhile, however, the increasing switching-board size becomes a challenge. Assume that each of the N incoming/outgoing links contains F fibers with W wavelength channels per fiber, the switching-board size would be as large as $NFW \times NFW$. Even if we install W boards, each of which operating on a single wavelength, the switching-board size of $NF \times NF$ can still be rather large, especially when F is of a large value (Meanwhile N is typically not of a big value since optical networks are generally sparsely connected.).

Figure 6.1: A schematic representation of the ideal single-board switch.

To solve this problem, a single-stage multi-board architecture could be adopted, as schematically illustrated in Figure 6.2(a). Specifically, we have multiple *transmission-*

Chapter 6: Node Configurations in Multi-fiber OPS Networks

switching boards (TSBs), each of which is connected to a number of incoming/outgoing channels. Meanwhile, each TSB has some of its input/output ports reserved for the *FDL switching boards* (FSBs), where a certain number of FDLs are installed. Each FSB is connected to multiple TSBs, so that its FDL capacities can be shared among TSBs, as shown in Figure 6.2(b). The specific connection schemes will be proposed later in Section 6.2.2. By doing so, each FSB also provides “interconnections” between different TSBs, such that a heavily loaded TSB could diverse its traffic to a less-loaded one. The advantages of such architecture include relatively simple node configuration and low degradation to signal quality. More significantly, it has good scalability for supporting future extensions [130]: when more fibers are lightened in each link, we only need to insert additional switching boards rather than replacing the big board(s) altogether. The packet-loss performance of such node configurations, however, needs to be carefully evaluated, which is the main objective of this chapter.

6.2.1 Different Connection Schemes: Main Idea

We study several schemes with reasonable main ideas and, in our experiences, lead to better packet-loss performance. We first introduce two different schemes for connecting incoming/outgoing fibers to TSBs (referred to as *TSB schemes* hereafter). Throughout this chapter, unless otherwise specified, we assume that different wavelengths are processed separately therefore the TSBs and FSBs under evaluation are operating on the same wavelength. Hence wavelength converters are not needed. The objectives of the two schemes are presented as follows:

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) The main architecture

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(b) The connections between TSBs and FSBs

Figure 6.2: A schematic representation of the multi-board switch.

1. The first TSB scheme has the objectives of (i) maximizing the number of different incoming/outgoing links connected to each TSB; and (ii) maximizing the number of TSBs with all their input/output ports fully utilized. When the TSB size is larger than the number of links it lets each TSB to be connected to the same number of fibers per link as far as possible.
2. The second TSB scheme has the objectives of (i) maximizing the number of different TSBs connected to each link; and (ii) utilizing evenly each TSB as far as possible. When the TSB size is larger than the number of links, it still lets each TSB to be connected to the same number of fibers per link as far as possible.

Remark 1: In both schemes, the first objective aims to maximize the number of outgoing links connected to each input link through the limited-sized switching boards. The same objective has different descriptions in different schemes because of the existence of the second objective. As later we will see, in most cases, we cannot

Chapter 6: Node Configurations in Multi-fiber OPS Networks

simultaneously achieve objective (i) of the first scheme and objective (ii) of the second scheme, and vice versa.

Remark 2: The second objective of the first TSB scheme is to have the largest number of incoming fibers each of which being connected to the maximum number of outgoing fibers (which equals to the TSB size). The second objective of the second TSB scheme, on the other hand, is to avoid having too few fibers connected to a certain switching board; which may cause high packet losses on that board. The two schemes are hence different when there are some idle input/output ports left on TSBs. For example, assume that we have totally 20 incoming/outgoing channels connected to 4 TSBs, each of which has 6 input/outputs reserved for such connections. In the first TSB scheme, we shall fully utilize all the 6 ports each of the first three TSBs and 2 ports of the fourth one. In the second TSB scheme, on the other hand, we shall utilize 5 ports each of the four TSBs.

The two different schemes for connecting TSBs to FSBs (referred to as *FSB schemes* hereafter) that we will evaluate in this chapter are similar to the TSB schemes. In fact, their objectives and main ideas could be defined simply by interchanging in the above descriptions the words “incoming/outgoing links” to “TSBs”, and “TSB” to “FSB”. Therefore this allows us to remove similar discussions.

6.2.2 Different Connection Schemes: Formal Definitions

Assume each TSB has M^T input/output ports, M_1^T of which are reserved for connecting to incoming/outgoing fibers, and M_2^T for connecting to FSBs ($M_1^T + M_2^T = M^T$). The TSB schemes can then be defined as follows:

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Figure 6.3: Examples of the two different TSB schemes.

The first TSB scheme: Index the j -th fiber ($j = 1, 2, \dots, F$) of the i -th link ($i = 1, 2, \dots, N$) as $k = (j-1) \cdot N + i$. Let the fiber k be connected to the K -th TSB where

$$K = \left\lceil \frac{k}{M_1^T} \right\rceil, \quad k = 1, 2, \dots, NF. \quad (6.1)$$

A simple example where $N = F = 3$ and $M_1^T = 4$ is presented in Figure 6.3(a).

The second TSB scheme: Index the j -th fiber ($j = 1, 2, \dots, F$) of the i -th link ($i = 1, 2, \dots, N$) as $k = (i-1) \cdot N + j$. Let the k -th fiber be connected to the K -th TSB where

$$K = \text{mod}(k-1, N_{TSB}) + 1, \quad k = 1, 2, \dots, NF. \quad (6.2)$$

Here N_{TSB} denotes the total number of TSBs,

Chapter 6: Node Configurations in Multi-fiber OPS Networks

$$N_{TSB} = \left\lceil \frac{NF}{M_1^T} \right\rceil. \quad (6.3)$$

A simple example where $N = F = 3$ and $M_1^T = 4$ can be found in Figure 6.3(b).

It is trivial yet lengthy to prove that the two schemes fulfill their respective objectives. Therefore we omit the proofs. In addition, Figure 6.3 also demonstrates the differences between the two schemes when there are idle ports left on TSBs.

The two FSB schemes can be defined similarly. Specifically, denoting the size of each FSB as $M^F \times M^F$, we have:

The first FSB scheme: Index the j -th input/output port ($j = 1, 2, \dots, M_2^T$) connecting the i -th TSB ($i = 1, 2, \dots, N_{TSB}$) to FSB as $k = (j-1) \cdot N_{TSB} + i$. Let input/output port k be connected to the K -th FSB where

$$K = \left\lceil \frac{k}{M^F} \right\rceil, \quad k = 1, 2, \dots, N_{TSB} \cdot M_2^T. \quad (6.4)$$

The second FSB scheme: Index the j -th input/output port ($j = 1, 2, \dots, M_2^T$) connecting the i -th TSB ($i = 1, 2, \dots, N_{TSB}$) to FSBs as $k = (i-1) \cdot N_{TSB} + j$. Let input/output port k be connected to the K -th TSB where

$$K = \text{mod}(k-1, N_{FSB}) + 1, \quad k = 1, 2, \dots, N_{TSB} \cdot M_2^T. \quad (6.5)$$

Here N_{FSB} denotes the number of FSBs,

Chapter 6: Node Configurations in Multi-fiber OPS Networks

$$N_{FSB} = \left\lceil \frac{N_{TSB} \cdot M_2^T}{M^F} \right\rceil. \quad (6.6)$$

Examples are omitted since they would look virtually the same as those illustrated in Figure 6.3.

6.3 Evaluations of Different Connection Schemes

We evaluate the *packet loss* of OPS, defined as the *statistic average number* of the lost packets in each time slot. We test three different traffic patterns:

- Uniform, where packets are uniformly destined for the different output links, i.e., $\rho_1 = \rho_2 = \dots = \rho_N$, where ρ_i denotes the probability that a newly-arrived packet is destined for output link i .
- Geometric, where $\rho_i / \rho_{i-1} = R$, $i = 2, 3, \dots, N$.
- Hotspot, where traffic destined for the first output link is 3 times as high as that for the other output links, i.e., $\rho_1 = 3\rho_i$, $i = 2, 3, \dots, N$.

The simulation results reported in this section are mainly for evaluating the effects of the different TSB and FSB schemes, respectively. Unless otherwise specified, we assume that $N = 4$ and $F = 20$ where the size of each TSB and FSB is 8×8 . In each TSB, we assume that $M_1^T = 6$ and $M_2^T = 2$. Simple calculations show that we need 14 TSBs. In the first TSB scheme, 13 of them will be fully utilized (i.e., being connected to 6 incoming/outgoing channels), while the last one is connected to only 2 fibers. In the second TSB scheme, 10 TSBs will be fully utilized; while each of the other 4 TSBs is connected to 5 fibers.

6.3.1 Effects of Different TSB Schemes

Figure 6.4 presents the simulation results under three different traffic patterns. The x -axis parameter ρ denotes the normalized average traffic load destined for each output port. The y -axis denotes the average number of dropped packets in the whole switch in each time slot. To focus on evaluating the TSB schemes, we let all the FDLs be disconnected from TSBs. From the simulation results, we see that under light traffic load where $\rho \leq 0.2$, the second TSB scheme outperforms the first one under all the three different traffic patterns. The differences, however, reduce to virtually zero under heavy traffic load. This is because that under light traffic load, the last TSB in the first scheme, with only two incoming/outgoing fibers connected to it, suffers a relatively large number of packet losses since the destination links of some incoming packets cannot be reached at all. The slightly lower packet losses in the first 13 TSBs cannot make enough compensation to such losses. Under heavy traffic load, on the other hand, most packet losses are caused by insufficient capacities in the output link rather than the limited flexibility in one (out of a total of 14) TSB. Therefore, the two schemes converge to perform nearly the same. In fact, in all our experiences with different parameter settings, unless we let the switching-board size be smaller than N (which is unlikely to happen in the real world as discussed previously), the first TSB scheme has never outperformed the second one.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic: $R = 1.2$.

(c) Hotspot traffic.

Figure 6.4: Comparisons between the two different TSB schemes (without FDLs).

Chapter 6: Node Configurations in Multi-fiber OPS Networks

The second TSB scheme continues to outperform the first one when FDLs are attached to TSBs, as we can see in Figure 6.5. Closer observations show that the differences between the two schemes become smaller under light traffic load. This is because FDLs help to lower the packet losses in the last TSB connected to a smaller number of incoming/outgoing fibers. Under medium traffic load (roughly where $0.2 \leq \rho \leq 0.6$), however, the difference between the two schemes' packet-loss performance is larger than that in the case without FDLs. This can be explained as follows: in the first TSB scheme, the last TSB extensively utilizes FDLs, which degrades the system performance when FDLs are not as redundant as that under light traffic load. Under heavy traffic load, once again, different TSB schemes perform nearly the same.

6.3.2 Effects of the Different FSB Schemes

Unless otherwise specified, hereafter we always adopt the second TSB scheme since it outperforms the first one. We firstly justify that it is worth the efforts to install FSBs such that FDLs can be shared among different TSBs. Figure 6.6 compares between the non-shared buffering scheme where two FDLs are dedicatedly attached to each TSB (such that FSBs are not needed), and the second FSB scheme. We see that the existence of FSBs obviously improves the packet-loss performance under various traffic patterns thanks to the more efficient utilizations of FDLs it has achieved.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) The first FSB scheme.

(b) The second FSB connection scheme.

Figure 6.5: Comparisons between the two different TSB connection schemes (with FDLs).

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic: $P = 1.5$.

(c) Hotspot traffic.

Figure 6.6: Comparisons between the different multi-board node configurations where FDLs are shared and not shared among different TSBs respectively.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) The first TSB scheme.

(a) The second TSB scheme.

Figure 6.7: Comparisons between the different FSB schemes.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Comparisons between the two different FSB schemes are presented in Figure 6.7, where 28 FDLs are installed in 4 FSBs. In the first FSB scheme, three FSBs are fully installed with 8 FDLs, while the fourth one is installed with only 4 FDLs. In the second scheme, each FSB is installed with 7 FDLs. Simulation results show that, unlike the case in the TSB schemes, different FSB schemes always perform nearly the same. In other words, having fewer FDLs in the last FSB does not degrade packet-loss performance significantly. This can be understood since the smaller size of the last FSB, after all, does not make any output link inaccessible to any incoming packets (as long as the TSB size is large enough). The less significant degradation therefore can be compensated by having slightly more FDLs installed in the other FSBs.

Unless otherwise specified, hereafter we adopt the second TSB scheme together with the second FSB scheme when simulating multi-board switches.

6.4 Schemes for Improving Switch Performance

As mentioned earlier, in circuit-switched WDM networks with global link-state information and proper routing-and-wavelength-assignment (RWA) algorithm, a single-stage multi-board switch can perform comparable to the strictly non-blocking switch, or even outperforms the latter one by having slightly more fibers per link [120]. In OPS networks, however, since global planning of data packet transmissions is not feasible, the conclusions become quite different. Specifically, we have

- The multi-board switch *cannot* easily achieve nearly the same performance as the ideal single-board switch. As we can see in Figure 6.8, even when there are only two boards, each of which with a large size of 40×40 , the performance is still

Chapter 6: Node Configurations in Multi-fiber OPS Networks

significantly worse than that of the single-board switch (In both switches, we assume that there is no FDL.). When the switching-board size is reduced to 8×8 , even with 8×8 FSBs, the performance of the multi-board switch basically is not comparable to that of the ideal switch with no FDL.

- The problem *cannot* be solved by slightly increasing the number of fibers per link. In Figure 6.9, we show that the performance of a multi-board switch with $F = 20$ and FSBs cannot compete against an ideal switch with $F = 15$ and no FDL, though the average traffic destined for each output port is 25% lower in the former case.

The above examples evidently show that the space dimension provided by the multi-fiber links in an ideal switch is of tremendous effects. When such effects are depressed by the reduced switching-board size, the degradation in packet-loss performance cannot easily be compensated by lowering the average traffic load in each fiber. To explore possible solutions, we evaluate three different cases as follows:

- The TSB size (M^T) is enlarged by increasing M_1^T with FDLs on board (For the case without FDLs, it has been shown in Figure 6.8 that a multi-board switch cannot perform comparably to an ideal switch);
- The FSB size (M^F) is enlarged, such that more flexible sharing of FDLs can be achieved;
- Each TSB has more input/output ports connected to FSBs, or in other words, the total buffer size is increased.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic ($R = 1.5$).

(c) Hotspot traffic.

Figure 6.8: Comparisons between the multi-board switches and the ideal single-board switch.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Figure 6.9: Effects of increasing the number of fibers per link,

$$\text{where } M^T = M^F = 8 \text{ and } M_2^T = 2.$$

Obviously, the above changes can be adopted simultaneously. To better understand the effects of each of them, in the following simulations, each time we only change a single parameter among M_1^T , M^F and M_2^T while keeping the others at fixed values.

Figure 6.10 evaluates the effects of having an increasing TSB size while M_2^T and M^F are fixed at 2 and 8 respectively. We test four different cases where $M_1^T = 8, 16, 20$, and 40, respectively. When the TSB size is increased, the number of FSBs is actually decreased. Nevertheless, we observe in Figure 6.10 that the enlarged space dimension outweighs the negative effects imposed by the reduced buffer size. The packet losses are lowered with a bigger TSB size. On the other hand, even when $M_1^T = 40$, the performance of the multi-board switch is not comparable to the ideal switch without FDLs.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic ($R = 1.5$).

(c) Hotspot traffic.

Figure 6.10: Effects of increasing TSB size, where $M^F = 8$ and $M_2^T = 2$.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic ($R = 1.5$).

(c) Hotspot traffic.

Figure 6.11: Effects of increasing FSB size, where $M^T = 8$ and $M_2^T = 2$.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Figure 6.11 demonstrates the switch performance with an increasing FSB size (while the TSB parameters are fixed at $M_1^T = 6$ and $M_2^T = 2$). We test three different cases where the FSB size is set as 8, 14 and 28, respectively. From the simulation results, we observe that, though larger FSBs help to improve switch performance, they do not provide an effective solution for approaching the performance of the ideal switch. In fact, under light traffic load ($\rho < 0.3$), variant FSB sizes make no significant difference in packet-loss performance at all: while free FDLs are mostly available under light traffic, the additional flexibility of sharing buffers in larger FSBs does not really help too much. The improvements are only significant under heavier traffic load.

To evaluate the effects of increasing buffer size, i.e., having larger values of M_2^T , we simulate the case where both TSBs and FSBs have a size of 8. As we can see in Figure 6.12, when $M_1^T = 6$ and $M_2^T = 2$, the packet-loss performance of the switch is much worse than that of the ideal switch. When we increase the buffer size to $M_2^T = 3$ and $M_1^T = 5$, the performance is significantly improved. In fact, the multi-board switch outperforms the ideal switch under geometric and hotspot traffic patterns when the traffic load is high enough ($\rho \geq 0.3$). If we further increase the buffer size to $M_2^T = 4$ and $M_1^T = 4$, the multi-board switch almost always outperforms the ideal switch, lowering packet losses by several orders most of time. The major drawback of such a solution, however, is that the numbers of TSBs and FSBs are increased. In the example case we have simulated, when M_2^T is increased from 2 to 4, the number of TSBs is increased from 14 to 20, and the number of FSBs from 4 to 10. The cost of having more switching boards has to be carefully weighted when building up a large switch.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

(a) Uniform traffic.

(b) Geometric traffic ($R = 1.5$).

(c) Hotspot traffic.

Figure 6.12: Effects of increasing M_2^T , where $M^T = M^F = 8$.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Figure 6.13: Trade-off between the packet-loss performance and the switch complexity ($M^T \times M^T$) with increasing values of M_2^T .

Finally, we study in further details the trade-off between the packet-loss performance and the switching-board complexity. We consider an OPS with a fixed value of $M_1^T = 8$. The complexity of TSB is defined as $M^T \times M^T$, where $M^T = M_1^T + M_2^T$. As we can see in Figure 6.13, when the complexity of TSB increases as a polynomial function of the buffer size M_2^T , the improvements to the packet-loss performance appear to be load-dependent. Specifically, more significant improvements can be made under lower traffic load. For example, when M_2^T increases from 0 to 7,

Chapter 6: Node Configurations in Multi-fiber OPS Networks

the switch complexity goes up from 64 to 225. Under light traffic load where $\alpha = 0.1$, the packet loss is drastically reduced from above 10^{-2} to be lower than 10^{-12} . Under a medium traffic load of $\alpha = 0.3$, the packet loss reaches around 10^{-7} . When traffic load is heavy such as $\alpha = 0.7$, however, adding FDLs only reduce packet loss to about 10^{-3} . Similar observations actually can be made in Figure 6.12 as well.

To summarize, we have proposed and tested three different potential solutions for the common objective of approaching the packet-loss performance of the ideal switch. Among them, the most efficient one turns out to be increasing the buffer size in each board. In fact, it is the only one among the three solutions that has managed to outperform the ideal switch. Between the other two approaches, namely to increase the TSB and FSB sizes respectively, the former one appears to be more effective. The latter one helps improve packet-loss performance only under medium to heavy traffic load. Rather than approaching the performance of the ideal switch, in many cases we may want to achieve a good balance between cost and packet-loss performance. While a comprehensive study may generate another chapter, our preliminary study shows that the selection of the best strategy would probably be load dependent: a moderate investment can significantly improve network performance under low to medium traffic load but not under heavy traffic load.

6.5 Conclusion

In this chapter, we have studied the feasibility of supporting multi-fiber OPS networks by the single-stage multi-board switches. Several different schemes for interconnecting multiple switching boards and sharing FDLs have been proposed.

Chapter 6: Node Configurations in Multi-fiber OPS Networks

Extensive simulation results have identified the best scheme among them. To solve the problem that the multi-board switches cannot easily perform comparable to the ideal, strictly non-blocking switch, several different solutions have been proposed and evaluated. It turns out that the most effective one is to increase the buffer size, at the cost of having an increased number of switching boards, rather than to increase the switching-board size or to enhance the FDL sharing between different boards. Finally, we showed that a good balance between moderate investment and satisfactory network performance may be achievable under light to medium traffic loads, but probably not under heavy traffic load. These insights would be useful for the future developments of cost-effective optical packet switches.

Chapter 7

Traffic Load Allocations in Optical Packet Switches

Various OPS node architectures and configurations have been extensively studied in the previous chapters, in which we focused on the different contention resolutions. In this chapter, we further study on another important factor that impacts the packet loss performance in optical packet switches (OPS), the *traffic load allocations*. In particular, two major cases are investigated where traffic load is distributed between different (i) time slots (time dimension); and (ii) output ports (space dimension), respectively. These two cases are of significant importance as each has different applications and forms the basis of some more complicated schemes. Our main contributions are two fold. First, we find that either time or space dimension, the packet loss appears to be a convex function of traffic load. Consequently it can be easily proved that the uniform traffic distribution leads to the lowest packet loss. Second, for OPS with limited capability of adjusting a given traffic distribution, we propose a *Load Balancing* (LB) algorithm which minimizes the packet loss after the traffic adjustments.

7.1 Introduction

Contention resolution technologies play a key role in the packet loss performance of OPS. On the other hand, it is also well known that traffic distributions have nontrivial impacts on the packet loss. For example, bursty traffic generally leads to a higher packet loss [26], [131]. Therefore many schemes have been proposed to lower the traffic burstiness or nonuniformity (e.g., [131], [132]). In this chapter, we consider the effects of traffic allocations in a general framework, which includes two classes of different cases as follows:

1. *Time dimension*, where the *ratio* of traffic load destined for different output links (i.e., *traffic pattern*) cannot be significantly modified, yet changing the distribution of traffic load in different time slots is allowed. Existing results of smoothing the burstiness of data traffic at network boundary [131], [132] can be viewed as roughly belonging to this class: traffic is re-distributed in different time slots to achieve better fairness in time dimension, while the traffic pattern is largely decided by traffic load between different source-destination pairs and hence cannot be significantly changed.
2. *Space dimension*, where the traffic pattern can be adjusted yet the traffic distribution in time dimension is fixed. Examples include allocating outgoing packets on several different alternative routes as that in the deflection routing scheme [31]; or allocating traffic on different fibers in a same multi-fiber output link [130], etc.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

The above two classes of load allocations may happen simultaneously in most real-world operations, though generally one of them would be of a more significant effect than the other. In this chapter, to achieve some straight yet insightful observations, however, we restrict our discussions to the most representative cases where traffic allocations happen in only *one* dimension. Our main contributions can be summarized as follows:

- We find that in both the time and space dimensions, the packet loss appears to a convex function of traffic load; therefore the lowest packet loss is achieved when the traffic load is evenly distributed among different time slots or output ports.
- For an optical switch with limited capability of adjusting a given traffic distribution, we propose a simple *Load Balancing* (LB) algorithm to minimize the packet loss after the adjustments.

Through theoretical analysis as well as extensive numerical simulations, we show that this convexity condition is probably fulfilled in most, if not all, of the different OPS configurations. Therefore, our contributions are general enough to help achieve best load allocation in various OPS networks.

7.2 Analysis of the Most Fundamental OPS Configuration

Throughout this chapter, we assume that the arriving packets are synchronized and contained in the fixed-length time slots before entering the switching matrix [26], [38], [46]. We evaluate the *packet loss* of OPS, defined as the *statistic average number* of the

Chapter 7: Traffic Load Allocations in Optical Packet Switches

lost packets due to contentions in each time slot. The most fundamental switch configuration we will study in this section is illustrated in Figure 7.1. It is a general OPS model with extensive existing studies on its packet-loss rate and average packet-delay performance (e.g., [24], [30], [109], [131]). In this model, there are N input/output links with M channels per link connected to the same switch matrix. The channels are assumed to be of the same wavelength, or equivalently, be equipped with redundant wavelength conversion capability. Apparently, in such a switch, there could be a maximum of $N \times M$ packets arriving in each time slot; and at most M packets be transmitted by each output link at one time. Note that the switch has no contention resolution component other than the multiple (if $M > 1$) channels in each link. Switch configurations with additional components, such as shared wavelength converters or shared buffers, will be evaluated later in Section 7.4.

Figure 7.1: The most fundamental OPS configuration.

In our analysis, the *average* traffic load, defined as the probability of having a packet arriving in each time slot from each input channel, is denoted as α . To define the *traffic pattern*, we denote ρ_i ($i = 1, 2, \dots, N$) as the probability that a newly arrived

Chapter 7: Traffic Load Allocations in Optical Packet Switches

packet is destined for the i -th output link. For the special case of uniform traffic pattern [24], [26], [30], [109], we have that $\rho_i = 1/N$, $i = 1, 2, \dots, N$. Based on these definitions, the probability of having j packets destined for the output link i can be calculated as:

$$P_i(j) = \binom{NM}{j} \cdot (\rho_i \alpha)^j \cdot (1 - \rho_i \alpha)^{N \cdot M - j}, \quad j = 0, 1, \dots, NM. \quad (7.1)$$

When $j \leq M$, all the j packets can be transmitted out. When $j > M$, however, $j - M$ packets will be lost. Therefore, the *packet loss* among those packets destined for output link i can be calculated as:

$$\begin{aligned} PL_i &= \sum_{j=M+1}^{NM} P_i(j) \cdot (j - M) \\ &= \left[\sum_{j=1}^{NM} P_i(j) \cdot j - \sum_{j=1}^M P_i(j) \cdot j \right] - M \cdot \left[\sum_{j=0}^{NM} P_i(j) - \sum_{j=0}^M P_i(j) \right] \\ &= NM \rho_i \alpha - \sum_{j=1}^M P_i(j) \cdot j - M + M \cdot \sum_{j=0}^M P_i(j) \\ &= NM \rho_i \alpha - M + \sum_{j=0}^M P_i(j) \cdot (M - j) \end{aligned} \quad (7.2)$$

The overall packet loss of the switch therefore can be obtained as:

$$PL = \sum_{i=1}^N PL_i \quad (7.3)$$

As later we would see that the traffic model in Eqs. (7.1) - (7.3) is highly accurate.

Now we consider the adjustments of traffic load distribution in two dimensions respectively. We see that the traffic distribution in time dimension can be reflected by the parameter α (A larger value of α denotes a higher traffic load, and vice versa.),

Chapter 7: Traffic Load Allocations in Optical Packet Switches

while the traffic distribution in space dimension can be defined by the traffic pattern $\{\rho_i, i = 1, 2, \dots, N\}$. Therefore, the two representative cases of traffic allocations can be re-phrased as follows:

1. *Traffic allocations in time dimension*, where traffic pattern $\{\rho_i, i = 1, 2, \dots, N\}$ is given and fixed. We are allowed to adjust $\alpha_t, t = 1, 2, \dots, T$, without changing the overall traffic load in T time slots $\sum_{t=1}^T \alpha_t$, to minimize the packet loss as defined in Eq. (7.3).
2. *Traffic allocations in space dimension*, where α is given and fixed. We are allowed to adjust the traffic pattern $\{\rho_i, i = 1, 2, \dots, N\}$ to minimize the packet loss in Eq. (7.3).

To figure out the optimal traffic distribution in time dimension that leads to the minimum packet loss. We have

Theorem 7.1: With any given set of non-negative parameters $\{\rho_i, i = 1, 2, \dots, N\}$, the packet loss PL defined in Eq. (7.3) is a convex function of α .

Proof of Theorem 7.1: We prove the convexity of the packet loss function by showing that $\frac{d^2}{d\alpha^2}(PL) \geq 0$. More specifically, we prove that

$$\frac{d^2}{d\alpha^2}(PL_i) > 0, i = 1, 2, \dots, N. \quad (7.4)$$

We have:

Chapter 7: Traffic Load Allocations in Optical Packet Switches

$$\begin{aligned}
\frac{d^2}{d\alpha^2}(PL_i) &= \frac{d^2}{d\alpha^2} \left[NM\rho_i\alpha - M + \sum_{j=0}^M P_i(j) \cdot (M-j) \right] \\
&= \frac{d^2}{d\alpha^2} \left[\sum_{j=0}^M P_i(j) \cdot (M-j) \right] \\
&= \sum_{j=0}^M \left[(M-j) \cdot \frac{d^2}{d\alpha^2} P_i(j) \right] \\
&= \sum_{j=0}^M (M-j) \cdot \frac{d^2}{d\alpha^2} \left[C_{NM}^j \cdot (\rho_i\alpha)^j \cdot (1-\rho_i\alpha)^{NM-j} \right] \\
&= \sum_{j=0}^M (M-j) \cdot \left[C_{NM}^j \cdot j \cdot (j-1) \cdot \rho_i^j \alpha^{j-2} (1-\rho_i\alpha)^{NM-j} \right. \\
&\quad \left. - 2C_{NM}^j \cdot j \cdot (N \cdot M - j) \cdot \rho_i^{j+1} \alpha^{j-1} (1-\rho_i\alpha)^{NM-j-1} \right. \\
&\quad \left. + C_{NM}^j \cdot (NM-j) \cdot (NM-j-1) \rho_i^{j+2} \alpha^j (1-\rho_i\alpha)^{NM-j-2} \right]
\end{aligned} \tag{7.5}$$

By expanding the above equation with $j=0,1,\dots,M$, we have that the coefficient of each item $\rho_i^{j+2}\alpha^j(1-\rho_i\alpha)^{NM-j-2}$, $j=0,1,\dots,M$, can be calculated as:

$$\begin{aligned}
&(M-j) \cdot C_{NM}^j \cdot (NM-j) \cdot (NM-j-1) - 2[M-(j+1)] \cdot [C_{NM}^{j+1} \cdot (j+1) \cdot (NM-j-1)] \\
&\quad + [M-(j+2)] \cdot [C_{NM}^{j+2} \cdot (j+2) \cdot (j+1)] \\
&= (M-j) \cdot \frac{NM \cdot (NM-1) \dots (NM-j+1)}{j!} \cdot (NM-j) \cdot (NM-j-1) \\
&\quad - 2 \cdot (M-j-1) \cdot \frac{NM \cdot (NM-1) \dots (NM-j)}{(j+1)!} \cdot (j+1) \cdot (NM-j-1) \\
&\quad + (M-j-2) \cdot \frac{NM \cdot (NM-1) \dots (NM-j-1)}{(j+2)!} \cdot (j+2) \cdot (j+1) \\
&= [NM \cdot (NM-1) \cdot (NM-2) \dots (NM-j-1)] \\
&\quad \times \left[\frac{M-j}{j} - 2 \cdot \frac{(M-j-1) \cdot (j+1)}{(j+1)!} + \frac{(M-j-2) \cdot (j+1) \cdot (j+2)}{(j+2)!} \right] \\
&= 0
\end{aligned}$$

Therefore, Eq. (7.5) can be simplified to be:

$$\frac{d^2}{d\alpha^2}(PL_i) = C_{N \cdot M}^{M-1} \cdot (N \cdot M - M + 1) \cdot (N \cdot M - M) \rho_i^{M+1} \alpha^{M-1} (1-\rho_i\alpha)^{N \cdot M - M - 1} \tag{7.6}$$

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Since $N \geq 1$, $M \geq 1$, $0 \leq \rho_i \leq 1$, $0 \leq \alpha \leq 1$, and $(1 - \rho_i \alpha) \geq 0$, we have

$\frac{d^2}{d\alpha^2}(PL_i) \geq 0$. Therefore, PL_i is a convex function of α . Consequently, PL is a convex function of α . $\square$

Since PL is convex function of α , from the Jensen's Inequality [133], we have

$$PL\left(\frac{1}{T} \sum_{i=1}^T \alpha_i\right) \leq \frac{1}{T} \sum_{i=1}^T PL(\alpha_i), \quad (7.7)$$

where equality is achieved when $\alpha_1 = \alpha_2 = \dots = \alpha_T$. Therefore, we have

Corollary 7.1: Under any given traffic pattern, among all the time-dimension traffic distributions, the uniform traffic allocation which evenly distributes traffic load among the different time slots achieves the lowest packet loss.

The optimality of the uniform distribution in time dimension is therefore verified.

For traffic allocations in space dimension, similar to Theorem 1, we have:

Theorem 7.2: With any given value of α ($0 \leq \alpha \leq 1$), the packet loss PL as defined in Eq. (7.3) is a convex function of ρ_i , $i = 1, 2, \dots, N$.

Proof of Theorem 7.2: As α and ρ_i are at symmetric positions in the function PL_i , $i = 1, 2, \dots, N$, by applying the same technique as above, we can prove that PL is a convex function of ρ_i , $i = 1, 2, \dots, N$. $\square$

Once again, from the Jensen's Inequality, we have that

Chapter 7: Traffic Load Allocations in Optical Packet Switches

$$PL\left(\frac{1}{N} \sum_{i=1}^N \rho_i\right) \leq \frac{1}{N} \sum_{i=1}^N PL(\rho_i) \quad (7.8)$$

which consequently leads to the conclusion as follows:

Corollary 7.2: Under any given traffic load, among all the space-dimension traffic distributions, the uniform traffic pattern leads to the lowest packet loss.

Theoretical results therefore confirm the optimality of the uniform traffic distribution in both time and space dimensions. In Section 7.4, these conclusions will be further verified by numerical simulations.

Note that all the above conclusions are based on the most fundamental OPS configuration as illustrated in Figure 7.1. For the other node configurations, such as those with shared wavelength converters (SWCs) and/or FDLs, as long as their packet losses are still convex functions of traffic load, the optimality of the uniform traffic distribution remains valid. By looking at the packet-loss performance of different node configurations, we see that they mostly share the common feature that when traffic load goes up, their *packet losses* increase faster under heavier traffic load (In other words, the derivative of their packet losses are non-decreasing functions of traffic load). Specifically, under low traffic load, since most of the increased traffic can go through the switch, the packet loss increases slowly. When traffic load becomes higher, however, an increasingly larger portion of the increased traffic will be blocked and dropped. Finally, packet loss will increase at nearly the same speed as that of the traffic load, where virtually all the increased traffic is dropped. The non-decreasing first-order derivative of the packet-loss function indicates a convex packet-loss function in these node configurations.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Though the above observation seems to be obvious in most real-world systems, to strictly prove the convexity of the packet-loss function in various optical switches generally remains as a challenge: we need to have very accurate yet not too complicated analytical models for the switches under general non-uniform traffic load, which are so far basically not available [30], [41]. Therefore, we adopt extensive simulations to study the packet-loss performance of some typical node configurations. As later we will see in Section 7.4, simulation results encouragingly indicate that packet losses in these OPS configurations all appear to be convex functions of traffic load, or at least closely resemble convex functions.

7.3 The Load Balancing (LB) Algorithm

The data flows in the networks are usually dynamic and bursty, which leads to uneven traffic load distributions in time and space dimensions. As discussed before, uneven traffic distributions cause more packet contentions and higher packet losses. To balance the traffic load, several approaches can be implemented in the OPS networks.

- In time dimension, traffic burstiness can be reduced by proper buffering, which includes the buffering at network edge nodes and the buffering at switching nodes. By holding some packets during the “peak time” and releasing them later, buffering operation helps to make traffic load in different time slots more balanced.
- In space dimension, we may explore space dimension in multi-fiber OPS by distributing traffic load to different fibers of the same output link. We may also adopt deflection routing by sending some packets to less-loaded output links. In multi-board OPS, the unevenly traffic distribution can be adjusted to some extent

Chapter 7: Traffic Load Allocations in Optical Packet Switches

by shifting traffic load from heavily-loaded boards to lower-loaded ones through shared buffers. Last but not least, wavelength converters help to achieve better load balancing between different wavelengths.

In the last section, we have proved that the uniform distribution leads to the best performance of the fundamental OPS model in either time or space dimension. In most practical switches, however, various constraints may make the uniform distribution beyond reachable. For example, we may have limited buffer size for adjusting the traffic distribution in time dimension, or limited wavelength conversion for adjusting the traffic distribution in space dimension. In this section, we consider the problem of minimizing the packet loss in an OPS with *limited* capability of adjusting a given traffic distribution. We propose a simple *Load Balancing* (LB) and prove its optimality under the simple condition that the packet loss of the switch is a convex function of traffic load (which probably holds in most node configurations).

The main idea of the LB algorithm is quite simple: we shift the traffic component(s) which is farthest away from the *mean value* ($\bar{\alpha} = \frac{1}{T} \sum_{t=1}^T \alpha_t$ in time dimension and $\bar{\rho} = \frac{1}{N} \sum_{i=1}^N \rho_i$ in space dimension respectively), or in other words the most unbalanced traffic component(s), to be closer to the mean value. Specifically, we shift traffic load from the most-loaded component(s) to the least-loaded component(s), until the traffic distribution is adjusted to be uniform, or the capability of adjusting traffic distribution has been exhausted. As an example, hereafter we discuss the traffic allocations in time dimension in detail.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Without loss of generality, we assume that the traffic distribution α_t , $t = 1, 2, \dots, T$, has $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_T$. The given capability of adjusting the traffic distribution is denoted as $\beta_{\max}$. In other words, $\sum_{t=1}^T |\alpha_t - \alpha'_t| \leq \beta_{\max}$, where $\{\alpha'_t\}$ denotes the traffic distribution after the adjustments.

LB Algorithm in Time Dimension

1. Calculate the *distance* of the given traffic distribution from the uniform distribution

$$\bar{\beta} = \frac{1}{2} \sum_{t=1}^T |\alpha_t - \bar{\alpha}| \quad (7.9)$$

where $\bar{\alpha} = \frac{1}{T} \sum_{t=1}^T \alpha_t$. If $\beta_{\max} \geq \bar{\beta}$, then the given traffic pattern can be adjusted to the uniform traffic distribution, stop; otherwise, let $\beta = \beta_{\max}$ and go to Step 2.

2. *Deduction operation*: Find all the traffic distribution components with the largest value $\alpha_m = \dots = \alpha_T$. Note that m can be equal to T . Let

$$\Delta\beta = \min\left(\alpha_T - \alpha_{m-1}, \frac{\beta}{T - m + 1}\right) \quad (7.10)$$

$$\alpha_j = \alpha_j - \Delta\beta, \quad j = m, m+1, \dots, T \quad (7.11)$$

$$\beta = \beta - \Delta\beta \times (T - m + 1) \quad (7.12)$$

Repeat the above procedure until $\beta = 0$.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

3. *Addition operation:* Re-initialize $\beta = \beta_{\max}$. Find all the components with the smallest value $\alpha_1 = \dots = \alpha_s$. Note that s can be equal to 1. Let

$$\Delta\beta = \min\left(\alpha_{s+1} - \alpha_1, \frac{\beta}{s}\right) \quad (7.13)$$

$$\alpha_j = \alpha_j + \Delta\beta, \quad j = 1, \dots, s \quad (7.14)$$

$$\beta = \beta - \Delta\beta \times s \quad (7.15)$$

Repeat the calculations in Eqs. (7.13) - (7.15) until $\beta = 0$. $\square$

The LB Algorithm in space dimension would be nearly the same, except that it is ρ_i , $i = 1, 2, \dots, N$, instead of α_t , $t = 1, 2, \dots, T$, that needs to be adjusted. The very similar descriptions of the detailed algorithm therefore are omitted.

Next we prove the LB algorithm's optimality in lowering packet loss with any given traffic distribution and $\beta_{\max}$. Once again we use traffic allocation in time domain as an example, while the optimality of the algorithm in space dimension can be proven similarly. As a starting point, we need to prove that the LB algorithm does help lower packet loss at all. We have

Lemma 7.1: Assume $f(x)$ is a convex function where $f'(x) \geq 0$ and $f''(x) \geq 0$.

For any $x_1 \leq x_2$ and $0 \leq \Delta x \leq x_2 - x_1$, we have

$$f(x_2) - f(x_2 - \Delta x) \geq f(x_1 + \Delta x) - f(x_1) \quad (7.16)$$

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Proof of Lemma 7.1: Since $f'(x) \geq 0$ and $f''(x) \geq 0$, we have that $f'(x)$ is a non-negative, non-decreasing function. Therefore, for any $x_1 \leq x_2$,

$$f'(x_1) \leq f'(x_2) \quad (7.17)$$

For any $x_1 \leq x_2$ and $0 \leq \Delta x \leq x_2 - x_1$, there are possibly two different cases: **Case (a)** $x_2 - \Delta x \geq x_1 + \Delta x$; and **Case (b)** $x_1 \leq x_2 - \Delta x < x_1 + \Delta x$. Hereafter we discuss the two cases separately.

Case (a) where $x_2 - \Delta x \geq x_1 + \Delta x$: for such case, from the well-known mean value theorem, we have that there exists an x_2^* , $x_2 - \Delta x \leq x_2^* \leq x_2$, such that

$$f(x_2) - f(x_2 - \Delta x) = f'(x_2^*) \cdot \Delta x \quad (7.18)$$

Meanwhile, there exists an x_1^* , $x_1 \leq x_1^* \leq x_1 + \Delta x$, such that

$$f(x_1 + \Delta x) - f(x_1) = f'(x_1^*) \cdot \Delta x \quad (7.19)$$

From Eq. (7.17), since $x_2^* \geq x_1^*$, we have that

$$f'(x_2^*) \geq f'(x_1^*) \quad (7.20)$$

Therefore, from Eqs. (7.18)-(7.20), we obtain that

$$f(x_2) - f(x_2 - \Delta x) \geq f(x_1 + \Delta x) - f(x_1) \quad (7.21)$$

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Case (b) where $x_1 \leq x_2 - \Delta x < x_1 + \Delta x$: Similar to the above, there exists a y_1 , $x_1 \leq y_1 \leq x_2 - \Delta x$, such that

$$f(x_2 - \Delta x) - f(x_1) = f'(y_1) \cdot (x_2 - x_1 - \Delta x) \quad (7.22)$$

Meanwhile, there exists a y_2 , $x_1 + \Delta x \leq y_2 \leq x_2$, such that

$$f(x_2) - f(x_1 + \Delta x) = f'(y_2) \cdot (x_2 - x_1 - \Delta x) \quad (7.23)$$

Since $y_1 \leq y_2$, similar to that in Eq. (7.20), we have

$$f'(y_1) \leq f'(y_2) \quad (7.24)$$

which in turn, similar to that in Eq. (7.21), leads to the conclusion that

$$f(x_2 - \Delta x) - f(x_1) \leq f(x_2) - f(x_1 + \Delta x) \quad (7.25)$$

or equivalently,

$$f(x_2) - f(x_2 - \Delta x) \geq f(x_1 + \Delta x) - f(x_1) \quad (7.26)$$

Therefore, the Lemma 7.1 is proven for both case (a) and case (b). $\square$

By writing the PL_i as defined in Eq. (7.2) as $PL_i(\alpha)$, a function of traffic load α , we have that $PL_i'(\alpha) \geq 0$ (i.e., $PL_i(\alpha)$ is a strictly non-decreasing function of traffic load). If $PL_i''(\alpha) \geq 0$, or equivalently, $PL_i(\alpha)$ is a convex function of the traffic load, from Lemma 7.1, we have that

Chapter 7: Traffic Load Allocations in Optical Packet Switches

$$PL_i(\alpha_1 + \Delta\beta) - PL_i(\alpha_1) \leq PL_i(\alpha_2) - PL_i(\alpha_2 - \Delta\beta) \quad (7.27)$$

or equivalently,

$$PL_i(\alpha_1 + \Delta\beta) + PL_i(\alpha_2 - \Delta\beta) \leq PL_i(\alpha_1) + PL_i(\alpha_2) \quad (7.28)$$

for any $\alpha_1 \leq \alpha_2$ and $0 \leq \Delta\beta \leq \alpha_2 - \alpha_1$. Therefore,

Theorem 7.3: If the packet loss in a switch is a non-decreasing convex function of traffic load, the packet loss can be lowered by shifting traffic load from heavily-loaded components to less-loaded components (without making the heavily-loaded components to be of lower values than those of the less-loaded components before the adjustments). $\square$

Next we proceed to prove the optimality of the LB algorithm under the same convexity condition. We have

Lemma 7.2: Assume $f(x)$ is a convex function where $f'(x) \geq 0$ and $f''(x) \geq 0$.

For any $x_1 \leq x_2$ and $0 \leq \Delta x \leq x_2 - x_1$, we have

$$\begin{cases} f(x_2) - f(x_2 - \Delta x) \geq f(x_1) - f(x_1 - \Delta x), & (a) \\ f(x_2 + \Delta x) - f(x_2) \geq f(x_1 + \Delta x) - f(x_1). & (b) \end{cases} \quad (7.29)$$

Proof of Lemma 7.2: The proof of Lemma 7.2 is straightforward. Specifically, by replacing the x_1 in Lemma 7.1 with $x_1 - \Delta x$, Eq. (7.29-a) could be proven; by replacing the x_2 in Lemma 7.1 with $x_2 + \Delta x$, Eq. (7.29-b) could be proven. $\square$

From Eq. (7.29-a), if $PL_i(\alpha)$ is a convex function of traffic load α , we have that

Chapter 7: Traffic Load Allocations in Optical Packet Switches

$$PL_i(\alpha_1) - PL_i(\alpha_1 - \Delta\beta) \leq PL_i(\alpha_2) - PL_i(\alpha_2 - \Delta\beta) \quad (7.30)$$

for any $\alpha_1 \leq \alpha_2$ and $0 \leq \Delta\beta \leq \alpha_2 - \alpha_1$. It indicates that in the Step 2 of the LB algorithm, reducing the most-loaded traffic component leads to the *quickest* decrease of packet loss. Meanwhile, from Eq. (7.29-b), we have that

$$PL_i(\alpha_1 + \Delta\beta) - PL_i(\alpha_1) \leq PL_i(\alpha_2 + \Delta\beta) - PL_i(\alpha_2) \quad (7.31)$$

for any $\alpha_1 \leq \alpha_2$ and $0 \leq \Delta\beta \leq \alpha_2 - \alpha_1$. It indicates that in the Step 3 of the LB algorithm, adding traffic to the least-loaded traffic component leads to the *slowest* increase of packet loss. Combining the observations in Eqs. (7.28), (7.30) and (7.31), we can draw the conclusion as follows:

Theorem 7.4: For any given traffic distribution and the given capability of adjustment $\beta_{\max}$, the LB algorithm achieves the lowest packet loss after the adjustments as long as the packet loss of the switch is a non-decreasing, convex function of traffic load. $\square$

To put the LB algorithm into practical applications, we may need to measure several parameters include the load distribution $\{\alpha_i\}$ (or $\{\rho_i\}$) and the adjustment capacity $\beta_{\max}$. In many cases, $\{\alpha_i\}$ can be calculated from historic records. When the maximum number of packets that can be shifted is known (For example, the number of packets that can be buffered equals to the buffer size.), it can be converted to $\beta_{\max}$. For the other cases where historic record-based calculations are not feasible or not reliable (e.g., the case where the traffic pattern changes quickly), the LB algorithm still

Chapter 7: Traffic Load Allocations in Optical Packet Switches

provides us the guideline that traffic load shifting should always be from the most-loaded components to the least-loaded components until the traffic distribution is uniform or until the adjustment capability has been saturated.

Note that the traffic adjustment we have discussed so far are all for operations in one and only one of the time and space dimensions. For the more complicated cases where traffic distributions in the two dimensions are strongly correlated to each other (i.e., an adjustment operation almost always changes traffic distributions in both dimensions simultaneously), further studies are expected.

7.4 Simulation Results and Discussions

7.4.1 Packet Loss in the Most Fundamental OPS Configuration

We start by considering the packet-loss performance in the fundamental OPS configuration as illustrated in Figure 7.1. To evaluate the effects of traffic allocations in time dimension, we assume that the overall traffic load in T time slots, measured by

$\sum_{t=1}^T \alpha_t$, is given and fixed. Specifically, we let $T = 4$ and $\alpha = \frac{1}{T} \sum_{t=1}^T \alpha_t = 0.5$, and

consider several different cases as follows:

- (a) $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{0.5, 0.5, 0.5, 0.5\}$
 - (b) $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{0.6, 0.6, 0.4, 0.4\}$
 - (c) $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{0.7, 0.7, 0.3, 0.3\}$
 - (d) $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{0.8, 0.8, 0.2, 0.2\}$
 - (e) $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{0.9, 0.9, 0.1, 0.1\}$
- (7.32)

For simplicity, we differentiate them by introducing the *deviation* degree γ :

Chapter 7: Traffic Load Allocations in Optical Packet Switches

$$\gamma = \frac{\sum_{t=1}^T |\alpha_t - \alpha|}{T} \quad (7.33)$$

The above five different cases then can be denoted as with different values of γ varying from 0 (case (a)) to 0.4 (case (e)).

As to the traffic distribution in space dimension, we consider several cases as follows:

- (i) *Uniform traffic pattern*, where $\rho_i = 1/N$, $i = 1, 2, \dots, N$;
- (ii) *Geometric traffic pattern*, where $\rho_i / \rho_{i-1} = R$, $i = 2, 3, \dots, N$. In other words, the average traffic load destined for different output links distribute as a geometric sequence. We present two different cases where $R = 1.1$ and $R = 1.2$ respectively; and
- (iii) *Hotspot traffic*, where we assume that each of the three hotspot output links has traffic destined for it five times as high as that destined for each of the other output links.

Both the analytical and the simulation results are presented in Figure 7.2, where we let $N = 8$, $M = 4$, $\alpha = 0.5$, and $T = 4$ (Note that such parameter values will be adopted in all the following simulations unless otherwise specified.). We see that for any value of γ , uniform traffic pattern steadily leads to the best performance in space-dimension traffic allocations. Meanwhile, for all the traffic patterns, $\gamma = 0$ leads to the best performance in time-dimension traffic allocations. Therefore, the optimality of the uniform traffic distribution in both dimensions is verified; and the global optimal

Chapter 7: Traffic Load Allocations in Optical Packet Switches

performance is achieved where traffic distribution is uniform in both time and space dimensions. Meanwhile, we see that the proposed analytical models appear to be highly accurate, generating analytical results almost perfectly matching the simulation results.

Figure 7.2: Performance of the most fundamental OPS under various load distributions in time and space dimensions.

Note that it can be proved that larger deviations lead to higher packet losses (since the load distributions with smaller deviations can be achieved by applying the LB algorithm to those load distributions with larger deviations). However, it basically remains as an open problem to quickly tell between two given traffic distributions which one leads to lower packet loss, unless one of them could be achieved by adjusting the other one with the LB algorithm.

7.4.2 Packet Loss Evaluation of Other OPS Configurations

Figure 7.3: Typical OPS configuration with SWCs and FDLs.

Optical packet switches are typically implemented with contention resolution components such as FDLs or SWCs or both, where a typical example is shown in Figure 7.3. Since it is notoriously difficult to achieve accurate yet simple analytical models for such types of node configurations under general non-uniform traffic load [30], [39], [67], [109], as we have discussed before, we investigate their packet-loss performance through numerical simulations.

Figure 7.4 shows the results where the switch is equipped with shared FDLs but no SWCs. We denote the number of FDLs as N_{FDL} and assume that they can be accessed by any input port. When there are more contending packets than that can be handled by FDLs, the contending packets will be randomly selected to be buffered. Figures 7.4(a)

Chapter 7: Traffic Load Allocations in Optical Packet Switches

and 7.4(b) show the results where $M = 1$ (may be viewed as single-fiber case) and $M = 4$ (may be viewed as multi-fiber case), respectively. For both cases, we see that the conclusion remains valid that the uniform traffic distribution leads to the best performance in both time and space dimensions. In other words, the existence of FDLs helps to lower packet loss, yet does not affect the optimality of the uniform traffic distribution. In fact, we observe that all the curves appear to be convex functions of traffic load.

Node configuration with SWCs but no FDLs is evaluated in Figure 7.5. We assume that there are totally W different wavelengths in each link, and a certain number (denoted as N_{swc}) of SWCs are installed in the switch, which can be accessed by any input port. We consider two different cases where M channels in each output link belong to a single-fiber ($W = 4$) or multiple fibers ($W = 2$), respectively. In both cases, we set $N_{swc} = 2$. The results in Figure 7.2 meanwhile can be viewed as for the special case with unlimited wavelength-conversion capability. When $W = 4$, all contentions have to be resolved by SWCs. When $W = 2$, the two channels on each wavelength help to resolve some contentions. From Figure 7.5, we see that for both cases, the uniform traffic distribution still performs the best, in both time and space dimensions.

(a) $M = 1$ (can be viewed as single-fiber case).

(b) $M = 4$ (can be viewed as multi-fiber case).

Figure 7.4: Packet-loss performance in the OPS with FDLs but no SWC.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

(a) $M = 4, W = 4$ (single-fiber case).

(b) $M = 4, W = 2$ (multi-fiber case).

Figure 7.5: Packet-loss performance in the OPS with SWCs but no FDL.

(a) $M = 4, W = 4$ (single-fiber case).

(b) $M = 4, W = 2$ (multi-fiber case).

Figure 7.6: Packet-loss performance in the OPS with both SWCs and FDLs.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

Finally, we evaluate the node configuration with both the SWCs and the FDLs, where the simulation results are presented in Figure 7.6. Again, we test the single-fiber ($W = 4$) and multi-fiber ($W = 2$) cases respectively. We assume that a packet contention will be resolved by wavelength conversion as long as such is applicable before an available FDL is used. We see that although having more contention resolution components helps improve the packet-loss performance under non-uniform traffic load, the best performance is nevertheless still achieved under uniform traffic distribution; and the packet losses of all the different cases still appear to be convex functions of traffic load.

7.4.3 Load Balancing Algorithm in the Most Fundamental Node Configuration

Figure 7.7 presents the simulation results where the LB algorithm is applied in the most fundamental node configuration as illustrated in Figure 7.1, where the multiple channels ($M = 4$) in each link are either of the same wavelength or equipped with full-range wavelength converters. The packet loss performance versus the adjustment capability $\beta_{\max}$ in time domain is shown in Figure 7.7(a), where we assume that the traffic pattern is uniform. Two different cases, namely Case 1 where $\alpha_i = \{0.2, 0.4, 0.6, 0.8\}$; and Case 2 where $\alpha_i = \{0.1, 0.3, 0.7, 0.9\}$ are considered respectively. The packet losses are steadily lowered with increased values of $\beta_{\max}$, until finally the uniform time-domain distribution is achieved (which corresponds to the cases where $\beta_{\max} = 0.4$ in Case 1 and $\beta_{\max} = 0.6$ in Case 2 respectively).

For traffic allocations in space dimension, it is presented in Figure 7.7(b) the effectiveness of the LB algorithm. Again we consider two different cases: geometric-

Chapter 7: Traffic Load Allocations in Optical Packet Switches

distributed case where $\rho_i / \rho_{i-1} = R = 1.3$, $i = 2, 3, \dots, N$, and the Hotspot case. For both cases, we let $\alpha = 0.5$. Both the analytical and the simulation results show that, the lowest packet losses are achieved when $\beta_{\max} = \frac{1}{2} \sum_{i=1}^N \left| \rho_i - \frac{1}{N} \right|$, that is, $\beta_{\max} \approx 0.246$ for the geometric-distributed case, and $\beta_{\max} = 0.375$ for the hotspot case respectively, where the traffic patterns are adjusted to be uniform.

Our theoretical results confirm that in space and time dimensions, the LB algorithm achieves the best performance subject to a given capacity of traffic load adjustment.

7.4.4 Load Balancing Algorithm in Other OPS Configurations

For OPS implemented with FDLs and/or SWCs as illustrated in Figure 7.3, their packet-loss performance versus the adjustment capacity $\beta_{\max}$, in both time and space dimensions, are presented in Figure 7.8 through Figure 7.10. For load allocations in time dimension, we still consider Case 1 where $\alpha_t = \{0.2, 0.4, 0.6, 0.8\}$; and Case 2 where $\alpha_t = \{0.1, 0.3, 0.7, 0.9\}$ respectively. For load allocations in space dimension, we still consider geometric-distribution and hotspot traffic, while assigning $\alpha = 0.5$. Figure 7.8 shows the performance of OPS with only FDLs, where $N_{FDL} = 2$ and $W = 1$. Compared with Figure 7.7, with FDLs the packet losses in Figure 7.8 are largely reduced. However, that does not affect the effectiveness of the LB algorithm, which manages to achieve the best performance at exactly the same values of $\beta_{\max}$ as those in Figure 7.7. Very similar observations could be made in Figure 7.9, where we have SWCs ($W = 4$ and $N_{SWC} = 2$) but no FDL, as well as in Figure 7.10, where we have both SWCs ($W = 4$ and $N_{SWC} = 2$) and FDLs ($N_{FDL} = 2$).

Chapter 7: Traffic Load Allocations in Optical Packet Switches

(a) Adjustments in time dimension (uniform traffic pattern).

(b) Adjustments in space dimension.

Figure 7.7: Performance of the LB algorithm in the most fundamental OPS configuration.

Chapter 7: Traffic Load Allocations in Optical Packet Switches

(a) Adjustments in time dimension (uniform traffic pattern).

(b) Adjustments in space dimension.

Figure 7.8: Performance of the LB algorithm in OPS with only FDLs ($N_{FDL} = 2$).

Chapter 7: Traffic Load Allocations in Optical Packet Switches

(a) Adjustments in time dimension (uniform traffic pattern).

(b) Adjustments in space dimension.

Figure 7.9: Performance of the LB algorithm in OPS with only SWCs. ($N_{SWC} = 2$).

Chapter 7: Traffic Load Allocations in Optical Packet Switches

(a) Adjustments in time dimension (uniform traffic pattern).

(b) Adjustments in space dimension.

Figure 7.10: Performance of the LB algorithm in OPS with both FDLs and SWCs ($N_{FDL} = 2, N_{SWC} = 2$).

7.5 Conclusions

In this chapter, we evaluated the impacts of traffic allocations on the packet-loss performance of various OPS configurations. For the most fundamental OPS configuration, we proved that the packet loss in it is a convex function of traffic load. Hence the uniform traffic distribution has the best performance in both time and space dimensions. Our extensive simulations also show that such conclusions probably apply in most of the other node configurations as well. For any given traffic distribution and any given capacity of traffic adjustment, we proposed the simple Load Balancing algorithm which yields the optimal performance in any switch as long as its packet loss is a convex function of traffic load. These results are not restricted to any particular node configurations or traffic distributions; therefore they are general enough to help achieve best performance in different OPS networks.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Optical packet switching has been viewed as a promising solution for future all-optical networks. It is expected that future OPS will provide ultra high switching speed, transparency of data rate and format, fine traffic granularity and upgrading flexibility. This dissertation reviewed the developments and the state-of-art technologies in optical packet switches and studied various OPS models, with the main focus on *contention resolution* technologies. The major contributions of this study can be concluded in the following aspects.

Firstly, to best utilize the available contention resolution resources, we evaluated the resource-sharing model with fiber delay lines and wavelength converters. Based on a generic simulation model and extensive simulation results, some guidelines for achieving best cost-effectiveness were therefore proposed. We demonstrated that utilizing TWCs at a higher priority than FDLs leads to a slightly better performance.

Chapter 8: Conclusions and Future Work

And we recommended the integrated schemes utilizing both TWCs and FDLs as attractive candidates among various contention resolution solutions.

In the second part of our work, we proposed a new scheme using the fixed wavelength converters instead of the expensive (and immature) tunable wavelength converters. It is shown that under either bursty or non-bursty traffic load, with or without FDL buffers, the fixed wavelength converters can achieve nearly the same performance as that of the full-range tunable wavelength converters. Therefore, a better cost-effectiveness could be achieved. Furthermore, we showed that the FWC scheme outperforms the scheme utilizing limited-range wavelength converters as well.

In the third part, we conducted the investigation on the multi-fiber OPS (MOPS) model, where the space dimension provided by the multiple fibers per link is explored to resolve the packet contentions. Analytical models of the MOPS configurations with/without FDLs and/or TWCs were presented, which accurately describe the respective packet loss performance. Numerical results evidently demonstrated the significance of exploring the space dimension and also showed that, with a limited number of FDLs and SWCs, MOPS even outperforms OPS with FTWC. Some guidelines have been proposed for achieving better cost-effectiveness where each switching board has to be of a limited and fixed size.

We then studied the single-stage multi-board switches in details. Different node configurations with different inter-board connections and buffer sharing methods were simulated and compared with each other. Several solutions for improving packet-loss performance have been proposed and carefully examined. We found that among them the most effective one is to increase buffer size rather than to increase the switching-

Chapter 8: Conclusions and Future Work

board size or to enhance buffer sharing. In further studies, it was shown that a good balance between moderate investment and satisfactory network performance may be achievable under light to medium traffic loads, but probably not under heavy traffic load.

Finally, we studied the impacts of traffic allocations to the packet-loss performance of OPS. By proving the convexity of the packet-loss function of the most fundamental node configuration, as well as conducting extensive simulations for the more general cases, we showed that it helps to lower packet loss by evenly allocating traffic load as far as possible. For any given traffic distribution and capability of traffic adjustment, we proposed the simple Load Balancing algorithm which achieves the optimal performance providing that the packet loss is a convex function of traffic load.

8.2 Future Work

Based on our experiences, we recommend several topics will be of value for the future OPS research.

For the resource-sharing OPS model where both FDLs and FTWCs are implemented, there is always a tradeoff between the cost and the packet-loss performance. We expect to find a cost model that can quantify the relation of performance and cost, based on which the optimal cost-effectiveness may be achieved.

When we perform analytical evaluations on various OPS models, usually we assume the traffic pattern to be uniform. However, the practical traffic patterns almost always show non-uniform characteristics. Analysis of packet-loss performance under

Chapter 8: Conclusions and Future Work

general non-uniform traffic load basically remains as an open problem which we would tackle in our future research.

Most of the existing results aim to achieving the lowest packet loss in a single node. However, we can reasonably expect that the solution optimizing the performance on a single node (achieving local optimum) may not guarantee the best performance of the whole network (achieving global optimum). Unfortunately the network-wide behaviors of OPS have never been thoroughly studied. In a network-wide prospect, a lot of problems still remain open, ranging from the fundamental ones such as the correlations between the local performance of each node and the global performance of the network, to the highly difficult ones like the optimal allocations of optical buffers and wavelength converters in a given network, etc. The network-level research may revolutionize our understandings of the “all optical networks of the future”.

To summarize, there are three topics we are most interested to study in our future work: a cost model of OPS node and the cost model-based optimization; general non-uniform traffic-based packet loss analysis; and network-wide packet-loss performance evaluation and optimization. We believe they are safely among the most important open problems on networking level in OPS networks.

Author's Publications

1. Yi Li and H. Ghafouri-Shiraz, "Resource-sharing for contention resolution in optical packet switching," *SPIE Asia-Pacific Optical and Wireless Communications, APOC 2002*, vol. 4907, pp. 268-273.
2. Yi Li and H. Ghafouri-Shiraz, "An efficient method in optical packet switching for contention resolution," 1st International Conference on Optical Communications and Networks, *ICOON 2002*, pp. 142-144.
3. Yi Li and H. Ghafouri-Shiraz, "Contention resolution by shared wavelength converters and fiber delay lines in an optical packet switch," *Microwave and Optical Technology Letters*, vol. 38, no. 5, pp. 395-398, Sept. 2003.
4. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "Fixed wavelength conversion for contention resolution in optical packet switches," *Microwave and Optical Technology Letters*, vol. 41, no. 3, pp. 185-187, May 2004.
5. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "On the benefits of multifiber optical packet switch," *Microwave and Optical Technology Letters*, vol. 43, no. 5, pp. 376-378, Dec 2004.
6. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "A new contention-resolution scheme for time-critical applications in multifiber optical packet-switched networks," *Microwave and Optical Technology Letters*, vol. 48, no. 4, pp. 717-719, Apr. 2006.

Author's Publications

7. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "Performance evaluation of multi-fiber optical packet switches," *Computer Networks*, vol. 51, issue 4, pp. 995-1012, 14 Mar. 2007.
8. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "On the performance of different node configurations in multi-fiber optical packet-switched networks," *Photonic Network Communications*, published online, 5 October 2006.
9. Yi Li, Gaoxi Xiao, and H. Ghafouri-Shiraz, "On traffic allocations in optical packet switches," revised and re-submitted to *IEEE Journal on Selected Areas in Communications*.

Bibliography

- [1] G. Keiser, *Optical Fiber Communications*. McGraw-Hill, 2000.
- [2] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*. Morgan Kaufmann Publishers, 1998.
- [3] R. Ramaswami, "Optical fiber communication: from transmission to networking," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 138-147, May. 2002.
- [4] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki, "The application of optical packet switching in future communication networks," *IEEE Commun. Mag.*, vol. 39, no.3, pp. 128-135, Mar. 2001.
- [5] V. W. S. Chan, K. L. Hall, E. Modiano, and K. A. Rauschenbach, "Architecture and technologies for high-speed optical data networks," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 12, pp. 2146-2168, Dec. 1998.
- [6] S. P. Monacos, J. M. Morookian, L. Davis, L. Bergman, S. Forouhar, and J. R. Sauer, "All-optical WDM packet networks," *IEEE/OSA J. Lightwave Technol.*, vol. 14, no. 6, pp. 1356-1370, Jun. 1996.
- [7] T. Battestilli and H. Perros, "An introduction to optical burst switching," *IEEE Commun. Mag.*, vol. 41, no. 8, pp. S10-S15, Aug. 2003.
- [8] C. Develder, J. Cheyns, E. V. Breusegem, E. Baert, D. Colle, M. Pickavet, P. Demeester, "Architectures for optical packet and burst switches," in *Proc. ECOC'03*, Mo4.4.1.

Bibliography

- [9] L. Xu, H. G. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," *IEEE Commun. Mag.*, vol. 39, no. 1, pp.136-142, Jan. 2001.
- [10] C. Qiao, "Labeled Optical burst switching for IP-over-WDM integration," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 104-114, Sept. 2000.
- [11] C. Qiao, Y. Chen, J. Staley, "The potentials of optical burst switching (OBS)," in *Proc. OFC'03*, vol. 1, pp. 219-220.
- [12] A. Ge, F. Callegati, and L. S. Tamil, "On optical burst switching and self-similar traffic," *IEEE Commun. Lett.*, vol. 4, no. 3, pp. 98-100, Mar. 2000.
- [13] H. Buchta, E. Patzak, J. Saniter, and C. Gauger, "Limits of effective throughput of optical burst switches based on semiconductor optical amplifiers," in *Proc. OFC'03*, vol. 1, pp. 215-217.
- [14] Y. Horiuchi and M. Suzuki, "Demonstration of dynamic self-routing in an all-optical burst switched mesh network with four optical label switch routers," in *Proc. OFC'03*, vol. 1, pp. 217-219.
- [15] D. K. Hunter and I. Andonovic, "Approaches to optical Internet packet switching," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 116-122, Sept. 2000.
- [16] M. Renaud, F. Masetti, C. Guillemot, and B. Bostica, "Network and system concept for optical packet switching," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 96-102, Apr. 1997.
- [17] S. Yao, B. Mukherjee, and S. Dixit "Advances in photonic packet switching: An overview," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 84-94, Feb. 2000.

Bibliography

- [18] S. Yao, S. J. B. Yoo, B. Mukherjee, and S. Dixit, "All-optical packet switching for metropolitan area networks: opportunities and challenges," *IEEE Commun. Mag.*, vol. 39, no. 3, pp. 142-148, Mar. 2001.
- [19] T. S. El-Bawab, A. Agrawal, F. Poppe, L. B. Sofman, D. Papadimitriou, B. Rousseau, "The evolution to optical-switching-based core networks," *Optical Networks Mag.*, pp. 7-19, Mar/Apr. 2003.
- [20] T. S. El-Bawab and J. Shin, "Optical packet switching core networks: Between vision and Reality," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 60-65, Sep. 2002.
- [21] G. I. Papadimitriou, C. Papazoglou, A. S. Pomportsis, "Optical switching: Switch fabrics, techniques, and architectures," *IEEE/OSA J. Lightwave Technol.*, vol. 21, no. 2, pp. 384-405, Feb. 2003.
- [22] I. Chlamtac *et al.*, "CORD: Contention resolution by delay lines," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 1014-1029, Jun. 1996.
- [23] Z. Hass, "The staggering switch: An electronically controlled optical packet switch," *IEEE/OSA J. Lightwave Technol.*, vol. 11, no. 5/6, pp. 925-936, May/June 1993.
- [24] 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.
- [25] W. D. Zhong and R. S. Tucker, "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching system," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 10, pp. 1737-1745, Oct. 1998.
- [26] S. L. Danielsen, B. Mikkelsen, C. Joergensen, T. Durhuus, and K. E. Stubkjaer, "WDM packet switch architecture and analysis of influence on tuneable

Bibliography

- wavelength converters on the performance,” *IEEE J. Lightwave Technol.*, vol. 15, no. 2, Feb. 1997.
- [27] S. L. Danielsen, C. Joergensen, B. Mikkelsen, and K. E. Stubkjaer, “Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tuneable wavelength converters,” *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 5, pp. 729-735, May 1998.
- [28] E. Ciaramella, G. Contestabile, F. Curti, and A. D’Ottavi, “Fast tunable wavelength conversion for all-optical packet switching,” *IEEE Photon. Technol. Lett.*, vol. 12, no. 10, pp. 1361-1363, Oct. 2000.
- [29] A. Narula-Tam, M. H. Brady, and S. G. Finn, “Blocking probability of a centralized switch with shared wavelength conversion,” in *Proc. OFC’04*, TuH5.
- [30] V. Eramo and M. Listanti, “Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters,” *IEEE/OSA J. Lightwave Technol.*, vol. 18, no. 12, pp. 1818-1833, Dec. 2000.
- [31] H. Morino, T. T. Bao, H. Aida, and T. Saito, “Performance evaluation of variable length packet switch based on deflection routing and input port distribution,” in *Proc. IEEE ICC’00*, vol. 2, pp. 1050-1053.
- [32] T. Chich, F. Clerot, “Wavelength converters for deflection routing networks,” in *Proc. ISCC 2000*, pp. 140-145.
- [33] D. Chiaroni *et al.*, “First demonstration of an asynchronous optical packet switching matrix prototype for multiterabit-class routers/switches,” in *Proc. ECOC ’01*, vol. 6, pp. 60-61.

Bibliography

- [34] K. Kitayama and M. Murata, "Almost packet loss-free asynchronous, variable-length optical packet switch with WDM buffering," in *Proc. ECOC '01*, vol. 3, pp. 476-477.
- [35] L. Tancevski *et al.*, "Optical routing of asynchronous, variable length packets," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 2084-2093, Oct. 2000.
- [36] F. Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, no. 9, pp.292-294, Sept. 2000.
- [37] S. L. Danielsen, P. B. Hansen, and K. E. Stubkjaer, "Wavelength conversion in optical packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 9, pp. 2095-2108, Sept.1998.
- [38] S. Bregni, A. Pattavina and G. Vegetti, "Architecture and performance of AWG-based optical switching nodes for IP networks," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1113-1121, Sept. 2003.
- [39] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, I. Andonovic, "SLOB: A switch with large optical buffers for packet switching," *IEEE/OSA J. Lightwave Technol.*, vol.16, no. 10, pp. 1725-1736, Oct. 1998.
- [40] D. K. Hunter *et al.*, "WASPNET: A wavelength switched packet network," *IEEE Commun. Mag.*, vol. 37, no. 3, pp. 120-129, Mar. 1999.
- [41] M. C. Chia *et al.*, "Packet loss and delay performance of feedback and feed-forward arrayed-waveguid grating-based optical packet switches with WDM inputs-outputs," *IEEE/OSA J. Lightwave Technol.*, vol. 19, no. 9, pp. 1241-1254, Sept. 2001.

Bibliography

- [42] C. Guillemot *et al.*, "Transparent optical packet switching: The European ACTS KEOPS project approach," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 12, pp. 2117-2134, Dec. 1998.
- [43] F. Masetti *et al.*, "Design and performance of a broadcast-and-select photonic packet switching architecture," in *Proc. ECOC '96*, vol. 3, pp. 309-312, Sept. 1996.
- [44] F. Masetti, et al., "High speed, high capacity ATM optical switches for future telecommunication transparent networks," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 979-998, Jun. 1996.
- [45] S. Yoo, "Optical-packet switching and optical-label switching technologies for the next generation optical Internet," in *Proc. OFC '03*, vol. 2, pp. 797-798.
- [46] L. Dittmann *et al.*, "The European IST project DAVID: a viable approach toward optical packet switching," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1026-1040, Sept. 2003.
- [47] S. J. B. Yoo *et al.*, "High-performance optical-label switching packet routers and smart edge routers for the next-generation Internet," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1041-1051, Sept. 2003.
- [48] M. C. Chia, D. K. Hunter, I. Andonovic, P. Ball, and I. Wright, "Optical packet switches: A comparison of designs," in *Proc. IEEE ICON'00*, pp. 365-369.
- [49] M. Klinkowski and M. Marciniak, "Development of IP/WDM optical networks," in *Proc. LFNM'03*, pp. 84-87, May 2001.
- [50] J. H. Kelm, B. C. Wang, "Application of O-CDMA techniques to all optical packet header processing," in *Proc. IEEE LEOS 2003*, vol. 1, pp. 232-233.

Bibliography

- [51] K. Chan, F. Tong, C. K. Chan, L. K. Chen, W. Hung, "An all-optical packet header recognition scheme for self-routing packet switch networks," in *Proc. OFC '02*, pp. 797-798.
- [52] T. Sakamoto *et al.*, "Optical packet synchronizer using wavelength and space switching," *IEEE Photon. Technol. Lett.*, vol. 14, no. 9, Sept. 2002.
- [53] P. Homan and J. Bester, "Enhanced synchronous packet switching for IP packets," *IEICE Trans. Commun.*, vol. E85-B, no. 1, pp. 247-256, Jan. 2002.
- [54] A. E. Willner, "All-optical packet-header-recognition techniques," in *Proc. IEEE LEOS 2002*, vol.1, pp. 47-48.
- [55] M. C. Cardakli, S. Lee, A. E. Willner, V. Grubsky, D. Starodubov, and J. Feinberg, "Reconfigurable optical packet header recognition and routing using time-to-wavelength mapping and tunable fiber Bragg gratings for correlation decoding," *IEEE Photon. Technol. Lett.*, vol. 12, no. 5, pp. 552-554, May 2000.
- [56] J. M. Martinez, F. Ramos and J. Marti, "All-optical packet header processor based on cascaded SOA-MZIs," *Electronics Letters*, vol. 40, no. 14, pp. 894-895, July 2004.
- [57] 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.
- [58] Y. Liu, M. T. Hill, N. Calabretta, H. de Waardt, G. D. Khoe, and 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.

Bibliography

- [59] L. Li, S. D. Scott, and J. S. Deogun, "A novel fiber delay line buffering architecture for optical packet switching," in *Proc. IEEE GLOBECOM'03*, vol. 5, pp. 2809-2813.
- [60] W. D. Zhong and R. S. Tucker, "A new wavelength-routed photonic packet buffer combining traveling delay lines with delay-line loops," *IEEE/OSA J. Lightwave Technol.*, vol. 19, no. 8, pp. 1085-1092, Aug. 2001.
- [61] 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.
- [62] 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.
- [63] 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.
- [64] N. McKeown, "Scheduling algorithm for input-queued cell switches," Ph.D. thesis, University of California at Berkeley, 1995.
- [65] S. T. Chuang, A. Geol, N. Mckeown, and B. Prabhakar, "Matching output queuing with a combined input/output-queued switch," *IEEE J. Select. Areas Commun.*, vol. 17, no. 6, pp. 1030-1039, Jun. 1999.
- [66] B. Ramamurthy and B. Mukherjee, "Wavelength conversion in WDM networking," *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, pp. 1061-1073, Sept. 1998.

Bibliography

- [67] X. Qin and Y. Yang, "Nonblocking WDM switching networks with full and limited wavelength conversion," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2032-2041, Dec. 2002.
- [68] G. Shen, S. K. Bose, T. H. Cheng, C. Lu, and T. Y. Chai, "Operation of WDM networks with different wavelength conversion capabilities," *IEEE Commun. Lett.*, vol. 4, no. 7, pp. 239-241, Jul. 2000.
- [69] A. S. Sudbø and S. Bjørnstad, "Scalable optical switch structure based on tunable wavelength converters and arrayed waveguide grating routers," *OSA J. Optical Networking*, vol. 2, no.9, pp. 340-349, Sept. 2003.
- [70] M. J. Potasek, "All-optical switching for high bandwidth optical networks," *Optical Networks Mag.*, vol. 3, no. 11/12, pp. 30-43, Nov./Dec. 2002.
- [71] J. Ramamirtham, J. Turner, and J. Friedman, "Design of wavelength converting switches for optical burst switching," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1122-1132, Sept. 2003.
- [72] I. White, R. Pentty, Matthew Webster, Y. J. Chai, and A. Wonfor, "Wavelength switching components for future photonic networks," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 74-81, Sept.2002.
- [73] K. E. Stubkjaer *et al.*, "Wavelength converter technology," *IEICE Trans. Commun.*, vol. E82-B, no. 2, pp. 390-400, Feb. 1999.
- [74] H. J. Lee, M. Sohn, K. Kim, and H. G. Kim, "Wavelength dependent performance of a wavelength converter based on cross-gain modulation and birefringence of a semiconductor optical amplifier," *IEEE Photon. Technol. Lett.*, vol. 11, no. 2, pp. 185-187, Feb. 1999.

Bibliography

- [75] X. Zheng, F. Liu, and A. Kloch, "Experimental investigation of the cascability of a cross-gain modulation wavelength converter," *IEEE Photon. Technol. Lett.*, vol. 12, no. 3, pp. 272-274, Mar. 2000.
- [76] S. Nakamura, Y. Ueno, and K. Tajima, "168-Gb/s all-optical wavelength conversion with a symmetric-Mach-Zehnder-type switch," *IEEE Photon. Technol. Lett.*, vol. 13, no. 10, pp. 1091-1093, Oct. 2001.
- [77] J. H. Lee, Z. Yusoff, W. Belardi, M. Ibsen, T. M. Monro, and D. J. Richardson, "A tunable WDM wavelength converter based on cross-phase modulation effects in normal dispersion holey fiber," *IEEE Photon. Technol. Lett.*, vol. 15, no. 3, pp. 437-439, Mar. 2003.
- [78] B. Olsson, P. Öhlén, L. Rau, and D. J. Blumenthal, "A simple and robust 40Gb/s wavelength converter using fiber cross-phase modulation and optical filtering," *IEEE Photon. Technol. Lett.*, vol. 12, no. 7, pp. 846-848, Jul. 2000.
- [79] R. Sato, "Tuning technique to optimize input power of a cross-phase modulation wavelength converter," *IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 8, pp.1883-1892, Aug. 2004.
- [80] J. M. Yates, J. P. R. Lacey, M. P. Rumsewicz, and M. A. Summerfield, "Performance of Networks Using Wavelength Converters Based on Four-Wave Mixing in semiconductor optical amplifiers," *IEEE/OSA J. Lightwave Technol.*, vol. 17, no. 5, pp.782-791, May 1999.
- [81] D. Hsu, S. Lee, P. Gong, Y. Lin, S. S. W. Lee, and M. C. Yuang, "High-efficiency wide-band SOA-based wavelength converters by using dual-pumped four-wave mixing and an assist beam," *IEEE Photon. Technol. Lett.*, vol. 16, no. 8, pp. 1903-1905, Aug. 2004.

Bibliography

- [82] S. Mikroulis, A. Bogris, E. Roditi, and D. Syvridis, "Investigation of an all-optical wavelength converter with reshaping properties based on four-wave mixing in passive microring resonators," *IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 12, pp. 2743-2748, Dec. 2004.
- [83] C. Chan, K. L. Sherman, and M. Zirngibl, "A fast 100-channel wavelength-tunable transmitter for optical packet switching," *IEEE Photon. Technol. Lett.*, vol. 13, no. 7, pp. 729-731, Jul. 2001.
- [84] H. S. Park, K. Y. Song, S. H. Yun, and B. Y. Kim, "All-fiber wavelength-tunable acoustooptic switch based on intermodal coupling in fibers," *IEEE/OSA J. Lightwave Technol.*, vol. 20, no. 10, pp. 1864-1868, Oct. 2002.
- [85] K. K. Chow, C. Shu, M. W. K. Mak, and H. K. Tsang, "Widely tunable wavelength converter using a double-ring fiber laser with a semiconductor optical amplifier," *IEEE Photon. Technol. Lett.*, vol. 14, no. 10, pp. 1445-1447, Oct. 2002.
- [86] G. Shen, T. H. Cheng, S. K. Bose, C. Lu, and T. Y. Chai, "The impact of number of transceivers and their tunabilities on WDM network performance," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 366-368, Nov. 2000.
- [87] G. Castanon, L. Tancevski, and L. Tamil, "Routing in all-optical packet switched irregular mesh networks," in *Proc. IEEE Globecom'99*, pp. 1017-1022.
- [88] S. L. Danielsen, C. Joergensen, B. Mikkelsen, and K. E. Stubkjaer, "Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tuneable wavelength converters," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 5, pp. 729-735, May 1998.

Bibliography

- [89] S. L. Danielsen, C. Joergensen, B. Mikkelsen, and K. E. Stubkjaer, "Optical packet switched network layer without optical buffers," *IEEE Photon. Technol. Lett.*, vol.10, no. 6, pp. 896-898, Jun. 1998.
- [90] F. Callegati, G. Corazza, and C. Raffaelli, "Exploitation of DWDM for optical packet switching with quality of service guarantees," *IEEE J. Select. Areas Comm.*, vol. 20, no. 1, pp. 190-201, Jan. 2002.
- [91] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *IEEE/OSA J. Lightwave Technol.*, vol. 21, no. 3, pp. 672-683, Mar. 2003.
- [92] S. J. B. Yoo *et al.*, "High-performance optical-label switching packet routers and smart edge routers for the next-generation Internet," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1041-1051, Sept. 2003.
- [93] N. McKeown, "The iSLIP scheduling algorithm for input-queued switches," *IEEE/ACM Trans. Networking.*, vol. 7, no. 2, pp.188-201, Apr. 1999.
- [94] C. Minkenberg, "Integrating unicast and multicast traffic scheduling in a combined input- and output-queued packet switching system," in *Proc. ICCCN* 2000, pp. 127-134.
- [95] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1260-1267, Aug. 1999.
- [96] M. A. Marsan, A. Bianco, E. Leonardi, and L. Milia, "RPA: A flexible scheduling algorithm for input buffered switches," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1921-1933, Dec. 1999.

Bibliography

- [97] M. A. Marsan, A. Bianco, P. Giaccone, E. Leonardi, and F. Neri, "Packet scheduling in input-queued cell-based switches," in *IEEE INFOCOM'01*, pp. 1085-1094.
- [98] B. Bensaou, D. H. K. Tsang, and K. T. Chan, "Credit-based fair queueing (CBFQQ): a simple service-scheduling algorithm for packet-switched networks," *IEEE/ACM Trans. Networking.*, vol. 9, no. 5, pp.591-604, Oct. 2001.
- [99] F. M. Chiussi and A. Francini, "A distributed scheduling architecture for scalable packet switches," *IEEE J. Select. Areas Comm.*, vol. 18, no. 12, pp. 2665-2683, Dec. 2000.
- [100] I. Elhanany and D. Sadot, "Analysis of non-uniform cell destination distribution in virtual output queueing systems," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 367-369, Sept. 2002.
- [101] R. O. LaMaire and D. N. Serpanos, "Two-dimensional round-robin schedulers for packet switches and multiple input queues," *IEEE/ACM Trans. Networking.*, vol. 2, no. 5, pp.471-482, Oct. 1994.
- [102] M. A. Bonuccelli, I. Gopal and C. K. Wong, "Incremental time-slot assignment in SS/TDMA satellite systems," *IEEE Trans. Commun.*, vol. 39, no. 7, pp. 1147-1156, Jul. 1991.
- [103] T. Anderson, S. Owicki, J. Saxe, and C. Thacker, "High speed switch scheduling for local area networks," *ACM Trans. Comput. Syst.* vol. 11, no. 4, pp. 319-352, Nov. 1993.
- [104] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services network," in *Proc. IEEE INFOCOM'02*, pp. 915-924, 1992.

Bibliography

- [105] B. Li and Y. Qin, "Traffic scheduling in a photonic packet switching system with QoS guarantee," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 12, pp. 2281-2295, Dec. 1998.
- [106] S. Y. Liew, G. Hu and H. J. Chao, "Scheduling algorithm for shared fiber-delay-line optical packet switches-Part I: The single-stage case," *IEEE/OSA J. Lightwave Technol.*, vol. 23, no.4, pp. 1586-1600, Apr. 2005.
- [107] S. Jiang, G. Hu, S. Y. Liew and H. J. Chao, "Scheduling algorithm for shared fiber-delay-line optical packet switches-Part II: The three-stage clos-network case," *IEEE/OSA J. Lightwave Technol.*, vol. 23, no.4, pp. 1601-1609, Apr. 2005.
- [108] M. J. Karol, "Shared-memory optical packet (ATM) switch," *SPIE* vol. 2024, *Multigigabit Fiber Communications Systems*, pp. 212-222, 1993.
- [109] J. Diao and P. L. Chu, "Analysis of partially shared buffering for WDM optical packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 17, no. 12, pp. 2461-2469, Dec. 1999.
- [110] F. M. Chiussi, Y. Xia, and V. P. Kumar, "Performance of shared-memory switches under multicast bursty traffic," *IEEE J. Select. Areas Comm.*, vol. 15, no. 3, pp. 473-487, Apr. 1997.
- [111] L. Tancevski, A. Ge, and G. Castanon, "Optical packet switch with partially shared buffers: Design principles," in *Proc. OFC'01*, vol. 2, TuK3-1 - TuK3-3.
- [112] M. R. N. Ribeiro and M. J. O'Mahony, "Improvement on performance of photonic packet switching nodes by priority assignment and buffer sharing," in *Proc. IEEE ICC'00*, vol. 3, pp. 1738-1742.

Bibliography

- [113] X. Li and M. Hamdi, "On scheduling optical packet switches with reconfiguration delay," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1156-1164, Sept. 2003.
- [114] F. Callegati, W. Cerroni, G. Corazza, C. Develder, M. Pickavet, P. Demeester, "Scheduling algorithms for a slotted packet switch with either fixed or variable length packets," *Photonic Networks Commun.*, pp. 163-176, vol. 8, no. 2, Sep. 2004.
- [115] H. J. Chao, Z. Jing, and K. Deng, "Packet scheduling scheme for a 3-stage crossbar network photonic switch," in *Proc. IEEE ICC '03*, vol. 2, pp. 1293-1298.
- [116] V. Sharma and E. A. Varvarigos, "Limited wavelength translation in all-optical WDM mesh networks," in *Proc. IEEE INFOCOM'98*, vol. 2, pp. 893-901.
- [117] R. Sato, "Tuning technique to optimize input power of a cross-phase modulation wavelength converter," *IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 8, pp. 1883-1892, Aug. 2004.
- [118] N. Nagatsu, A. Watanabe, S. Okamoto, and K. Sato, "Performance and node architecture of WDM multiple fiber ring networks," in *Proc. IEEE ICC'98*, pp. 931-936.
- [119] X. Cao, V. Anand, Y. Xiong, and C. Qiao, "Performance evaluation of wavelength band switching in multi-fiber all-optical networks," in *Proc. IEEE INFOCOM'03*, vol. 3, pp. 2251-2261, 2003.
- [120] Y. Leung, G. Xiao, and K. Hung, "Design of node configuration for all-optical multi-fiber networks," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 135-145, Jan. 2002.

Bibliography

- [121] L. Li and A. K. Somani, "A new analytical model for multifiber WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 2138-2145, Oct. 2000.
- [122] Y. Luo and N. Ansari, "Performance evaluation of survivable multifiber WDM networks," in *Proc. IEEE Globecom'03*, pp. 2524-2528.
- [123] A. Odlyzko, "Internet growth: myth and reality, use and abuse," *Information Impacts Magazine*, Nov. 2000.
- [124] O. Gerstel and H. Raza, "Merits of low-density WDM line system for long-haul networks," *IEEE/OSA J. Lightwave Technol.*, vol. 21, no. 11, pp. 2470-2475, Nov. 2003.
- [125] M. Jonsson, "Optical interconnection technology in switches, routers, and optical cross connects," *Optical Network Mag.*, vol. 4, no. 7/8, pp. 20-34, July/August 2003.
- [126] V. Eramo, M. Listanti, and M. Spaziani, "Resources sharing in optical packet switches with limited-range wavelength converters," *IEEE/OSA J. Lightwave Technol.*, vol. 23, no. 2, pp. 671-687, Feb. 2005.
- [127] A. E. Eckberg and T. C. Hou, "Effect of output buffer sharing on buffer requirements in an ATDM packet switch," in *Proc. IEEE INFOCOM'88*, pp. 459-466, Mar. 1988.
- [128] M. A. Marsan, A. Fumagalli, E. Leonardi, F. Neri, and P. Poggiolini, "Daisy: a scalable all-optical packet network with multifiber ring topology," *Comput. Networks and ISDN Systems*, vol. 30, pp. 1065-1082, 1998.

Bibliography

- [129] T. Gipsper, H. A. Jager and L. Rapp, "Broadcasting, scalability, and reconfigurability aspects in an all-optical network architecture," *Fiber and Integrated Optics*, vol. 17, no. 1, pp.21-40, Jan. 1998.
- [130] Y. Li, G. Xiao, and H. Ghafouri-Shiraz, "On the benefits of multifiber optical packet switch," *Microwave and Optical Technology Letters*, vol. 43, no. 5, pp. 376-378, Dec. 2004.
- [131] S. Yao, F. Xue, B. Mukherjee, S. J. B. Yoo, and S. Dixt, "Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 66-72, Sep. 2002.
- [132] W. Yan, E. Hou, and N. Ansari, "Anomaly detection and traffic shaping under self-similar aggregated traffic in optical switched networks," in *Proc. ICCT'03*, vol. 1, pp. 378-381.
- [133] G. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, 2nd ed., Cambridge University Press, 1952.